Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Multivariate Analysis 95 (2005) 345–369
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

A multivariate empirical characteristic function test of independence with normal marginals

M. Bilodeau[a,*], P. Lafaye de Micheaux[b]

[a]*Département de Mathématiques et de Statistique, Centre de Recherches Mathématiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal Que., Canada H3C 3J7*
[b]*Université Pierre Mendès France, Laboratoire de Statistique et Analyse de Données (LabSAD), BP 47/F-38040 Grenoble Cedex 9, France*

Received 30 December 2002

Abstract

This paper proposes a semi-parametric test of independence (or serial independence) between marginal vectors each of which is normally distributed but without assuming the joint normality of these marginal vectors. The test statistic is a Cramér–von Mises functional of a process defined from the empirical characteristic function. This process is defined similarly as the process of Ghoudi et al. [J. Multivariate Anal. 79 (2001) 191] built from the empirical distribution function and used to test for independence between univariate marginal variables. The test statistic can be represented as a *V*-statistic. It is consistent to detect any form of dependence. The weak convergence of the process is derived. The asymptotic distribution of the Cramér–von Mises functionals is approximated by the Cornish–Fisher expansion using a recursive formula for cumulants and inversion of the characteristic function with numerical evaluation of the eigenvalues. The test statistic is finally compared with Wilks statistic for testing the parametric hypothesis of independence in the one-way MANOVA model with random effects.

AMS 1991 subject classification: 62H15; 62M99; 62E20; 60F05

Keywords: Characteristic function; Independence; Multivariate analysis; Serial independence; Stochastic processes

* Corresponding author.
E-mail addresses: bilodeau@dms.umontreal.ca (M. Bilodeau), pierre.lafaye-de-micheaux@upmf-grenoble.fr (P. Lafaye de Micheaux).

0047-259X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2004.08.011

1. Introduction

Different characterizations have led to various tests of independence. Let $p \geqslant 1$ be a fixed integer. Consider a partitioned random vector $\boldsymbol{\epsilon} = \left(\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}\right)$ made up of p q-dimensional subvectors and a corresponding partitioned $\boldsymbol{t} = \left(\boldsymbol{t}^{(1)}, \ldots, \boldsymbol{t}^{(p)}\right)$, for any fixed vector $\boldsymbol{t}$. Independence of the subvectors may be characterized with the joint distribution function or characteristic function as

$$K_p(\boldsymbol{t}) = \prod_{k=1}^{p} K^{(k)}(\boldsymbol{t}^{(k)}), \tag{1.1}$$

$$C_p(\boldsymbol{t}) = \prod_{k=1}^{p} C^{(k)}(\boldsymbol{t}^{(k)}), \tag{1.2}$$

where K_p and C_p are, respectively, the joint distribution function and joint characteristic function. The marginal versions are $K^{(k)}$ and $C^{(k)}$ for $k = 1, \ldots, p$. In the univariate setting ($q = 1$) Blum et al. [4] proposed an empirical process based on (1.1), whereas Csörgő [10] defined a similar process based on (1.2). Feuerverger [19] proposed an empirical characteristic function version of the Blum et al. [4] test statistic. He pointed out difficulties with dimensions above 2.

Recently, in the univariate setting, Ghoudi et al. [21] introduced a new process based on a characterization of independence which is now presented. This characterization for $p = 3$ is implicit in the early paper of Blum et al. [4]. For any $A \subset I_p = \{1, \ldots, p\}$ and any $\boldsymbol{t} \in \mathbb{R}^p$, let

$$\mu_A(\boldsymbol{t}) = \sum_{B \subset A} (-1)^{|A \backslash B|} K_p(\boldsymbol{t}_B) \prod_{j \in A \backslash B} K^{(j)}(t^{(j)}).$$

The notation $|A|$ stands for cardinality of the set A and the convention $\prod_{\emptyset} = 1$ is adopted. The vector $\boldsymbol{t}_B$ is used to make a selection of components of $\boldsymbol{t}$ according to the set B,

$$(\boldsymbol{t}_B)^{(i)} = \begin{cases} t^{(i)}, & i \in B; \\ \infty, & i \in I_p \backslash B. \end{cases}$$

Independence can be characterized as follows: $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$ are independent if and only if $\mu_A \equiv 0$, for all $A \subset I_p$ satisfying $|A| > 1$. This characterization was also given previously in a slightly different form in Deheuvels [15]. A Cramér–von Mises functional of each process in a decomposition of the empirical dependence process, obtained originally by Deheuvels [14], led them to a non-parametric test of independence in the non-serial and serial situations. The interest of this decomposition resides in the mutual asymptotic independence of all the processes in the decomposition and the simple form of the covariance which is expressed as a product of covariance functions of the Brownian bridge.

In the multivariate setting ($q \geqslant 1$), the present paper proposes tests of independence, built from a process relying on a similar independence characterization based on characteristic functions, when subvectors or marginals are normally distributed. Namely, the marginals

$\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$ are independent if and only if $v_A \equiv 0$, for all $A \subset I_p$, $|A| > 1$, where

$$v_A(\boldsymbol{t}) = \sum_{B \subset A} (-1)^{|A \setminus B|} C_p(\boldsymbol{t}^B) \prod_{j \in A \setminus B} C^{(j)}(\boldsymbol{t}^{(j)})$$

and where the selection of subvectors is modified to

$$(\boldsymbol{t}^B)^{(i)} = \begin{cases} \boldsymbol{t}^{(i)}, & i \in B; \\ \boldsymbol{0}, & i \in I_p \setminus B. \end{cases}$$

The proof of this characterization of independence is by mathematical induction much like the proof of Proposition 2.1 using μ_A in Ghoudi et al. [21].

Note that the subvectors are not assumed to be jointly multinormal in which case independence can be tested parametrically with covariances using likelihood ratio tests. Assuming a normal distribution for the marginals is convenient since it allows a numerical computation of the eigenvalues of the integral operator in Section 4. A similar assumption was made by Deheuvels [16] who derived some tests of independence for a bivariate distribution with exponential marginals. When the distribution of the marginals is unknown, the asymptotic covariance function of the processes considered in Sections 2 and 3 still has the same product structure but it contains the unknown characteristic function of the marginals. As an alternative construction, the characteristic functions could be replaced by half-space probabilities, as in Beran and Millar [2], which reduce to distribution functions in the univariate case. This would provide a generalization of Ghoudi et al. [21] to the multivariate case. The asymptotic distribution of the processes would again depend on the unknown probability distribution of the marginals. Validity of the bootstrap in this situation is under investigation.

The non-serial problem of testing independence between normally distributed subvectors of a random vector and the serial problem of testing independence between normally distributed observations of a multivariate stationary sequence are considered. It is shown that the asymptotic distribution of the proposed process is the same in both cases under the null hypothesis of independence. Moreover, it is established that the estimation of the unknown mean vector and positive definite covariance matrix of the normal marginals does not affect the asymptotic distribution of the process. The proposed Cramér–von Mises type of test statistic is related to *V*-statistics for which de Wet and Randles [13] studied the effect of estimating the unknown parameters.

In matters of notation, convergence in distribution of a stochastic process, or of its functionals, will be denoted by the symbol $\xrightarrow{D}$. All the proofs are deferred to Section 6.

2. Testing independence: the non-serial situation

2.1. The case of known parameters

Let $\boldsymbol{\epsilon} = \left(\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}\right) \in \mathbb{R}^{pq}$ denote a partition into p q-dimensional subvectors and let $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ be an i.i.d. sample of such (pq)-dimensional random vectors. Suppose that the subvectors of the random vectors $\boldsymbol{\epsilon}_i$ all have the same $N_q(\boldsymbol{0}, \boldsymbol{I})$ normal distribution, with characteristic function ϕ. The problem is that of testing the independence of the marginals, that is, the independence of $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$. This non-serial problem with known parameters is of very limited practical importance. However, it serves as a prototype on which subsequent results are based. Unlike the serial problem in Section 3 where the components of $\boldsymbol{\epsilon}$ are necessarily of the same dimension, the non-serial problem with marginals $\boldsymbol{\epsilon}^{(k)}$'s of different dimensions is also a problem of interest. Its solution in the case of known parameters would not require any modification. In the case of unknown parameters, their estimation would have to be adapted to this new situation. The marginals of equal dimension in the non-serial problem is also a sufficient prototype to solve the serial problem.

With this aim, for any $A \subset I_p$ and any $\boldsymbol{t} = (\boldsymbol{t}^{(1)}, \ldots, \boldsymbol{t}^{(p)}) \in \mathbb{R}^{pq}$, let

$$R_{n,A}(\boldsymbol{t}) = \sqrt{n} \sum_{B \subset A} (-1)^{|A \setminus B|} \phi_{n,p}(\boldsymbol{t}^B) \prod_{i \in A \setminus B} \phi(\boldsymbol{t}^{(i)}), \tag{2.1}$$

where

$$\phi_{n,p}(\boldsymbol{t}) = \frac{1}{n} \sum_{j=1}^{n} \exp(\mathrm{i}\langle \boldsymbol{t}, \boldsymbol{\epsilon}_j \rangle)$$

is the empirical characteristic function of the sample. The notation $\langle \cdot, \cdot \rangle$ is for the usual inner product between vectors. The process $R_{n,A}$ is the empirical version of the quantity v_A on which the independence characterization is based. In fact, it follows readily that $E[R_{n,A}(\boldsymbol{t})/\sqrt{n}] = v_A(\boldsymbol{t})$.

The asymptotic behavior of these processes is stated next. It is shown that for different subsets A the associated processes are asymptotically independent, each process being asymptotically Gaussian with a covariance function of a particularly simple form. Specifically, the covariance function is a product of covariance functions of the type encountered by Feuerverger and Mureika [20], Csörgő [8] or Marcus [29], for the empirical characteristic function process. This implies in Section 4 an advantageous product representation of eigenvalues which are the solutions of an integral equation over the lower-dimensional space $\mathbb{R}^q$ instead of $\mathbb{R}^{pq}$. Another process defined by Csörgő [10] has a covariance of a more complicated structure. Some notations are now in order. The norm of vectors or matrices, real or complex, will be denoted $\|\cdot\|$, whereas $|\cdot|$ will be the norm of real or complex scalars. The conjugate of $u \in \mathbb{C}$ will be $c(u)$. We also let $C(\mathbb{R}^{pq}, \mathbb{C})$ be the space of continuous functions from $\mathbb{R}^{pq}$ to $\mathbb{C}$.

