Bernoulli **20**(3), 2014, 1344–1371
DOI: 10.3150/13-BEJ524

On the empirical multilinear copula process for count data

CHRISTIAN GENEST[1,*,**], JOHANNA G. NEŠLEHOVÁ[1,†,‡] and
BRUNO RÉMILLARD[2]

[1]*Department of Mathematics and Statistics, McGill University, 805, rue Sherbrooke ouest, Montréal (Québec), Canada H3A 0B9. E-mail:* [*]*cgenest@math.mcgill.ca;* [†]*johanna@math.mcgill.ca; url:* [**]*www.math.mcgill.ca/cgenest/;* [‡]*www.math.mcgill.ca/neslehova/*
[2]*Service de l'enseignement des méthodes quantitatives de gestion, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal (Québec), Canada H3T 2A7. E-mail: bruno.remillard@hec.ca; url: neumann.hec.ca/pages/bruno.remillard/*

Continuation refers to the operation by which the cumulative distribution function of a discontinuous random vector is made continuous through multilinear interpolation. The copula that results from the application of this technique to the classical empirical copula is either called the multilinear or the checkerboard copula. As shown by Genest and Nešlehová (*Astin Bull.* **37** (2007) 475–515) and Nešlehová (*J. Multivariate Anal.* **98** (2007) 544–567), this copula plays a central role in characterizing dependence concepts in discrete random vectors. In this paper, the authors establish the asymptotic behavior of the empirical process associated with the multilinear copula based on d-variate count data. This empirical process does not generally converge in law on the space $\mathcal{C}([0,1]^d)$ of continuous functions on $[0,1]^d$, equipped with the uniform norm. However, the authors show that the process converges in $\mathcal{C}(K)$ for any compact $K \subset \mathcal{O}$, where $\mathcal{O}$ is a dense open subset of $[0,1]^d$, whose complement is the Cartesian product of the ranges of the marginal distribution functions. This result is sufficient to deduce the weak limit of many functionals of the process, including classical statistics for monotone trend. It also leads to a powerful and consistent test of independence which is applicable even to sparse contingency tables whose dimension is sample size dependent.

Keywords: checkerboard copula; contingency table; count data; empirical process; Kendall's tau; mid-ranks; multilinear extension copula; Spearman's rho; test of independence

1. Introduction

This paper's central message is that there are advantages, both conceptual and technical, to viewing a contingency table as arising from a multivariate distribution having uniform margins on the unit interval, that is, a copula. As will be shown here, this approach leads to new statistical methodology that can be used to analyze tables that are sparse or whose number of categories grows with the sample size.

To go straight to the point, consider the simple case of a $K \times L$ contingency table derived from a random sample of size n of ordinal or interval responses in ordered categories $A_1 < \cdots < A_K$ and $B_1 < \cdots < B_L$. For arbitrary $k \in \{1, \ldots, K\}$ and $\ell \in \{1, \ldots, L\}$, let $f_{k\ell}$ be the relative frequency of the pair (A_k, B_ℓ) and denote by f_{k+} and $f_{+\ell}$ the row-wise and column-wise totals, respectively. Further set

$$F_{k+} = \sum_{i=1}^{k} f_{i+}, \qquad F_{+\ell} = \sum_{i=1}^{\ell} f_{+i},$$

and let $F_{0+} = F_{+0} = 0$. A density $\hat{c}_n^{\maltese}$ with respect to the Lebesgue measure can then be defined (almost everywhere) on $[0, 1]^2$ by setting

$$\hat{c}_n^{\maltese}(u, v) = \frac{f_{k\ell}}{f_{k+} f_{+\ell}}$$

whenever $u \in (F_{(k-1)+}, F_{k+})$ and $v \in (F_{+(\ell-1)}, F_{+\ell})$. As shown in Section 2, the corresponding distribution function $\widehat{C}_n^{\maltese}$ is a copula, that is, its margins are uniform on [0, 1]. Moreover, when $f_{k\ell} = f_{k+} f_{+\ell}$ for all $k \in \{1, \dots, K\}$ and $\ell \in \{1, \dots, L\}$, $\widehat{C}_n^{\maltese}$ becomes the independence copula Π whose Lebesgue density is identically equal to 1 on $[0, 1]^2$.