Theorem 2.1. *If $\boldsymbol{\epsilon}_1^{(1)}, \ldots, \boldsymbol{\epsilon}_1^{(p)}$ are independent,the collection of processes $\{R_{n,A} : |A| > 1\}$ converge in $C(\mathbb{R}^{pq}, \mathbb{C})$ to independent zero mean complex Gaussian processes $\{R_A : |A| > 1\}$ having covariance function given by*

$$C_A(\boldsymbol{s}, \boldsymbol{t}) = E[R_A(\boldsymbol{s}) c(R_A(\boldsymbol{t}))] = \prod_{k \in A} [\phi(\boldsymbol{t}^{(k)} - \boldsymbol{s}^{(k)}) - \phi(\boldsymbol{t}^{(k)})\phi(\boldsymbol{s}^{(k)})] \tag{2.2}$$

and pseudo-covariance function given by

$$P_A(\boldsymbol{s}, \boldsymbol{t}) = E[R_A(\boldsymbol{s}) R_A(\boldsymbol{t})] = C_A(-\boldsymbol{s}, \boldsymbol{t}). \tag{2.3}$$

It is interesting to note that the multinomial formula in Proposition 6.1 of Ghoudi et al. [21] yields the equivalent representation

$$R_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \prod_{k \in A} [\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)} \rangle) - \phi(\boldsymbol{t}^{(k)})]. \tag{2.4}$$

This i.i.d. average representation is used in the proof of Theorem 2.1.

2.2. The case of unknown parameters

The context is the same as in the preceding subsection except that the components of the random vectors $\boldsymbol{\epsilon}_i$ in the sample now all have the same $N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ normal distribution, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, positive definite, are unknown. The problem again is that of testing for independence of the marginals, that is, the independence of the components $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$.

First, define the standardized residual vectors $\boldsymbol{e}_j^{(k)} = \boldsymbol{S}^{-\frac{1}{2}}(\boldsymbol{\epsilon}_j^{(k)} - \overline{\boldsymbol{\epsilon}})$, where $\boldsymbol{S} = \frac{1}{np}\sum_{j=1}^{n}\sum_{k=1}^{p}(\boldsymbol{\epsilon}_j^{(k)} - \overline{\boldsymbol{\epsilon}})(\boldsymbol{\epsilon}_j^{(k)} - \overline{\boldsymbol{\epsilon}})^{\mathrm{T}}$ and $\overline{\boldsymbol{\epsilon}} = \frac{1}{np}\sum_{j=1}^{n}\sum_{k=1}^{p}\boldsymbol{\epsilon}_j^{(k)}$ are, respectively, the sample covariance matrix and the sample mean. Also, let $\overline{\boldsymbol{\epsilon}}^{(k)} = \frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\epsilon}_j^{(k)}$ be the sample mean of the kth subvector, $k = 1, \ldots, p$.

The underlying process is the same as the one considered in Section 2.1, apart from the unknown parameters which are replaced by their sample estimates. The plug-in process is thus

$$\hat{R}_{n,A}(\boldsymbol{t}) = \sqrt{n} \sum_{B \subset A} (-1)^{|A \setminus B|} \hat{\phi}_{n,p}(\boldsymbol{t}^B) \prod_{i \in A \setminus B} \phi(\boldsymbol{t}^{(i)}), \tag{2.5}$$

where

$$\hat{\phi}_{n,p}(\boldsymbol{t}) = \frac{1}{n} \sum_{j=1}^{n} \exp(\mathrm{i}\langle \boldsymbol{t}, \boldsymbol{e}_j \rangle)$$

is the empirical characteristic function based on the standardized residuals $\boldsymbol{e}_j = (\boldsymbol{e}_j^{(1)}, \ldots, \boldsymbol{e}_j^{(p)}) \in \mathbb{R}^{pq}$. Attention should be brought to the unusual fact that the second ϕ function in (2.1) and (2.5) does not have a subscript n since the distribution of the marginals is known up to some unknown parameters. The asymptotic behavior of these processes is stated next, the main conclusion being that the estimation of the unknown parameters does not affect the asymptotic distribution.

Theorem 2.2. *If $\boldsymbol{\epsilon}_1^{(1)}, \ldots, \boldsymbol{\epsilon}_1^{(p)}$ are independent, the processes $\{\hat{R}_{n,A} : |A| > 1\}$ converge in $C(\mathbb{R}^{pq}, \mathbb{C})$ to independent zero mean complex Gaussian processes $\{R_A : |A| > 1\}$ having covariance and pseudo-covariance functions, respectively, given by $C_A(\boldsymbol{s}, \boldsymbol{t})$ and $P_A(\boldsymbol{s}, \boldsymbol{t})$ in* (2.2) *and* (2.3).

The same multinomial formula of Ghoudi et al. [21] yields the representation

$$\hat{R}_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \prod_{k \in A} [\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{e}_j^{(k)} \rangle) - \phi(\boldsymbol{t}^{(k)})].$$

The Cramér–von Mises test statistic proposed is $nT_{n,b,A}$, where for a given subset A

$$T_{n,b,A} = \frac{1}{n} \int |\hat{R}_{n,A}(\boldsymbol{t})|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t} \tag{2.6}$$

and where the weight function φ_b is the $N_{pq}(\boldsymbol{0}, b^2\boldsymbol{I})$ density which permits the mathematical integration. The multinomial representation and this appropriate weighting allow this test statistic to be computed explicitly as

$$\begin{aligned} T_{n,b,A} = \frac{1}{n^2} \sum_{l=1}^{n} \sum_{l'=1}^{n} \prod_{k \in A} \Bigg\{ & \exp\left[-\frac{b^2}{2} \|\boldsymbol{e}_l^{(k)} - \boldsymbol{e}_{l'}^{(k)}\|^2 \right] \\ & - (b^2+1)^{-\frac{q}{2}} \exp\left[-\frac{1}{2} \frac{b^2}{b^2+1} \|\boldsymbol{e}_l^{(k)}\|^2 \right] \\ & - (b^2+1)^{-\frac{q}{2}} \exp\left[-\frac{1}{2} \frac{b^2}{b^2+1} \|\boldsymbol{e}_{l'}^{(k)}\|^2 \right] + (2b^2+1)^{-\frac{q}{2}} \Bigg\}. \end{aligned}$$

Since squared Mahalanobis-type statistics are affine invariant it follows that $T_{n,b,A}$ is affine invariant. Thus, the asymptotic distribution of this statistic does not depend on the unknown parameters.

It should be noted that the functional (2.6) defining this test statistic is not continuous; it is not even defined on $C(\mathbb{R}^{pq}, \mathbb{C})$ but only on the subset of squared-integrable functions with respect to the measure $\varphi_b(\boldsymbol{t})\, d\boldsymbol{t}$. Thus, the continuous mapping theorem as in Billingsley [3, p. 31] cannot be invoked. In order to obtain the asymptotic distribution of this functional, the following generalization of Theorem 3.3 of Kellermeier [26] on a uniform integrability condition is proposed. Let $\mathcal{B}_j^{pq}$ be the ball of radius j centered at zero in $\mathbb{R}^{pq}$.

Theorem 2.3. *Let $\boldsymbol{y}_n$ and $\boldsymbol{y}$ be random elements of $C(\mathbb{R}^{pq}, \mathbb{C})^2$ such that $\boldsymbol{y}_n \xrightarrow{D} \boldsymbol{y}$ on all compact balls. Let $f : \mathbb{C}^2 \to \mathbb{R}$ be a continuous function and let G be a probability measure on $\mathbb{R}^{pq}$. Define $w_n = \int f(\boldsymbol{y}_n(\boldsymbol{t}))\, dG(\boldsymbol{t})$ and $w = \int f(\boldsymbol{y}(\boldsymbol{t}))\, dG(\boldsymbol{t})$. Suppose that w_n and w are well defined with probability one. Moreover, suppose that there exists $\alpha \geqslant 1$ such that*

$$\lim_{j \to \infty} \limsup_{n \to \infty} \int_{\mathbb{R}^{pq} \setminus \mathcal{B}_j^{pq}} E|f(\boldsymbol{y}_n(\boldsymbol{t}))|^{\alpha}\, dG(\boldsymbol{t}) = 0. \tag{2.7}$$

Then, $w_n \xrightarrow{D} w$ as $n \to \infty$.

Using Theorem 2.3, the joint convergence of the Cramér–von Mises functionals can be established.

Theorem 2.4. *For two different subsets A and B of I_p,*

$$\int \left(|\hat{R}_{n,A}(\boldsymbol{t})|^2, |\hat{R}_{n,B}(\boldsymbol{t})|^2\right) \varphi_b(\boldsymbol{t})\, d\boldsymbol{t} \xrightarrow{D} \int \left(|R_A(\boldsymbol{t})|^2, |R_B(\boldsymbol{t})|^2\right) \varphi_b(\boldsymbol{t})\, d\boldsymbol{t},$$

where integrals are computed componentwise.

All possible subsets A can then be simultaneously accounted for by combining the test statistics as in

$$S_n = n \sum_{|A|>1} T_{n,b,A} \tag{2.8}$$

or

$$M_n = n \max_{|A|>1} T_{n,b,A}. \tag{2.9}$$

2.3. Relation to V-statistics

The statistic $T_{n,b,A}$ is in fact a V-statistic as in de Wet and Randles [13]. It can be represented as

$$T_{n,b,A} = \frac{1}{n^2} \sum_{l=1}^{n} \sum_{l'=1}^{n} h(\boldsymbol{\epsilon}_l, \boldsymbol{\epsilon}_{l'}; \hat{\boldsymbol{\lambda}}_n), \tag{2.10}$$

where $\hat{\boldsymbol{\lambda}}_n = (\bar{\boldsymbol{\epsilon}}, \boldsymbol{S})$ consistently estimates the true parameter $\boldsymbol{\lambda} = (\boldsymbol{0}, \boldsymbol{I})$. The function h at an arbitrary $\boldsymbol{\gamma} = (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is defined as

$$h(\boldsymbol{\epsilon}_l, \boldsymbol{\epsilon}_{l'}; \gamma) = \int g(\boldsymbol{\epsilon}_l, \boldsymbol{t}; \gamma) g(\boldsymbol{\epsilon}_{l'}, \boldsymbol{t}; \gamma) \varphi_b(\boldsymbol{t})\, d\boldsymbol{t},$$

where, from elementary properties of integrals of odd functions, the function g can be taken real-valued

$$\begin{aligned} g(\boldsymbol{\epsilon}_l, \boldsymbol{t}; \gamma) = \prod_{k \in A} \Big[& \cos(\langle \boldsymbol{t}^{(k)}, \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\epsilon}_l^{(k)} - \boldsymbol{\mu})\rangle) \\ & + \sin(\langle \boldsymbol{t}^{(k)}, \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\epsilon}_l^{(k)} - \boldsymbol{\mu})\rangle) - \phi(\boldsymbol{t}^{(k)}) \Big]. \end{aligned}$$

Letting $\mu(\boldsymbol{t}; \gamma) = E_{(\boldsymbol{0},\boldsymbol{I})} g(\boldsymbol{\epsilon}_l, \boldsymbol{t}; \gamma)$, it is seen that $T_{n,b,A}$ is a V-statistic which falls into case I situation in de Wet and Randles [13]. These authors refer to case I when all first-order partial derivatives of $\mu(\boldsymbol{t}; \boldsymbol{\gamma})$ evaluated at the true parameter $\boldsymbol{\gamma} = \boldsymbol{\lambda}$ vanish. Otherwise, they refer to case II. This is case I here since only A's such that $|A| > 1$ are considered. Thus, the asymptotic distribution of $T_{n,b,A}$ is the same whether one uses $\hat{\boldsymbol{\lambda}}_n$ or $\boldsymbol{\lambda}$ in (2.10). It is not clear, however, how this argument would apply to the joint distribution of $T_{n,b,A}$ and $T_{n,b,B}$. The proof of Theorem 2.4 does not use de Wet and Randles [13].

For subsets A, $|A| = 1$, the statistic $T_{n,b,A}$ reduces to the statistic used by Baringhaus and Henze [1] and Henze and Zirkler [23] to test normality of a given marginal. They showed that the asymptotic distribution is affected by the estimation of the unknown parameters by establishing case II of de Wet and Randles [13]. Henze and Wagner [22] treated the same problem with an approach based on stochastic processes.

2.4. Consistency

Consider the alternatives whereby $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$ are distributed as $N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, but are not independent. Then, $S_n \overset{\text{a.s.}}{\to} \infty$ (almost surely) and $M_n \overset{\text{a.s.}}{\to} \infty$. Thus, the test statistics S_n and M_n in (2.8) and (2.9) are consistent against any such alternatives.

The argument to establish consistency is rather trivial as in Baringhaus and Henze [1]. Recall that $C_p(\cdot)$ is the joint characteristic function of $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$. This argument consists of the following almost sure convergence:

$$\begin{aligned} T_{n,b,A} &= \int \Bigg| \sum_{B \subset A} (-1)^{|A \setminus B|} \exp\left(-\mathrm{i} \sum_{k \in B} \langle \boldsymbol{t}^{(k)}, \boldsymbol{S}^{-\frac{1}{2}} \overline{\boldsymbol{\epsilon}} \rangle \right) \\ &\quad \times \frac{1}{n} \sum_{j=1}^{n} \exp\left(\mathrm{i} \sum_{k \in B} \langle \boldsymbol{t}^{(k)}, \boldsymbol{S}^{-\frac{1}{2}} \boldsymbol{\epsilon}_j^{(k)} \rangle \right) \prod_{i \in A \setminus B} \phi(\boldsymbol{t}^{(i)}) \Bigg|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t} \\ &\overset{\text{a.s.}}{\to} \int \Bigg| \sum_{B \subset A} (-1)^{|A \setminus B|} \exp\left(-\mathrm{i} \sum_{k \in B} \langle \boldsymbol{t}^{(k)}, \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{\mu} \rangle \right) \\ &\quad \times C_p((\boldsymbol{I}_p \otimes \boldsymbol{\Sigma}^{-\frac{1}{2}}) \boldsymbol{t}^B) \prod_{i \in A \setminus B} \phi(\boldsymbol{t}^{(i)}) \Bigg|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t} \end{aligned}$$

which equals 0 for all A, $|A| > 1$, if and only if $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$ are independent $N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Therefore, if $\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(p)}$ are dependent $N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then, there exists an A such that $nT_{n,b,A} \overset{\text{a.s.}}{\to} \infty$ which suffices to have $S_n \overset{\text{a.s.}}{\to} \infty$ and $M_n \overset{\text{a.s.}}{\to} \infty$.

3. Testing independence: the serial situation

3.1. The case of known parameters

Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$ be a stationary sequence of random vectors $\boldsymbol{u}_i$ distributed as $N_q(\boldsymbol{0}, \boldsymbol{I})$. It is desired to verify whether the $\boldsymbol{u}_i$'s are independent. For a given p, introduce the partitioned random vectors $\boldsymbol{\epsilon}_i = (\boldsymbol{u}_i, \ldots, \boldsymbol{u}_{i+p-1}) \in \mathbb{R}^{pq}$, $i = 1, \ldots, n - p + 1$. Also, let $S_{n,A}(\boldsymbol{t})$ be as in (2.1) with the slight modification $\phi_{n,p}(\boldsymbol{t}) = \frac{1}{n} \sum_{j=1}^{n-p+1} \exp(\mathrm{i}\langle \boldsymbol{t}, \boldsymbol{\epsilon}_j \rangle)$. In practice a selection of p must be made. For example, if the sequence is dependent at lag k, i.e. $\boldsymbol{u}_i$ and $\boldsymbol{u}_{i+k}$ are dependent, the selection $p = k+1$ (or an even larger value of p) should eventually, for n sufficiently large, detect this dependence with the subset $A = \{1, k+1\}$.

The main result related to the asymptotic distribution is that the m-dependence (the value of m is $p - 1$) introduced by the overlapping of the $\boldsymbol{u}_i$'s does not affect the asymptotic distribution. It is the same as in the non-serial case.

Theorem 3.1. *If the $\boldsymbol{u}_i$'s are independent, the collection of processes $\{S_{n,A} : |A| > 1\}$ converge in $C(\mathbb{R}^{pq}, \mathbb{C})$ to independent zero mean complex Gaussian processes $\{R_A : |A| > 1\}$ having covariance and pseudo-covariance functions, respectively, given by $C_A(\boldsymbol{s}, \boldsymbol{t})$ and $P_A(\boldsymbol{s}, \boldsymbol{t})$ in* (2.2) *and* (2.3).

As in (2.4), the multinomial formula of Ghoudi et al. [21] yields

$$S_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p+1} \prod_{k \in A} [\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)} \rangle) - \phi(\boldsymbol{t}^{(k)})]$$

which is useful in the proof of Theorem 3.1.

3.2. The case of unknown parameters

The context is the same as in the preceding section, but here the $\boldsymbol{u}_i$'s all have the same $N_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ normal distribution, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, positive definite, are assumed unknown. Again, we want to test whether the $\boldsymbol{u}_i$'s are independent. To this aim, define the random vectors $\boldsymbol{\epsilon}_i = (\boldsymbol{u}_i, \ldots, \boldsymbol{u}_{i+p-1}) \in \mathbb{R}^{pq}$ and $\boldsymbol{e}_i = (\hat{\boldsymbol{u}}_i, \ldots, \hat{\boldsymbol{u}}_{i+p-1}) \in \mathbb{R}^{pq}$, $i = 1, \ldots, n-p+1$. Also, define the standardized residuals $\hat{\boldsymbol{u}}_i = \boldsymbol{S}^{-\frac{1}{2}}(\boldsymbol{u}_i - \overline{\boldsymbol{u}})$ with the sample covariance matrix $\boldsymbol{S} = \frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{u}_j - \overline{\boldsymbol{u}})(\boldsymbol{u}_j - \overline{\boldsymbol{u}})^{\mathrm{T}}$ and the sample mean $\overline{\boldsymbol{u}} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{u}_j$. Now, let

$$\hat{S}_{n,A}(\boldsymbol{t}) = \sqrt{n} \sum_{B \subset A} (-1)^{|A \setminus B|} \hat{\phi}_{n,p}(\boldsymbol{t}^B) \prod_{i \in A \setminus B} \phi(\boldsymbol{t}^{(i)}),$$

where $\hat{\phi}_{n,p}(\boldsymbol{t}) = \frac{1}{n}\sum_{j=1}^{n-p+1} \exp(\mathrm{i}\langle \boldsymbol{t}, \boldsymbol{e}_j \rangle)$. The asymptotic behavior of these processes is stated next. The main conclusion is that the estimation of the unknown parameters does not affect the asymptotic distribution.

Theorem 3.2. *If the $\boldsymbol{u}_i$'s are independent, the processes $\{\hat{S}_{n,A} : |A| > 1\}$ converge in $C(\mathbb{R}^{pq}, \mathbb{C})$ to independent zero mean complex Gaussian processes $\{R_A : |A| > 1\}$ having covariance and pseudo-covariance functions, respectively, given by $C_A(\boldsymbol{s}, \boldsymbol{t})$ and $P_A(\boldsymbol{s}, \boldsymbol{t})$ in* (2.2) *and* (2.3).

Note that the multinomial formula yields

$$\hat{S}_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n-p+1} \prod_{k \in A} [\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{e}_j^{(k)} \rangle) - \phi(\boldsymbol{t}^{(k)})]$$

and so the Cramér–von Mises test statistic

$$\tilde{T}_{n,b,A} = \frac{1}{n} \int |\hat{S}_{n,A}(\boldsymbol{t})|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t}$$

can be computed as

$$\tilde{T}_{n,b,A} = \frac{1}{n^2}\sum_{l=1}^{n-p+1}\sum_{l'=1}^{n-p+1}\prod_{k\in A}\Bigg\{\exp\left[-\frac{b^2}{2}\|\boldsymbol{e}_l^{(k)}-\boldsymbol{e}_{l'}^{(k)}\|^2\right] -(b^2+1)^{-\frac{q}{2}}\exp\left[-\frac{1}{2}\frac{b^2}{b^2+1}\|\boldsymbol{e}_l^{(k)}\|^2\right] -(b^2+1)^{-\frac{q}{2}}\exp\left[-\frac{1}{2}\frac{b^2}{b^2+1}\|\boldsymbol{e}_{l'}^{(k)}\|^2\right]+(2b^2+1)^{-\frac{q}{2}}\Bigg\}.$$

This representation shows that $\tilde{T}_{n,b,A}$ is affine invariant. Here, again we can use Theorem 2.3 to obtain

Theorem 3.3. *For two different subsets A and B of I_p,*

$$\int\left(|\hat{S}_{n,A}(\boldsymbol{t})|^2, |\hat{S}_{n,B}(\boldsymbol{t})|^2\right)\varphi_b(\boldsymbol{t})\,d\boldsymbol{t} \xrightarrow{D} \int\left(|R_A(\boldsymbol{t})|^2, |R_B(\boldsymbol{t})|^2\right)\varphi_b(\boldsymbol{t})\,d\boldsymbol{t},$$

where integrals are computed componentwise.

In the serial situation, a subset A and its translate $A+k$ lead essentially to the same statistic $\tilde{T}_{n,b,A}$. Hence, when considering these statistics, only A's such that $1 \in A$ can be considered. We can then use the following statistics to perform the statistical test

$$\tilde{S}_n = n\sum_{|A|>1, 1\in A}\tilde{T}_{n,b,A}, \quad \tilde{M}_n = n\max_{|A|>1, 1\in A}\tilde{T}_{n,b,A}.$$

4. Asymptotic critical values of $T_{n,b,A}$ and $\tilde{T}_{n,b,A}$

It has already been established that the asymptotic distributions of $T_{n,b,A}$ and $\tilde{T}_{n,b,A}$ are identical. This section shows how to compute the critical values of the Cramér–von Mises variable $T_{b,A} \equiv \int |R_A(\boldsymbol{t})|^2\varphi_b(\boldsymbol{t})\,d\boldsymbol{t}$. This can be achieved either by inversion of the characteristic function (see [24] or an improved version of this algorithm introduced by Davies [11,12] or Deheuvels and Martynov [17]) after evaluation of the eigenvalues of C_A or by computing its cumulants and then applying the Cornish–Fisher asymptotic expansion (see [27,28]).

The Cramér–von Mises test statistic in (2.6) can also be written as

$$T_{n,b,A} = \frac{1}{n}\int W_{n,A}^2(\boldsymbol{t})\varphi_b(\boldsymbol{t})\,d\boldsymbol{t},$$

where $W_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^n\prod_{k\in A}\left[\cos(\langle\boldsymbol{t}^{(k)},\boldsymbol{e}_j^{(k)}\rangle)+\sin(\langle\boldsymbol{t}^{(k)},\boldsymbol{e}_j^{(k)}\rangle)-\phi(\boldsymbol{t}^{(k)})\right]$ is a real process which converges to a real Gaussian process with the same covariance function $C_A(\boldsymbol{s},\boldsymbol{t})$ as in (2.2).