More significantly, several standard measures of association in the pair (X, Y), and classical tests of independence between X and Y, are based on $\widehat{C}_n^{\maltese}$. For example, Pearson's χ^2 statistic and the likelihood ratio statistic G^2 are immediately seen to satisfy

$$\chi^2 = n \sum_{k=1}^{K} \sum_{\ell=1}^{L} \frac{(f_{k\ell} - f_{k+} f_{\ell+})^2}{f_{k+} f_{\ell+}} = n \int_0^1 \int_0^1 \{\hat{c}_n^{\maltese}(u, v) - 1\}^2 \, \mathrm{d}v \, \mathrm{d}u, \tag{1.1}$$

$$G^2 = 2n \sum_{k=1}^{K} \sum_{\ell=1}^{L} f_{k\ell} \ln\left(\frac{f_{k\ell}}{f_{k+} f_{+\ell}}\right) = 2n \int_0^1 \int_0^1 \ln\{\hat{c}_n^{\maltese}(u, v)\} \, \mathrm{d}\widehat{C}_n^{\maltese}(u, v). \tag{1.2}$$

With some additional work (Nešlehová [14]) it can also be shown that the well-known Spearman and Kendall statistics for testing monotone trend (Agresti [1]) can be rewritten in terms of $\widehat{C}_n^{\maltese}$. Many other examples could be given.

The introduction of the *multilinear empirical copula* $\widehat{C}_n^{\maltese}$ in this context is not merely a neat way of unifying various known statistics for frequency data analysis. Because integral expressions such as (1.1) and (1.2) make sense even when the number of categories changes with n, $\widehat{C}_n^{\maltese}$ is rather a key tool for the investigation of new or existing procedures that can be used even in cases where the table is sparse or of varying dimension.

Further, it may be seen that when X and Y are continuous, $\widehat{C}_n^{\maltese}$ is a smoothed version of the classical empirical copula (Deheuvels [3]) from which it differs by at most a factor of $1/n$ uniformly. Statistical tools based on $\widehat{C}_n^{\maltese}$ can thus bridge the gap between continuous and discrete outcomes. In particular, the problems associated with ties, which invalidate many of the procedures developed for continuous data (Genest, Nešlehová and Ruppert [8]), are then automatically taken care of. While it seems intuitively reasonable to base inference on $\widehat{C}_n^{\maltese}$, this new approach generally requires the knowledge of its limit $C^{\maltese}$ and the asymptotic behavior of the corresponding empirical process

$$\widehat{\mathbb{C}}_n^{\maltese} = \sqrt{n}\left(\widehat{C}_n^{\maltese} - C^{\maltese}\right), \tag{1.3}$$

which has hitherto never been studied in the literature.

This paper contributes to the problem by determining the asymptotic behavior of the process (1.3) in general dimension $d \geq 2$ when the components of the underlying random vector $\mathbf{X} = (X_1, \ldots, X_d)$ are either integer-valued or strictly increasing transformations thereof. As will be seen, $\widehat{C}_n^{\maltese}$ is a consistent estimator of the so-called multilinear extension (or checkerboard) copula $C^{\maltese}$ of $\mathbf{X}$. This limiting copula, defined in Section 2, has been studied earlier, for example, by Genest and Nešlehová [7] and Nešlehová [14], who showed that it captures many important dependence properties of $\mathbf{X}$ when $d = 2$. In particular, when the components of $\mathbf{X}$ are independent, $C^{\maltese}$ is the independence copula Π.

The main result, stated in Section 3, gives the asymptotic behavior of the process (1.3). Unless the components of $\mathbf{X}$ are mutually independent, $\widehat{\mathbb{C}}_n^{\maltese}$ does not generally converge on the space $\mathcal{C}([0,1]^d)$ of continuous functions on $[0,1]^d$ equipped with the uniform norm because $C^{\maltese}$ has discontinuous partial derivatives. Fortunately, $\widehat{\mathbb{C}}_n^{\maltese}$ converges – without any regularity conditions – in the subspace $\mathcal{C}(K)$ for any compact subset $K \subset \mathcal{O}$, where $\mathcal{O}$ is a dense open subset of $[0,1]^d$ whose complement is the Cartesian product of the ranges of the marginal distribution functions. The proof of the main result is involved; it is outlined in Section 4 and detailed in the Appendix.