Let $k = |A|$. Also, let $\sim$ be the equality in distribution symbol. It is well known from the Karhunen–Loève expansion that

$$\int_{\mathbb{R}^{pq}} |R_A(\boldsymbol{t})|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t} \sim T_{b,A} = \sum_{(i_1,\ldots,i_k)\in\mathbb{N}^{*k}} \lambda_{(i_1,\ldots,i_k)} \chi^2_{1;(i_1,\ldots,i_k)}, \tag{4.1}$$

where $\chi^2_{1;(i_1,\ldots,i_k)}$ are independent chi-square variables with one degree of freedom. Also, it is easy to show that $\lambda_{(i_1,\ldots,i_k)} = \prod_{l=1}^{k} \lambda_{i_l}$ is a product of the eigenvalues λ_j of the integral operator O defined by

$$O(f)(\boldsymbol{x}) = \int_{\mathbb{R}^q} f(\boldsymbol{y}) K(\boldsymbol{x},\boldsymbol{y}) \tilde{\varphi}_b(\boldsymbol{y})\, d\boldsymbol{y}$$

with the kernel

$$K(\boldsymbol{x},\boldsymbol{y}) = \exp\left(-\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2\right) - \exp\left(-\frac{1}{2}(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2)\right)$$

and where $\tilde{\varphi}_b$ is the $N_q(\boldsymbol{0}, b^2\boldsymbol{I})$ density. That is to say, the problem is to solve, in λ (and f), the linear second-order homogeneous Fredholm integral equation

$$\lambda f(\boldsymbol{x}) = \int_{\mathbb{R}^q} f(\boldsymbol{y}) K(\boldsymbol{x},\boldsymbol{y}) \tilde{\varphi}_b(\boldsymbol{y})\, d\boldsymbol{y}. \tag{4.2}$$

See [5] for an introduction to integral operators. It does not seem possible to solve (4.2) explicitly, but one can approximate the eigenvalues. A change of variables in (4.2) gives the integral equation

$$\lambda f(b\sqrt{2}\boldsymbol{x}) = \int_{\mathbb{R}^q} \pi^{-\frac{q}{2}} f(b\sqrt{2}\boldsymbol{y}) K(b\sqrt{2}\boldsymbol{x}, b\sqrt{2}\boldsymbol{y}) \exp(-\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y})\, d\boldsymbol{y},$$

which results in the discretization

$$\lambda f(b\sqrt{2}\boldsymbol{y}_i) = \sum_{j=1}^{N} B_j \pi^{-\frac{q}{2}} f(b\sqrt{2}\boldsymbol{y}_j) K(b\sqrt{2}\boldsymbol{y}_i, b\sqrt{2}\boldsymbol{y}_j), \quad i = 1,\ldots,N. \tag{4.3}$$

Then, with $\boldsymbol{D} = \mathrm{diag}(\pi^{-\frac{q}{2}} B_j)$, $\boldsymbol{f} = (f_1,\ldots,f_N)^{\mathrm{T}}$, where $f_j = f(b\sqrt{2}\boldsymbol{y}_j)$, and $\boldsymbol{K} = (K(b\sqrt{2}\boldsymbol{y}_i, b\sqrt{2}\boldsymbol{y}_j)) : N \times N$, (4.3) is equivalent to $\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{K}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{h} = \lambda\boldsymbol{h}$, where $\boldsymbol{h} = \boldsymbol{D}^{\frac{1}{2}}\boldsymbol{f}$, from which the first N approximated eigenvalues can be obtained.

The parameters B_j and $\boldsymbol{y}_j = (y_{j1},\ldots,y_{jq})$, $j = 1,\ldots,N$ could be, respectively, the coefficients and the points of a cubature formula (CF), or else they could be obtained by a Monte Carlo experiment, in which case $B_j = \pi^{\frac{q}{2}}/N$ and $\boldsymbol{y}_j \sim N_q(\boldsymbol{0}, \frac{1}{2}\boldsymbol{I})$, $j = 1,\ldots,N$. A good rule of thumb is to use a cubature formula when b is small, for example less than one, otherwise use the Monte Carlo method.

The cubature formulas we used are the following: the Nth degree Gauss quadrature formula when $q = 1$, the 15th degree CF $E_2^{r^2} : 15-1$ (see [32, p. 326]) when $q = 2$ and the seventh degree CF $E_q^{r^2}$: 7 − 2 appearing in [32, p. 319] for $q \geqslant 3$. This last formula contains an error, see [31] for details. See [6,7] for a comprehensive list of such formulas.

It is also the case that all the cumulants $\kappa_{b,A}(m)$ of $T_{b,A}$ can be computed explicitly. As a first method, it follows immediately from (4.1) and the factorization of the eigenvalues that the mth cumulant is given by

$$\kappa_{b,A}(m) = 2^{m-1}(m-1)! \left[\sum_{j=1}^{\infty} \lambda_j^m \right]^{|A|}.$$

A second method, which permits to double-check the preceding computation of the eigenvalues, is through the following relation using iterated kernels:

$$\sum_{j=1}^{\infty} \lambda_j^m = \int_{\mathbb{R}^q} K^{(m)}(\boldsymbol{x}, \boldsymbol{x}) \tilde{\varphi}_b(\boldsymbol{x})\, d\boldsymbol{x}, \tag{4.4}$$

where $K^{(1)}(\boldsymbol{x}, \boldsymbol{y}) = K(\boldsymbol{x}, \boldsymbol{y})$ and

$$K^{(m)}(\boldsymbol{x}, \boldsymbol{y}) = \int_{\mathbb{R}^q} K^{(m-1)}(\boldsymbol{x}, \boldsymbol{z}) K(\boldsymbol{z}, \boldsymbol{y}) \tilde{\varphi}_b(\boldsymbol{z})\, d\boldsymbol{z}, \quad m \geqslant 2.$$

The integral involving an iterated kernel in (4.4) can be evaluated as follows. Define

$$\begin{aligned} I^{(1)}_{\boldsymbol{x},\boldsymbol{y}}(\alpha, \beta, \gamma) &= \exp(\alpha\|\boldsymbol{x}\|^2 + \beta\|\boldsymbol{y}\|^2 + \gamma\langle \boldsymbol{x}, \boldsymbol{y}\rangle), \\ I^{(m)}_{\boldsymbol{x},\boldsymbol{y}}(\alpha, \beta, \gamma) &= \int_{\mathbb{R}^q} I^{(m-1)}_{\boldsymbol{x},\boldsymbol{z}}(\alpha, \beta, \gamma) K(\boldsymbol{z}, \boldsymbol{y}) \tilde{\varphi}_b(\boldsymbol{z})\, d\boldsymbol{z}, \quad m \geqslant 2. \end{aligned}$$

One can show the identity

$$K^{(m)}(\boldsymbol{x}, \boldsymbol{x}) = I^{(m)}_{\boldsymbol{x},\boldsymbol{x}}\left(-\frac{1}{2}, -\frac{1}{2}, 1\right) - I^{(m)}_{\boldsymbol{x},\boldsymbol{x}}\left(-\frac{1}{2}, -\frac{1}{2}, 0\right)$$

and the recurrence relation

$$\begin{aligned} I^{(m)}_{\boldsymbol{x},\boldsymbol{x}}(\alpha, \beta, \gamma) = \chi_\beta \Bigg[& I^{(m-1)}_{\boldsymbol{x},\boldsymbol{x}}\left(\alpha - \frac{\gamma^2}{4c}, -\frac{1}{4c} - \frac{1}{2}, -\frac{2\gamma}{4c}\right) \\ & - I^{(m-1)}_{\boldsymbol{x},\boldsymbol{x}}\left(\alpha - \frac{\gamma^2}{4c}, -\frac{1}{2}, 0\right) \Bigg], \end{aligned}$$

where $\chi_\beta = (2\pi)^{-\frac{q}{2}} b^{-q} \int_{\mathbb{R}^q} \exp\left(c\|\boldsymbol{z}\|^2\right) d\boldsymbol{z} = (1+b^2-2b^2\beta)^{-\frac{q}{2}}$ and $-\frac{1}{4c} = \frac{b^2}{2+2b^2-4b^2\beta}$. Thus, one can express $K^{(m)}(\boldsymbol{x}, \boldsymbol{x})$ in terms of several $I^{(1)}_{\boldsymbol{x},\boldsymbol{x}}$ values and use the relation

$$\int_{\mathbb{R}^q} I^{(1)}_{\boldsymbol{x},\boldsymbol{x}}(\alpha, \beta, \gamma) \tilde{\varphi}_b(\boldsymbol{x})\, d\boldsymbol{x} = [1 - 2b^2(\alpha + \beta + \gamma)]^{-\frac{q}{2}}$$

to obtain all the cumulants recursively.

Table 1 provides an approximation of the cut-off values obtained from the Cornish–Fisher asymptotic expansion based on the first six cumulants, for $b = 0.1$. Recall that in all the tables $k = |A|$.

Table 1
Critical values of the distribution of $T_{b,A}$ for $b = 0.1$

$1-\alpha$	$q=2$			$q=3$		
	$k=2$	$k=3$	$k=4$	$k=2$	$k=3$	$k=4$
0.900	0.000733	1.230e − 05	2.122e − 07	0.00137	3.347e − 05	8.707e − 07
0.905	0.000745	1.244e − 05	2.140e − 07	0.00138	3.370e − 05	8.741e − 07
0.910	0.000758	1.260e − 05	2.159e − 07	0.00140	3.393e − 05	8.777e − 07
0.915	0.000771	1.276e − 05	2.179e − 07	0.00142	3.418e − 05	8.815e − 07
0.920	0.000785	1.293e − 05	2.200e − 07	0.00143	3.444e − 05	8.854e − 07
0.925	0.000800	1.311e − 05	2.222e − 07	0.00145	3.472e − 05	8.896e − 07
0.930	0.000815	1.330e − 05	2.246e − 07	0.00147	3.501e − 05	8.939e − 07
0.935	0.000832	1.350e − 05	2.271e − 07	0.00149	3.531e − 05	8.986e − 07
0.940	0.000850	1.372e − 05	2.298e − 07	0.00152	3.564e − 05	9.035e − 07
0.945	0.000870	1.395e − 05	2.326e − 07	0.00154	3.600e − 05	9.088e − 07
0.950	0.000891	1.421e − 05	2.357e − 07	0.00157	3.638e − 05	9.145e − 07
0.955	0.000915	1.449e − 05	2.392e − 07	0.00160	3.680e − 05	9.207e − 07
0.960	0.000941	1.480e − 05	2.429e − 07	0.00163	3.726e − 05	9.275e − 07
0.965	0.000971	1.514e − 05	2.471e − 07	0.00167	3.777e − 05	9.351e − 07
0.970	0.001005	1.554e − 05	2.519e − 07	0.00171	3.836e − 05	9.437e − 07
0.975	0.001045	1.601e − 05	2.575e − 07	0.00176	3.904e − 05	9.537e − 07
0.980	0.001093	1.657e − 05	2.643e − 07	0.00182	3.985e − 05	9.656e − 07
0.985	0.001156	1.729e − 05	2.728e − 07	0.00190	4.088e − 05	9.805e − 07
0.990	0.001243	1.828e − 05	2.845e − 07	0.00200	4.229e − 05	1.000e − 06
0.995	0.001390	1.994e − 05	3.039e − 07	0.00217	4.460e − 05	1.033e − 06

Table 2
Cumulative distribution function, $P[T^*_{b,A} \leqslant x]$, for $b = 0.1$

x	$q=2$			$q=3$		
	$k=2$	$k=3$	$k=4$	$k=2$	$k=3$	$k=4$
0.0	0.593	0.566	0.546	0.562	0.536	0.520
0.2	0.665	0.640	0.622	0.637	0.612	0.598
0.4	0.725	0.705	0.691	0.702	0.683	0.671
0.6	0.777	0.761	0.751	0.759	0.745	0.737
0.8	0.819	0.809	0.802	0.807	0.798	0.794
1.0	0.854	0.848	0.845	0.848	0.843	0.842
1.2	0.883	0.881	0.880	0.880	0.880	0.881
1.4	0.906	0.907	0.908	0.907	0.910	0.913
1.6	0.925	0.928	0.931	0.928	0.933	0.937
1.8	0.941	0.944	0.948	0.945	0.951	0.955
2.0	0.953	0.957	0.961	0.958	0.964	0.969
2.2	0.963	0.967	0.972	0.968	0.974	0.979
2.4	0.970	0.975	0.979	0.976	0.982	0.986
2.6	0.977	0.981	0.985	0.982	0.987	0.990
2.8	0.982	0.986	0.989	0.986	0.991	0.994
3.0	0.985	0.989	0.992	0.990	0.994	0.996
3.2	0.988	0.992	0.994	0.992	0.996	0.997
3.4	0.991	0.994	0.996	0.994	0.997	0.998
3.6	0.993	0.995	0.997	0.996	0.998	0.999
3.8	0.994	0.996	0.998	0.997	0.998	0.999
4.0	0.995	0.997	0.998	0.997	0.999	0.999

Table 3
Empirical percentage points of $nT_{n,b,A}$ based on $N = 10\,000$ Monte Carlo replications: non-serial case

$1-\alpha$	b	n	$q=2$			$q=3$		
			$k=2$	$k=3$	$k=4$	$k=2$	$k=3$	$k=4$
		20	0.000718	1.249e − 05	2.371e − 07	0.00134	3.460e − 05	1.033e − 06
	0.1	50	0.000718	1.253e − 05	2.358e − 07	0.00136	3.420e − 05	9.748e − 07
		100	0.000733	1.229e − 05	2.242e − 07	0.00137	3.389e − 05	9.430e − 07
		20	0.170	0.049	0.016	0.269	0.110	0.050
	0.5	50	0.167	0.0481	0.0156	0.266	0.107	0.0484
0.9		100	0.168	0.0471	0.0149	0.266	0.106	0.0472
		20	0.561	0.333	0.221	0.727	0.555	0.450
	1.0	50	0.558	0.327	0.213	0.724	0.547	0.440
		100	0.559	0.326	0.209	0.723	0.544	0.436
		20	0.938	0.864	0.820	0.981	0.969	0.958
	3.0	50	0.940	0.860	0.814	0.983	0.967	0.955
		100	0.938	0.858	0.811	0.983	0.966	0.954
		20	0.000854	1.485e − 05	3.036e − 07	0.00151	3.849e − 05	1.212e − 06
	0.1	50	0.000883	1.465e − 05	2.840e − 07	0.00156	3.751e − 05	1.103e − 06
		100	0.000874	1.429e − 05	2.653e − 07	0.00155	3.724e − 05	1.027e − 06
		20	0.191	0.0539	0.0184	0.289	0.114	0.053
	0.5	50	0.192	0.0520	0.0168	0.289	0.111	0.0505
0.95		100	0.191	0.0513	0.0158	0.289	0.110	0.0486
		20	0.606	0.345	0.228	0.749	0.562	0.456
	1.0	50	0.603	0.338	0.218	0.748	0.553	0.445
		100	0.593	0.336	0.213	0.745	0.549	0.439
		20	0.955	0.868	0.823	0.984	0.970	0.959
	3.0	50	0.954	0.863	0.816	0.987	0.967	0.956
		100	0.953	0.861	0.813	0.986	0.966	0.955

As in Ghoudi et al. [21], define $T^*_{b,A} \equiv [T_{b,A} - E(T_{b,A})]/\sqrt{\mathrm{Var}(T_{b,A})}$ to be the standardized version of $T_{b,A}$. This standardization could be preferable since variables $T_{b,A}$ with small values of $|A|$ have a larger variance. Then, Table 2 provides the distribution function of this statistic for $|A| = 2, 3, 4$ and $q = 2, 3$, with $b = 0.1$ as approximated by the technique of Davies. Besides, a C++ program is available from the authors which permits to compute any cut-off value given the nominal level and vice versa.

One can find in Table 3 the empirical percentage points of the null distribution of $nT_{n,b,A}$ ($n = 20, 50, 100$; $b = 0.1, 0.5, 1.0, 3.0$; $\alpha = 0.1, 0.05$) based on $N = 10\,000$ Monte Carlo replications, in the non-serial case.

Table 4 gives the empirical percentage points of $nT_{n,b,A}$ ($n = 20, 50, 100$; $b = 0.1, 0.5, 1.0, 3.0$; $\alpha = 0.1, 0.05$) based on $N = 10\,000$ Monte Carlo replications, in the serial case, for $p = 4$.

The empirical percentage points in Tables 3 and 4 are in close agreement. This should be expected since the asymptotic distribution of the test statistics is the same in the serial and non-serial cases. Also, these empirical percentage points for $b = 0.1$ can be compared to the asymptotic critical values in Table 1. This comparison supports the use of the asymptotic critical values even for moderate sample sizes.

Table 4
Empirical percentage points of $nT_{n,b,A}$ based on $N = 10\,000$ Monte Carlo replications for $p = 4$: serial case

$1-\alpha$	b	n	$q=2$			$q=3$		
			$k=2$	$k=3$	$k=4$	$k=2$	$k=3$	$k=4$
		20	0.000712	1.174e − 05	2.139e − 07	0.00132	3.384e − 05	9.931e − 07
	0.1	50	0.000721	1.205e − 05	2.288e − 07	0.00135	3.357e − 05	9.867e − 07
		100	0.000716	1.217e − 05	2.261e − 07	0.00138	3.335e − 05	9.530e − 07
		20	0.173	0.0504	0.0171	0.276	0.1160	0.0547
	0.5	50	0.168	0.0484	0.0162	0.269	0.109	0.050
0.9		100	0.170	0.0476	0.0155	0.268	0.107	0.048
		20	0.567	0.353	0.241	0.737	0.586	0.486
	1.0	50	0.561	0.334	0.223	0.728	0.559	0.458
		100	0.561	0.328	0.216	0.727	0.55	0.447
		20	0.937	0.880	0.842	0.984	0.975	0.966
	3.0	50	0.939	0.867	0.825	0.983	0.970	0.960
		100	0.939	0.862	0.818	0.983	0.968	0.957
		20	0.000841	1.414e − 05	2.734e − 07	0.00148	3.804e − 05	1.176e − 06
	0.1	50	0.000897	1.412e − 05	2.780e − 07	0.00154	3.772e − 05	1.128e − 06
		100	0.000861	1.435e − 05	2.737e − 07	0.00154	3.688e − 05	1.071e − 06
		20	0.196	0.0556	0.0193	0.296	0.122	0.0587
	0.5	50	0.192	0.0528	0.0179	0.292	0.114	0.0539
0.95		100	0.189	0.0517	0.0168	0.290	0.111	0.0509
		20	0.615	0.367	0.255	0.759	0.598	0.498
	1.0	50	0.604	0.346	0.231	0.753	0.567	0.466
		100	0.602	0.340	0.222	0.751	0.557	0.452
		20	0.955	0.886	0.849	0.985	0.976	0.968
	3.0	50	0.953	0.871	0.829	0.987	0.971	0.961
		100	0.952	0.865	0.822	0.987	0.969	0.958

5. One-way MANOVA model with random effects

The one-way linear model with random effects

$$\boldsymbol{\epsilon}_i^{(j)} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\delta}_i^{(j)}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, p,$$

where $\boldsymbol{\alpha}_i \sim N_q(\mathbf{0}, \boldsymbol{\Psi})$ and $\boldsymbol{\delta}_i^{(j)} \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$ are all mutually independent, provides a joint normal model for the non-serial case. This means that in this variance component model

$$\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_i^{(1)}, \ldots, \boldsymbol{\epsilon}_i^{(p)}) \sim N_{pq}(\mathbf{1}_p \otimes \boldsymbol{\mu}, (\boldsymbol{I}_p \otimes \boldsymbol{\Sigma}) + (\mathbf{1}_p \mathbf{1}_p^{\mathrm{T}}) \otimes \boldsymbol{\Psi}), \quad i = 1, \ldots, n,$$

are i.i.d. The test of independence amounts to the parametric test of the hypothesis $\mathrm{H}_0 : \mathbf{\Psi} = \mathbf{0}$. The MANOVA decomposition

$$\sum_{i=1}^{n}\sum_{j=1}^{p}(\epsilon_i^{(j)} - \overline{\epsilon}_{\bullet}^{(\bullet)})(\epsilon_i^{(j)} - \overline{\epsilon}_{\bullet}^{(\bullet)})^{\mathrm{T}} = \sum_{i=1}^{n}\sum_{j=1}^{p}(\epsilon_i^{(j)} - \overline{\epsilon}_i^{(\bullet)})(\epsilon_i^{(j)} - \overline{\epsilon}_i^{(\bullet)})^{\mathrm{T}} + p\sum_{i=1}^{n}(\overline{\epsilon}_i^{(\bullet)} - \overline{\epsilon}_{\bullet}^{(\bullet)})(\overline{\epsilon}_i^{(\bullet)} - \overline{\epsilon}_{\bullet}^{(\bullet)})^{\mathrm{T}} = E + H,$$

(see [30]), leads to the usual MANOVA table. A dot means averaging over the corresponding index. The test of $\mathrm{H}_0 : \mathbf{\Psi} = \mathbf{0}$ is usually done with Wilks statistic

$$\Lambda = \frac{\det \mathrm{E}}{\det(\mathrm{E}+\mathrm{H})} = \prod_{k=1}^{q}\left[1 + \lambda_k(E^{-1}H)\right]^{-1},$$

where $\lambda_k(E^{-1}H)$ are the eigenvalues of $E^{-1}H$. The null distribution of Λ is the Λ_{q,ν_H,ν_E} distribution, where the degrees of freedom for hypothesis and error are, respectively, $\nu_H = n-1$ and $\nu_E = n(p-1)$. Tables of exact critical points for Λ are available but for $q = 2$ the relation

$$\frac{(\nu_E - 1)}{\nu_H}\frac{(1-\sqrt{\Lambda})}{\sqrt{\Lambda}} \sim F_{2\nu_H, 2(\nu_E - 1)}$$

holds.