To illustrate the usefulness of the process (1.3) for inference, Section 5 provides a few initial examples of application. It is first shown that the main result is sufficient to deduce the limiting distribution of classical statistics for monotone trend such as Spearman's rho and Kendall's tau. Moreover, a new and consistent Cramér–von Mises type test of independence is proposed that can be used whatever the margins. As illustrated through a small simulation study, it performs very well even for sparse contingency tables whose dimension is sample size dependent; in all cases considered, it is consistently more powerful than the classical chi-squared test. Section 6 concludes.

2. The multilinear extension copula

Suppose that $\mathbf{X} = (X_1, \ldots, X_d)$ is a vector of discrete random variables with joint cumulative distribution function H. For each $j \in \{1, \ldots, d\}$, let F_j denote the distribution function of X_j and assume that there exists a strictly increasing function $A_j \colon \mathbb{N} \to \mathbb{R}$ such that $\operatorname{supp}(X_j) \subseteq \{A_j(k) \colon k \in \mathbb{N}\}$. Note that the inclusion may be strict; in particular, it is not assumed that $\Pr\{X_j = A_j(k)\} > 0$ holds for all $k \in \mathbb{N}$ or that the support of X_j is infinite. Furthermore, observe that the closure of the range of F_j, viz. $\mathcal{R}_j = \{0, 1, F_j\{A_j(0)\}, F_j\{A_j(1)\}, \ldots\}$, defines a partition of $[0,1]$. In what follows, $A_j(-1) = A_j(0) - 1$ for all $j \in \{1, \ldots, d\}$ by convention.

Definition 2.1. *The multilinear extension copula $C^{\maltese}$ of H is the unique copula whose density with respect to the Lebesgue measure is given by*

$$c^{\maltese}(u_1, \ldots, u_d) = \frac{\Pr\{X_1 = A_1(k_1), \ldots, X_d = A_d(k_d)\}}{\Pr\{X_1 = A_1(k_1)\} \times \cdots \times \Pr\{X_d = A_d(k_d)\}}$$

whenever for all $j \in \{1, \ldots, d\}$, $F_j\{A_j(k_j - 1)\} < u_j \leq F_j\{A_j(k_j)\}$ for some $k_j \in \mathbb{N}$.

An explicit form of $C^{\maltese}$, which is easily verified by differentiation, is given in Proposition 2.1 below. For each $j \in \{1,\dots,d\}$ and $u \in [0,1]$, let u_j^- and u_j^+ be, respectively, the greatest and the least element of $\mathcal{R}_j$ such that $u_j^- \le u \le u_j^+$. Further let

$$\lambda_{F_j}(u) = \begin{cases} (u - u_j^-)/(u_j^+ - u_j^-), & \text{if } u_j^- \neq u_j^+, \\ 1, & \text{otherwise.} \end{cases}$$

Thus when $k \in \mathbb{N}$ is such that $\Delta F_j\{A_j(k)\} = \Pr\{X_j = A_j(k)\} > 0$, then for all $u \in (F_j\{A_j(k-1)\}, F_j\{A_j(k)\})$, one has $u_j^- = F_j\{A_j(k-1)\}$, $u_j^+ = F_j\{A_j(k)\}$ and

$$\lambda_{F_j}(u) = \frac{u - F_j\{A_j(k-1)\}}{\Delta F_j\{A_j(k)\}}.$$

Furthermore, if F_j^{-1} is the pseudo-inverse of F_j, then $F_j \circ F_j^{-1}(u_j^-) = F_j\{A_j(k-1)\}$ and $F_j \circ F_j^{-1}(u_j^+) = F_j\{A_j(k)\}$. Finally, for any $S \subset \{1,\dots,d\}$ and $u_1,\dots,u_d \in [0,1]$, set

$$\lambda_{H,S}(u_1,\dots,u_d) = \prod_{\ell \in S} \lambda_{F_\ell}(u_\ell) \prod_{\ell \notin S} \{1 - \lambda_{F_\ell}(u_\ell)\},$$

which depends on H only through its margins $F_1,\dots,F_d$.