Looking at Table 5, we see that Wilks test has power superior to the test proposed in this paper. This is not surprising since Wilks test is specifically designed for the parametric hypothesis $\mathbf{\Psi} = \mathbf{0}$ in the linear model which holds in the simulation. However, our test is more generally applicable and will yield reasonable power for reasonably large sample sizes. Moreover, it is not difficult to construct another model where Wilks test would fail. For example, contrary to our test, Wilks test would be unable to detect the dependence in the normal mixture model

$$\begin{pmatrix}\boldsymbol{\epsilon}^{(1)}\\ \boldsymbol{\epsilon}^{(2)}\end{pmatrix} \sim 0.5N_4\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\boldsymbol{I}_2 & \mathbf{\Psi}\\ \mathbf{\Psi} & \boldsymbol{I}_2\end{pmatrix}\right) + 0.5N_4\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\boldsymbol{I}_2 & -\mathbf{\Psi}\\ -\mathbf{\Psi} & \boldsymbol{I}_2\end{pmatrix}\right),$$

in which $\mathrm{cov}(\boldsymbol{\epsilon}^{(1)}, \boldsymbol{\epsilon}^{(2)}) = 0$ holds, since the expected value of the between sum of squares H is then a quantity independent of $\mathbf{\Psi}$.

6. Proofs

Define the metric ρ on $C := C(\mathbb{R}^{pq}, \mathbb{C})$ by

$$\rho(x, y) = \sum_{j=1}^{\infty} 2^{-j}\frac{\rho_j(x, y)}{1 + \rho_j(x, y)},$$

Table 5
Empirical power of $nT_{n,b,A}$ and Wilks test based on $N = 10\,000$ Monte Carlo replications for $p = 2$, $q = 2$, $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \gamma \boldsymbol{I}_2$ and $\boldsymbol{\Psi} = \theta \boldsymbol{I}_2$

$1-\alpha$	$\frac{\theta}{\gamma}$	n	$nT_{n,b,A}$					
			$b = 0.01$	$b = 0.1$	$b = 0.5$	$b = 1.0$	$b = 3.0$	Wilks
	0.0	20	10.14	10.27	10.06	10.21	10.33	9.96
		50	9.46	10.14	10.15	10.20	9.94	10.34
	0.2	20	14.27	14.39	13.78	12.64	11.12	39.15
0.90		50	28.61	29.87	27.50	19.00	10.50	63.63
	0.4	20	28.43	28.57	26.61	19.28	11.81	67.69
		50	68.31	69.70	63.43	40.63	13.67	94.72
	1.0	20	77.09	77.36	72.66	50.48	16.49	97.86
		50	99.80	99.81	99.43	92.68	27.66	99.99
	0.0	20	4.67	4.91	5.52	4.88	5.16	5.22
		50	4.84	4.58	4.98	4.82	5.04	5.22
	0.2	20	7.35	7.63	7.94	6.39	5.53	26.10
0.95		50	18.96	18.26	17.11	10.60	5.14	49.60
	0.4	20	16.77	17.20	16.93	10.80	6.03	53.44
		50	56.83	55.72	50.23	28.08	7.37	89.24
	1.0	20	64.57	65.61	61.12	36.17	8.98	94.91
		50	99.45	99.42	98.61	86.81	17.72	99.98

where

$$\rho_j(x, y) = \sup_{\|\boldsymbol{t}\| \leqslant j} |x(\boldsymbol{t}) - y(\boldsymbol{t})|$$

is the usual sup norm. Endowed with this metric, $C(\mathbb{R}^p, \mathbb{C})$ is a separable Fréchet space, that is, a linear complete and separable metric space. Moreover, convergence in this metric corresponds to the uniform convergence on all compact sets, $\lim_{n\to\infty} \rho(x_n, x) = 0$ if and only if $\forall j \geqslant 1$, $\lim_{n\to\infty} \rho_j(x_n, x) = 0$. For random elements, it also holds $\rho(X_n, Y_n) \xrightarrow{P} 0$ if and only if $\forall j \geqslant 1$, $\rho_j(X_n, Y_n) \xrightarrow{P} 0$. For all mappings $x \in C$, let $r_j(x)$ be the restriction of x to the ball $\mathcal{B}_j^{pq}$ of radius j. The symbol $\xrightarrow{fd}$ will represent convergence in distribution of the process at a finite number of points in the index set.

6.1. *Proof of Theorem 2.1*

By Propositions 14.6, 14.2 and Theorem 14.3 of Kallenberg [25], it suffices to show that $r_j(R_{n,A}) \xrightarrow{fd} r_j(R_A)$ and $\{r_j(R_{n,A})\}$ is a tight family in order to show the convergence $R_{n,A} \xrightarrow{D} R_A$. Convergence of the finite-dimensional distributions to Gaussian limit is a direct consequence of the multivariate central limit theorem and the representation (2.4) of the process. The covariance function C_A of (2.2) and the independence of the processes R_A

are also easy to obtain. Then, one can write

$$R_{n,A}(\boldsymbol{t}) = \sum_{B\subset A} (-1)^{|A\setminus B|}\phi_{A,B}(\boldsymbol{t})Y_{n,B}(\boldsymbol{t}), \tag{6.1}$$

where

$$Y_{n,B}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{\exp(\mathrm{i}\langle \boldsymbol{t}^B, \boldsymbol{\epsilon}_j\rangle) - E\left[\exp(\mathrm{i}\langle \boldsymbol{t}^B, \boldsymbol{\epsilon}_1\rangle)\right]\right\}$$

and $\phi_{A,B}(\boldsymbol{t}) = \prod_{l\in A\setminus B}\phi(\boldsymbol{t}^{(l)})$. The process $Y_{n,B}$ on a compact is an empirical characteristic function process which was shown by Csörgő [9] to be tight if the sufficient condition in [10, p. 294] is satisfied. Hence, $r_j(Y_{n,B})$ is tight. Since there is only a finite number of B's in (6.1), it follows that $\{r_j(R_{n,A})\}$ is also tight. □

6.2. *Proof of Theorem 2.2*

By invariance assume $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}$. Let $\boldsymbol{e}_j^{(k)} = \boldsymbol{S}^{-\frac{1}{2}}(\boldsymbol{\epsilon}_j^{(k)} - \overline{\boldsymbol{\epsilon}}) = \boldsymbol{\epsilon}_j^{(k)} + \boldsymbol{\Delta}_j^{(k)}$, with $\boldsymbol{\Delta}_j^{(k)} = (\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I})\boldsymbol{\epsilon}_j^{(k)} - \boldsymbol{S}^{-\frac{1}{2}}\overline{\boldsymbol{\epsilon}}$. Two intermediate processes are now defined. Firstly,

$$\begin{aligned}\tilde{R}_{n,A}(\boldsymbol{t}) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\Bigg\{\prod_{k\in A}(\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})) \\ &\quad + \mathrm{i}\sum_{\alpha\in A} a_{\alpha,j}\prod_{k\in A;k\neq\alpha}\left[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})\right]\exp(\mathrm{i}\langle \boldsymbol{t}^{(\alpha)}, \boldsymbol{\epsilon}_j^{(\alpha)}\rangle)\Bigg\} \\ &\equiv \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\prod_{k\in A}\left[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})\right] + \mathrm{i}U_n(\boldsymbol{t}),\end{aligned} \tag{6.2}$$

where $a_{\alpha,j} = \langle \boldsymbol{t}^{(\alpha)}, \boldsymbol{\Delta}_j^{(\alpha)}\rangle$. Secondly, define also

$$\breve{R}_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\prod_{k\in A}[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})]. \tag{6.3}$$

The proof proceeds with the following steps:

$$\rho(\hat{R}_{n,A}, \tilde{R}_{n,A}) \xrightarrow{P} 0, \tag{6.4}$$

$$\rho(\tilde{R}_{n,A}, \breve{R}_{n,A}) \xrightarrow{P} 0, \tag{6.5}$$

$$\{\breve{R}_{n,A} : |A| > 1\} \xrightarrow{D} \{R_A : |A| > 1\}. \tag{6.6}$$

The last step was proven in Theorem 2.1. A Taylor expansion of

$$\prod_{k\in A}\left[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle + a_{k,j}) - \phi(\boldsymbol{t}^{(k)})\right]$$

around $(a_{k,j})_{k\in A} = \mathbf{0}$ and the Schwarz inequality yield

$$\left|\hat{R}_{n,A}(\boldsymbol{t}) - \tilde{R}_{n,A}(\boldsymbol{t})\right| \leqslant 2^{|A|-2}|A|^2\|\boldsymbol{t}\|^2 \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\|\boldsymbol{\Delta}_j\|^2, \tag{6.7}$$

where $\boldsymbol{\Delta}_j = (\boldsymbol{\Delta}_j^{(1)}, \ldots, \boldsymbol{\Delta}_j^{(p)})$. Now

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\|\boldsymbol{\Delta}_j\|^2 = \sum_{k=1}^{p}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\|\boldsymbol{\Delta}_j^{(k)}\|^2,$$

where each of the p terms is expressed as

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\|\boldsymbol{\Delta}_j^{(k)}\|^2 = \sqrt{n}\,\mathrm{tr}\left[(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I})^2\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\epsilon}_j^{(k)}\boldsymbol{\epsilon}_j^{(k)\mathrm{T}}\right] \\ -2\overline{\boldsymbol{\epsilon}}^{(k)\mathrm{T}}\boldsymbol{S}^{-\frac{1}{2}}\sqrt{n}(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I})\overline{\boldsymbol{\epsilon}} + \sqrt{n}\overline{\boldsymbol{\epsilon}}^{\mathrm{T}}\boldsymbol{S}^{-1}\overline{\boldsymbol{\epsilon}}.$$

In view of $\overline{\boldsymbol{\epsilon}}^{(k)} = O_P(n^{-\frac{1}{2}})$, $\overline{\boldsymbol{\epsilon}} = O_P(n^{-\frac{1}{2}})$,

$$\sqrt{np}(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I}) = -\frac{1}{2\sqrt{np}}\sum_{j=1}^{n}\sum_{k=1}^{p}\left(\boldsymbol{\epsilon}_j^{(k)}\boldsymbol{\epsilon}_j^{(k)\mathrm{T}} - \boldsymbol{I}\right) + o_P(1)$$

and

$$\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\epsilon}_j^{(k)}\boldsymbol{\epsilon}_j^{(k)\mathrm{T}} = \boldsymbol{I} + O_P(n^{-\frac{1}{2}})$$