Proposition 2.1. *The multilinear extension copula $C^{\maltese}$ of H is given by*

$$C^{\maltese}(u_1,\dots,u_d) = \sum_{S \subseteq \{1,\dots,d\}} \lambda_{H,S}(u_1,\dots,u_d) H\{F_1^{-1}(u_{S_1}),\dots,F_d^{-1}(u_{S_d})\},$$

where for each $j \in \{1,\dots,d\}$, $u_{S_j} = u_j^+$ if $j \in S$ and $u_{S_j} = u_j^-$ otherwise. In particular, $C^{\maltese}(u_{S_1},\dots,u_{S_d}) = H\{F_1^{-1}(u_{S_1}),\dots,F_d^{-1}(u_{S_d})\}$ for any $S \subseteq \{1,\dots,d\}$.

It is easily seen that $C^{\maltese}$ satisfies Sklar's representation, that is, for all $x_1,\dots,x_d \in \mathbb{R}$,

$$H(x_1,\dots,x_d) = C^{\maltese}\{F_1(x_1),\dots,F_d(x_d)\}.$$

This is because in effect, this identity needs only be verified if for all $j \in \{1,\dots,d\}$, $x_j = A_j(k_j)$ for some $k_j \in \mathbb{N}$ such that $\Delta F_j\{A_j(k_j)\} > 0$. In fact, $C^{\maltese}$ is precisely the construction used to extend a sub-copula to a copula in the proof of Sklar's theorem; see, for example, Nelsen [13] for details in the bivariate case.

The copula $C^{\maltese}$ is known to capture many important dependence properties of H, as summarized by Genest and Nešlehová [7]. As shown by Nešlehová ([14], Corollary 6), $C^{\maltese}$ is invariant with respect to strictly increasing transformations of the margins.

Now consider a random sample $\mathcal{X} = \{(X_{11},\dots,X_{1d}),\dots,(X_{n1},\dots,X_{nd})\}$ from H and let H_n be the corresponding empirical distribution function. Because H_n is itself a discrete distribution, one can define its multilinear extension copula $\widehat{C}_n^{\maltese}$ and its corresponding density $\hat{c}_n^{\maltese}$ with respect to the Lebesgue measure as above. To be explicit, fix $j \in \{1,\dots,d\}$ and denote by

$A_{nj}(0) < \cdots < A_{nj}(n_j)$ the distinct values of $X_{1j}, \ldots, X_{nj}$. Let also $A_{nj}(-1) = A_{nj}(0) - 1$. The range $\mathcal{R}_{nj}$ of F_{nj} then consists of

$$0 = F_{nj}\{A_{nj}(-1)\} < F_{nj}\{A_{nj}(0)\} < \cdots < 1 = F_{nj}\{A_{nj}(n_j)\}.$$

If $(u_1, \ldots, u_d)$ is such that for all $j \in \{1, \ldots, d\}$, $F_{nj}\{A_{nj}(k_j - 1)\} < u_j \leq F_{nj}\{A_{nj}(k_j)\}$ for some $k_j \in \{0, \ldots, n_j\}$, then

$$\hat{c}_n^{\maltese}(u_1, \ldots, u_d) = \frac{h_n\{A_{n1}(k_1), \ldots, A_{nd}(k_d)\}}{\Delta F_{n1}\{A_{n1}(k_1)\} \times \cdots \times \Delta F_{nd}\{A_{nd}(k_d)\}},$$

whose numerator is the proportion of data with $X_{ij} = A_{nj}(k_j)$ for $j \in \{1, \ldots, d\}$, and

$$\widehat{C}_n^{\maltese}(u_1, \ldots, u_d) = \sum_{S \subseteq \{1, \ldots, d\}} \lambda_{H_n, S}(u_1, \ldots, u_d) H_n\{F_{n1}^{-1}(u_{S_1}), \ldots, F_{nd}^{-1}(u_{S_d})\}.$$

Observe that $\hat{c}_n^{\maltese}$ and $\widehat{C}_n^{\maltese}$ are both functions of the component-wise ranks. As announced in the Introduction, $\widehat{C}_n^{\maltese}$ is a consistent estimator of the multilinear extension copula $C^{\maltese}$ of H. This fact will be a consequence of this paper's main result, Theorem 3.1, which characterizes the limit of the process $\widehat{\mathbb{C}}_n^{\maltese}$ defined in (1.3).

Remark 2.1. When $X_1, \ldots, X_d$ are continuous, $\widehat{C}_n^{\maltese}$ was actually used by Deheuvels [4] to construct tests of independence. It is then asymptotically equivalent to the empirical copula $\widehat{C}_n$ given, for all $u_1, \ldots, u_d \in [0, 1]$, by

$$\widehat{C}_n(u_1, \ldots, u_d) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{F_{n1}(X_{i1}) \leq u_1, \ldots, F_{nd}(X_{id}) \leq u_d\}.$$

Indeed, if for all $j \in \{1, \ldots, d\}$, $F_{nj}\{A_{nj}(k_j - 1)\} \leq u_j < F_{nj}\{A_{nj}(k_j)\}$ for some $k_j \in \{0, \ldots, n_j\}$, then $\widehat{C}_n(u_1, \ldots, u_d) = H_n\{A_{n1}(k_1 - 1), \ldots, A_{nd}(k_d - 1)\}$. Because the coefficients $\lambda_{H_n,S}$ are non-negative and add up to 1 by the multinomial formula, the fact that H_n is non-decreasing component-wise implies that

$$\widehat{C}_n(u_1, \ldots, u_d) \leq \widehat{C}_n^{\maltese}(u_1, \ldots, u_d) \leq H_n\{A_{n1}(k_1), \ldots, A_{nd}(k_d)\}.$$

Hence, $|\widehat{C}_n(u_1, \ldots, u_d) - \widehat{C}_n^{\maltese}(u_1, \ldots, u_d)|$ is bounded above by

$$\begin{aligned} & H_n\{A_{n1}(k_1), \ldots, A_{nd}(k_d)\} - H_n\{A_{n1}(k_1 - 1), \ldots, A_{nd}(k_d - 1)\} \\ & \quad \leq \sum_{j=1}^{d} \left| F_{nj}\{A_{nj}(k_j)\} - F_{nj}\{A_{nj}(k_j - 1)\} \right|, \end{aligned}$$

from which it follows that $\|\widehat{C}_n - \widehat{C}_n^{\maltese}\| \leq d/n$ almost surely. This also implies that $\widehat{C}_n^{\maltese}$ is asymptotically equivalent to other versions of the empirical copula commonly used in the literature; see, for example, Fermanian, Radulović and Wegkamp [6].

To ease the notation, it will be assumed henceforth, without loss of generality, that $X_1, \dots, X_d$ are integer-valued. In this case, one has the following alternative representation of $C^{\maltese}$, which is useful to study the process (1.3).

Proposition 2.2. *Let $(X_1, \dots, X_d)$ be a random vector in $\mathbb{N}^d$ with distribution function H. Let also $U_1, \dots, U_d$ be independent standard uniform random variables, independent of $(X_1, \dots, X_d)$. Then $C^{\maltese}$ is the unique copula of the distribution function $H^{\maltese}$ of $(X_1 + U_1 - 1, \dots, X_d + U_d - 1)$ with margins $F_1^{\maltese}, \dots, F_d^{\maltese}$, that is, for all $u_1, \dots, u_d \in [0, 1]$,*

$$C^{\maltese}(u_1, \dots, u_d) = H^{\maltese}\{F_1^{\maltese-1}(u_1), \dots, F_d^{\maltese-1}(u_d)\}.$$

Given an empirical distribution function H_n based on a random sample from a multivariate integer-valued distribution H, one can proceed as in Proposition 2.2 to define a multilinear extension $H_n^{\maltese}$ whose margins $F_{n1}^{\maltese}, \dots, F_{nd}^{\maltese}$ are continuous extensions of the margins $F_{n1}, \dots, F_{nd}$ of H_n. Furthermore,

$$\widehat{C}_n^{\maltese}(u_1, \dots, u_d) = H_n^{\maltese}\{F_{n1}^{\maltese-1}(u_1), \dots, F_{nd}^{\maltese-1}(u_d)\}$$

holds for all $u_1, \dots, u_d \in [0, 1]$, which will come in handy in Section 3.