(see [22, p. 9]), we obtain $\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\|\boldsymbol{\Delta}_j\|^2 \xrightarrow{P} 0$. Using (6.7), $\rho_k(\hat{R}_{n,A}, \tilde{R}_{n,A}) \xrightarrow{P} 0$ and so (6.4) is proved. To establish (6.5) consider $\rho_k(\tilde{R}_{n,A}(\boldsymbol{t}), \breve{R}_{n,A}(\boldsymbol{t})) = \sup_{\|\boldsymbol{t}\|\leqslant k}|U_n(\boldsymbol{t})|$, where

$$U_n(\boldsymbol{t}) = \sum_{\alpha\in A}\left\{\boldsymbol{t}^{(\alpha)\mathrm{T}}\sqrt{n}(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I})\left[A_{n,\alpha}(\boldsymbol{t}^{(A)}) - \overline{\boldsymbol{\epsilon}}B_{n,\alpha}(\boldsymbol{t}^{(A)})\right] - \boldsymbol{t}^{(\alpha)\mathrm{T}}\sqrt{n}\overline{\boldsymbol{\epsilon}}B_{n,\alpha}(\boldsymbol{t}^{(A)})\right\} \tag{6.8}$$

with

$$A_{n,\alpha}(\boldsymbol{t}^{(A)}) = \frac{1}{n}\sum_{j=1}^{n}\left\{\boldsymbol{\epsilon}_j^{(\alpha)}\exp(\mathrm{i}\langle\boldsymbol{t}^{(\alpha)}, \boldsymbol{\epsilon}_j^{(\alpha)}\rangle)\prod_{k\in A;k\neq\alpha}\left[\exp(\mathrm{i}\langle\boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})\right]\right\}$$

and

$$B_{n,\alpha}(\boldsymbol{t}^{(A)}) = \frac{1}{n}\sum_{j=1}^{n}\left\{\exp(\mathrm{i}\langle\boldsymbol{t}^{(\alpha)}, \boldsymbol{\epsilon}_j^{(\alpha)}\rangle)\prod_{k\in A;k\neq\alpha}\left[\exp(\mathrm{i}\langle\boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})\right]\right\}.$$

Note that both expressions $A_{n,\alpha}$ and $B_{n,\alpha}$ are i.i.d. averages of zero mean variables. By the uniform strong law of large numbers on compact sets (see [18, p. 108]), one can show that

$$\max_{\|\boldsymbol{t}^{(A)}\|\leqslant k} \|A_{n,\alpha}(\boldsymbol{t}^{(A)})\| \stackrel{\text{a.s.}}{\to} 0$$

and

$$\max_{\|\boldsymbol{t}^{(A)}\|\leqslant k} |B_{n,\alpha}(\boldsymbol{t}^{(A)})| \stackrel{\text{a.s.}}{\to} 0,$$

so that (6.5) is proved.

6.3. *Proof of Theorem 2.3*

Let $x_{n,j} = \int_{\mathcal{B}_j^{pq}} f(\boldsymbol{y}_n(\boldsymbol{t}))\, dG(\boldsymbol{t})$ and $x_j = \int_{\mathcal{B}_j^{pq}} f(\boldsymbol{y}(\boldsymbol{t}))\, dG(\boldsymbol{t})$. Define the mapping $h_j : C(\mathcal{B}_j^{pq}, \mathbb{C})^2 \to \mathbb{R}$ by $h_j(\boldsymbol{y}) = \int_{\mathcal{B}_j^{pq}} f(\boldsymbol{y}(\boldsymbol{t}))\, dG(\boldsymbol{t})$. Then, $x_{n,j} = h_j(r_j(\boldsymbol{y}_n))$ and $x_j = h_j(r_j(\boldsymbol{y}))$. By assumption, for all j, $r_j(\boldsymbol{y}_n) \stackrel{D}{\to} r_j(\boldsymbol{y})$. Thus, Theorem 5.1 of Billingsley [3] and the continuity of h_j imply $x_{n,j} \stackrel{D}{\to} x_j$ as $n \to \infty$, for all j. Also, the dominated convergence theorem yields $x_j \stackrel{D}{\to} w$, as $j \to \infty$. Using Theorem 4.2 of Billingsley [3], in order to establish $w_n \stackrel{D}{\to} w$ as $n \to \infty$, it suffices to show that for all $\varepsilon > 0$,

$$\lim_{j\to\infty} \limsup_{n\to\infty} P\{|x_{n,j} - w_n| \geqslant \varepsilon\} = 0.$$

By the Markov inequality it suffices to show that for some $\alpha > 0$,

$$\lim_{j\to\infty} \limsup_{n\to\infty} E|x_{n,j} - w_n|^{\alpha} = 0.$$

But

$$\begin{aligned} E|x_{n,j} - w_n|^{\alpha} &= E\left|\int_{\mathbb{R}^{pq}\setminus\mathcal{B}_j^{pq}} f(\boldsymbol{y}_n(\boldsymbol{t}))\, dG(\boldsymbol{t})\right|^{\alpha} \\ &\leqslant E\left[\int_{\mathbb{R}^{pq}\setminus\mathcal{B}_j^{pq}} |f(\boldsymbol{y}_n(\boldsymbol{t}))|\, dG(\boldsymbol{t})\right]^{\alpha} \\ &= E_{\boldsymbol{y}_n}\left(E_G\left[\mathbb{I}\{\|\boldsymbol{t}\| > j\}|f(\boldsymbol{y}_n(\boldsymbol{t}))|\right]\right)^{\alpha}. \end{aligned}$$

Then, using the Jensen inequality with $\alpha \geqslant 1$ and the Fubini theorem,

$$\begin{aligned} E|x_{n,j} - w_n|^{\alpha} &\leqslant E_{\boldsymbol{y}_n} E_G[\mathbb{I}\{\|\boldsymbol{t}\| > j\}|f(\boldsymbol{y}_n(\boldsymbol{t}))|^{\alpha}] \\ &= E_G E_{\boldsymbol{y}_n}[\mathbb{I}\{\|\boldsymbol{t}\| > j\}|f(\boldsymbol{y}_n(\boldsymbol{t}))|^{\alpha}] \\ &= \int_{\mathbb{R}^{pq}\setminus\mathcal{B}_j^{pq}} E|f(\boldsymbol{y}_n(\boldsymbol{t}))|^{\alpha}\, dG(\boldsymbol{t}). \end{aligned}$$

The theorem is proved. □

6.4. Proof of Theorem 2.4

Define the functional (norm) $|[x]|^2 = \int |x(\boldsymbol{t})|^2 \varphi_b(\boldsymbol{t})\, d\boldsymbol{t}$ on the subset of squared-integrable functions with respect to $\varphi_b(\boldsymbol{t})\, d\boldsymbol{t}$. Following Henze and Wagner [22, pp. 10–11], in order to prove the theorem, it suffices to show

$$\left(|[\breve{R}_{n,A}]|^2, |[\breve{R}_{n,B}]|^2\right) \xrightarrow{\mathcal{D}} \left(|[R_A]|^2, |[R_B]|^2\right) \tag{6.9}$$

and

$$|[\tilde{R}_{n,A} - \breve{R}_{n,A}]|^2 \xrightarrow{P} 0. \tag{6.10}$$

This is sufficient since the triangle inequality produces

$$\Big||[\breve{R}_{n,A}]| - |[\tilde{R}_{n,A}]|\Big| \leqslant |[\breve{R}_{n,A} - \tilde{R}_{n,A}]| \xrightarrow{P} 0.$$

This implies

$$\left(|[\tilde{R}_{n,A}]|^2, |[\tilde{R}_{n,B}]|^2\right) \xrightarrow{\mathcal{D}} \left(|[R_A]|^2, |[R_B]|^2\right).$$

But from (6.7), $\left|\hat{R}_{n,A}(\boldsymbol{t}) - \tilde{R}_{n,A}(\boldsymbol{t})\right| = \|\boldsymbol{t}\|^2 o_P(1)$ which yields

$$|[\hat{R}_{n,A} - \tilde{R}_{n,A}]|^2 \xrightarrow{P} 0.$$

Again, the triangle inequality produces

$$\Big||[\hat{R}_{n,A}]| - |[\tilde{R}_{n,A}]|\Big| \leqslant |[\hat{R}_{n,A} - \tilde{R}_{n,A}]| \xrightarrow{P} 0.$$

Then, the desired result

$$\left(|[\hat{R}_{n,A}]|^2, |[\hat{R}_{n,B}]|^2\right) \xrightarrow{\mathcal{D}} \left(|[R_A]|^2, |[R_B]|^2\right)$$

could be concluded.

The convergence in (6.9) is now proved by means of arbitrary linear combinations with the use of Theorem 2.3 with $\boldsymbol{y}_n = (\breve{R}_{n,A}, \breve{R}_{n,B})$, and the continuous function $f : \mathbb{C}^2 \to \mathbb{R}$ defined by $f(x_1, x_2) = a_1|x_1|^2 + a_2|x_2|^2$ for arbitrary constants a_1 and a_2. Theorem 2.1 states that $(\breve{R}_{n,A}, \breve{R}_{n,B}) \xrightarrow{\mathcal{D}} (R_A, R_B)$. From Proposition 14.6 of Kallenberg [25], this remains true on all the closed balls. Note that

$$E|f(\boldsymbol{y}_n(\boldsymbol{t}))| \leqslant E|a_1||\breve{R}_{n,A}(\boldsymbol{t})|^2 + E|a_2||\breve{R}_{n,B}(\boldsymbol{t})|^2.$$

It can be readily shown that $E|\breve{R}_{n,A}(\boldsymbol{t})|^2 = \prod_{k\in A}(1 - \phi^2(\boldsymbol{t}^{(k)}))$, which is independent of n and integrable with respect to $\varphi_b(\boldsymbol{t})\, d\boldsymbol{t}$. Hence, condition (2.7) is satisfied with $\alpha = 1$. The convergence in (6.9) thus holds.

To prove (6.10), proceed as follows. From (6.2), (6.3), (6.8) and the inequality $(\sum_{j=1}^{s} a_j)^2 \leqslant s \sum_{j=1}^{s} a_j^2$, it follows:

$$\begin{aligned}
&|\tilde{R}_{n,A}(\boldsymbol{t}) - \breve{R}_{n,A}(\boldsymbol{t})|^2 \\
&\leqslant |A| \sum_{\alpha \in A} \left[\frac{\|\boldsymbol{t}\|^2}{4} \left\| \frac{1}{\sqrt{n}p} \sum_{j=1}^{n} \sum_{k=1}^{p} (\boldsymbol{\epsilon}_j^{(k)} \boldsymbol{\epsilon}_j^{(k)\mathrm{T}} - \boldsymbol{I}) \right\|^2 \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 \right. \\
&\quad + \|\boldsymbol{t}\|^2 \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 o_P(1) + \|\boldsymbol{t}\|^2 n \, \|\bar{\boldsymbol{\epsilon}}\|^2 \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \\
&\quad \left. + \|\boldsymbol{t}\|^2 \left\| \sqrt{n}(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I}) \right\|^2 \|\bar{\boldsymbol{\epsilon}}\|^2 \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \right]. \qquad (6.11)
\end{aligned}$$

Let

$$\begin{aligned}
W_n &= \int \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \|\boldsymbol{t}\|^2 \varphi_b(\boldsymbol{t}) \, d\boldsymbol{t}, \\
\tilde{W}_n &= \int \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 \|\boldsymbol{t}\|^2 \varphi_b(\boldsymbol{t}) \, d\boldsymbol{t}.
\end{aligned}$$

By the Tonelli theorem,

$$E W_n = \int E \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \|\boldsymbol{t}\|^2 \varphi_b(\boldsymbol{t}) \, d\boldsymbol{t},$$

where by direct calculation

$$E \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 = \frac{1}{n} \prod_{k \in A; k \neq \alpha} (1 - \phi^2(\boldsymbol{t}^{(k)})), \qquad (6.12)$$

so that $W_n = o_P(1)$. Similarly, since

$$E \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 = \frac{p}{n} \prod_{k \in A; k \neq \alpha} (1 - \phi^2(\boldsymbol{t}^{(k)})), \qquad (6.13)$$

it follows that $\tilde{W}_n = o_P(1)$. Thus, $|[\tilde{R}_{n,A} - \breve{R}_{n,A}]|^2 \leqslant o_P(1)$. $\square$

6.5. Proof of Theorem 3.1

As in the proof of Theorem 2.1, it suffices to show that $r_j(S_{n,A}) \xrightarrow{fd} r_j(R_A)$ and $\{r_j(S_{n,A})\}$ is a tight family. We have

$$\begin{aligned}
S_{n,A}(\boldsymbol{t}) &= \frac{1}{\sqrt{n}} \sum_{l=1}^{pc_n} \prod_{k \in A} \left[\exp\left(\mathrm{i} \langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_l^{(k)} \rangle \right) - \phi(\boldsymbol{t}^{(k)}) \right] \\
&\quad + \frac{1}{\sqrt{n}} \sum_{l=pc_n+1}^{n-p+1} \prod_{k \in A} \left[\exp\left(\mathrm{i} \langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_l^{(k)} \rangle \right) - \phi(\boldsymbol{t}^{(k)}) \right] \\
&\equiv S_{n,A}^*(\boldsymbol{t}) + r_n(\boldsymbol{t}),
\end{aligned}$$

where $c_n = \left[\frac{n-p+1}{p}\right]$ is the integer part. The process $S_{n,A}$ is asymptotically equivalent to $S^*_{n,A}$ since

$$\rho_j(S_{n,A}, S^*_{n,A}) \leqslant \frac{1}{\sqrt{n}} \sum_{l=pc_n+1}^{n-p+1} 2^{|A|} \xrightarrow{P} 0.$$

Let

$$Y_l^A(\boldsymbol{t}) = \prod_{k\in A}\left[\exp\left(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_l^{(k)}\rangle\right) - \phi(\boldsymbol{t}^{(k)})\right], \quad l = 1, 2, \ldots,$$

which is an m-dependent sequence. Thus, using a multivariate central limit theorem for m-dependent complex random vectors obtained by applying the Cramér–Wold device in connection with a complex variant of Theorem 11 in Ferguson [18], the finite dimensional convergence to Gaussian limit and the covariance function are obtained. It remains to show that $\{r_j(S^*_{n,A})\}$ is a tight sequence. Since $\{1, 2, \ldots, pc_n\} = \cup_{h=1}^{p}\{pl + h;\ 0 \leqslant l < c_n\}$, we have $S^*_{n,A}(\boldsymbol{t}) = \sum_{h=1}^{p}\sum_{B\subset A}(-1)^{|A\setminus B|}\phi_{A,B}(\boldsymbol{t})Y_{n,h,B}(\boldsymbol{t})$, where $\phi_{A,B}(\boldsymbol{t}) = \prod_{l\in A\setminus B}\phi(\boldsymbol{t}^{(l)})$ and

$$Y_{n,h,B}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{l=0}^{c_n-1}\left[\prod_{j\in B}\exp\left(\mathrm{i}\langle \boldsymbol{t}^{(j)}, \boldsymbol{\epsilon}_{pl+h}^{(j)}\rangle\right) - \prod_{j\in B}\phi(\boldsymbol{t}^{(j)})\right].$$

To show that $\{r_j(Y_{n,h,B})\}$ is tight, proceed as in proof of Theorem 2.1. Since there is only a finite number of h's and B's, tightness follows and the theorem is proved. □

6.6. *Proof of Theorem 3.2*

Let

$$\tilde{S}_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p+1}\prod_{k\in A}[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})] + \mathrm{i}U_n(\boldsymbol{t}),$$

where $U_n(\boldsymbol{t})$ is defined as in (6.2) and where $\boldsymbol{\Delta}_j^{(k)} = (\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I})\boldsymbol{\epsilon}_j^{(k)} - \boldsymbol{S}^{-\frac{1}{2}}\bar{\boldsymbol{u}}$. Let also

$$\breve{S}_{n,A}(\boldsymbol{t}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n-p+1}\prod_{k\in A}\left[\exp(\mathrm{i}\langle \boldsymbol{t}^{(k)}, \boldsymbol{\epsilon}_j^{(k)}\rangle) - \phi(\boldsymbol{t}^{(k)})\right].$$

Proceed as in the proof of Theorem 2.2 to establish the analogues of (6.4), (6.5) and (6.6). The proofs of (6.4) and (6.5) are established along the lines in the proof of Theorem 2.2. Step (6.6) was proven in Theorem 3.1. □

6.7. Proof of Theorem 3.3

Follow the lines in the proof of Theorem 2.4 but use Theorem 3.1 instead of Theorem 2.1, and replace (6.11) with

$$
\begin{aligned}
&|\tilde{S}_{n,A}(\boldsymbol{t}) - \breve{S}_{n,A}(\boldsymbol{t})|^2 \\
&\quad \leqslant |A| \sum_{\alpha \in A} \left[\frac{\|\boldsymbol{t}\|^2}{4} \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}} - \boldsymbol{I}) \right\|^2 \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 \right. \\
&\qquad + \|\boldsymbol{t}\|^2 \left\| A_{n,\alpha}(\boldsymbol{t}^{(A)}) \right\|^2 o_P(1) + \|\boldsymbol{t}\|^2 n \, \|\overline{\boldsymbol{u}}\|^2 \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \\
&\qquad \left. + \|\boldsymbol{t}\|^2 \left\| \sqrt{n}(\boldsymbol{S}^{-\frac{1}{2}} - \boldsymbol{I}) \right\|^2 \|\overline{\boldsymbol{u}}\|^2 \left| B_{n,\alpha}(\boldsymbol{t}^{(A)}) \right|^2 \right].
\end{aligned}
$$

Replace also n with $n - p + 1$ in (6.12) and (6.13). □

Acknowledgments

M. Bilodeau thanks the Natural Sciences and Engineering Research Council of Canada for financial support through Grant 97303-99. This paper is part of P. Lafaye de Micheaux's Ph.D. Thesis. The authors are grateful for the comments of two referees which helped to improve the presentation.

References

[1] L. Baringhaus, N. Henze, A consistent test for multivariate normality based on the empirical characteristic function, Metrika 35 (1988) 339–348.
[2] R. Beran, P.W. Millar, Confidence sets for a multivariate distribution, Ann. Statist. 14 (1986) 431–443.
[3] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[4] J.R. Blum, J. Kiefer, M. Rosenblatt, Distribution free tests of independence based on the sample distribution function, Ann. Math. Statist. 32 (1961) 485–498.
[5] J.B. Conway, A Course in Functional Analysis, Springer, New York, 1985.
[6] R. Cools, Monomial cubature rules since "Stroud": a compilation. II. Numerical evaluation of integrals, J. Comput. Appl. Math. 112 (1999) 21–27.
[7] R. Cools, P. Rabinowitz, Monomial cubature rules since "Stroud": a compilation, J. Comput. Appl. Math. 48 (1993) 309–326.
[8] S. Csörgő, Limit behaviour of the empirical characteristic function, Ann. Probab. 9 (1981) 130–144.
[9] S. Csörgő, Multivariate empirical characteristic functions, Z. Wahrsch. Verw. Gebiete 55 (1981) 203–229.
[10] S. Csörgő, Testing for independence by the empirical characteristic function, J. Multivariate Anal. 16 (1985) 290–299.
[11] R.B. Davies, Numerical inversion of a characteristic function, Biometrika 60 (1973) 415–417.
[12] R.B. Davies, Algorithm AS 155: the distribution of a linear combination of χ^2 random variables, Appl. Statist. 29 (1980) 323–333.
[13] T. de Wet, R.H. Randles, On the effect of substituting parameter estimators in limiting $\chi 2$ U and V statistics, Ann. Statist. 15 (1987) 398–412.
[14] P. Deheuvels, An asymptotic decomposition for multivariate distribution-free tests of independence, J. Multivariate Anal. 11 (1981) 102–113.

[15] P. Deheuvels, Indépendance multivariée partielle et inégalités de Fréchet, in: M.C. Demetrescu, M. Iosifescu (Eds.), Studies in Probability and Related Topics, Nagard, Rome, 1983, pp. 145–155.
[16] P. Deheuvels, Tests of independence with exponential marginals, in: Goodness-of-fit tests and model validity (Paris, 2000), Statistical and Industrial Technology, Birkhäuser Boston, Boston, MA, 2002, pp. 463–476.
[17] P. Deheuvels, G.V. Martynov, Cramér–von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions, Comm. Statist. Theory Methods 25 (1996) 871–908.
[18] T.S. Ferguson, A Course in Large Sample Theory, Chapman & Hall, London, 1996.
[19] A. Feuerverger, A consistent test for bivariate dependence, Internat. Statist. Rev. 61 (1993) 419–433.
[20] A. Feuerverger, R.A. Mureika, The empirical characteristic function and its applications, Ann. Statist. 5 (1977) 88–97.
[21] K. Ghoudi, R.J. Kulperger, B. Rémillard, A nonparametric test of serial independence for time series and residuals, J. Multivariate Anal. 79 (2001) 191–218.
[22] N. Henze, T. Wagner, A new approach to the bhep tests for multivariate normality, J. Multivariate Anal. 62 (1997) 1–23.
[23] N. Henze, B. Zirkler, A class of invariant consistent tests for multivariate normality, Comm. Statist. Theory Methods 19 (1990) 3595–3617.
[24] J.P. Imhof, Computing the distribution of quadratic forms in normal variables, Biometrika 48 (1961) 419–426.
[25] O. Kallenberg, Foundations of Modern Probability, Springer, New York, 1997.
[26] J. Kellermeier, The empirical characteristic function and large sample hypothesis testing, J. Multivariate Anal. 10 (1980) 78–87.
[27] Y.-S. Lee, M.C. Lee, On the derivation and computation of the Cornish–Fisher expansion, Austral. J. Statist. 34 (1992) 443–450.
[28] Y.-S. Lee, T.-K. Lin, Algorithm AS 269: High order Cornish–Fisher expansion, Appl. Statist. 41 (1992) 233 –240.
[29] M.B. Marcus, Weak convergence of the empirical characteristic function, Ann. Probab. 9 (1981) 194–201.
[30] A.C. Rencher, Methods of multivariate analysis, second ed., Wiley, New York, 2002.
[31] A.H. Stroud, Some seventh degree integration formulas for symmetric regions, SIAM J. Numer. Anal. 4 (1967) 37–44.
[32] A.H. Stroud, Approximate Calculation of Multiple Integrals, Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, NJ, 1971.