3. The empirical multilinear copula process

In what follows, $\mathcal{C}(K)$ stands for the space of all continuous functions from a compact set $K \subseteq [0, 1]^d$ to $\mathbb{R}$ equipped with the uniform norm, that is, $\|f\|_K = \sup\{|f(u_1, \dots, u_d)|: (u_1, \dots, u_d) \in K\}$. When $K = [0, 1]^d$, the index on $\|\cdot\|$ is suppressed. Similarly, let $\ell^\infty(K)$ denote the space of all bounded functions from K to $\mathbb{R}$ equipped with the uniform norm. For each $j \in \{1, \dots, d\}$ and all $u_1, \dots, u_d \in (0, 1)$ where the partial derivatives exist, set

$$\dot{C}_j^{\maltese}(u_1, \dots, u_d) = \frac{\partial}{\partial u_j} C^{\maltese}(u_1, \dots, u_d).$$

Furthermore, let $\mathbb{B}_{C^{\maltese}}$ be a $C^{\maltese}$-Brownian bridge, that is, a centred Gaussian process on $[0, 1]^d$ with covariance given, for all $s_1, \dots, s_d, t_1, \dots, t_d \in [0, 1]$, by

$$C^{\maltese}(s_1 \wedge t_1, \dots, s_d \wedge t_d) - C^{\maltese}(s_1, \dots, s_d)C^{\maltese}(t_1, \dots, t_d).$$

Here, $a \wedge b = \min(a, b)$ for arbitrary $a, b \in \mathbb{R}$. The limit of $\widehat{\mathbb{C}}_n^{\maltese}$ can be expressed in terms of a transformation of $\mathbb{B}_{C^{\maltese}}$ involving the following operator.

Definition 3.1. *Let H be a multivariate distribution function with support included in $\mathbb{N}^d$ and margins $F_1, \dots, F_d$. The multilinear interpolation operator $\mathfrak{M}_H$: $\ell^\infty([0, 1]^d) \to \ell^\infty([0, 1]^d)$: $g \mapsto \mathfrak{M}_H(g)$ is defined, for every $g \in \ell^\infty([0, 1]^d)$, by*

$$\mathfrak{M}_H(g)(u_1, \dots, u_d) = \sum_{S \subseteq \{1, \dots, d\}} \lambda_{H,S}(u_1, \dots, u_d) g(u_{S_1}, \dots, u_{S_d}).$$

As was the case with $\lambda_{H,S}$, the operator $\mathfrak{M}_H$ depends on H only through its margins. Although the paths of the process $\widehat{\mathbb{C}}_n^{\maltese}$ are continuous on $[0,1]^d$ for every n, it cannot possibly converge in $\mathcal{C}([0,1]^d)$ in general. This is because unless $C^{\maltese} = \Pi$, its partial derivatives exist only on the open set

$$\mathcal{O} = \bigcup_{(k_1,\ldots,k_d)\in\mathbb{N}^d} \big(F_1(k_1-1), F_1(k_1)\big) \times \cdots \times \big(F_d(k_d-1), F_d(k_d)\big).$$

Fortunately, the convergence of $\widehat{\mathbb{C}}_n^{\maltese}$ can be established in $\mathcal{C}(K)$ for any compact $K \subset \mathcal{O}$. The symbol $\rightsquigarrow$ is used henceforth to denote weak convergence.

Theorem 3.1. *Let $\mathbb{C}^{\maltese} = \mathfrak{M}_H(\mathbb{B}_{C^{\maltese}})$ and let K be any compact subset of $\mathcal{O}$. Then, as $n \to \infty$, $\widehat{\mathbb{C}}_n^{\maltese} \rightsquigarrow \widehat{\mathbb{C}}^{\maltese}$ in $\mathcal{C}(K)$, where, for all $(u_1,\ldots,u_d) \in \mathcal{O}$,*

$$\widehat{\mathbb{C}}^{\maltese}(u_1,\ldots,u_d) = \mathbb{C}^{\maltese}(u_1,\ldots,u_d) - \sum_{j=1}^{d} \dot{C}_j^{\maltese}(u_1,\ldots,u_d)\mathbb{C}^{\maltese}(1,\ldots,1,u_j,1,\ldots,1).$$

This theorem can be strengthened when $X_1,\ldots,X_d$ are mutually independent, which is the case if and only if $C^{\maltese}$ is the independence copula Π.

Corollary 3.1. *Suppose that $C^{\maltese} = \Pi$. Then, as $n \to \infty$, $\widehat{\mathbb{C}}_n^{\maltese} \rightsquigarrow \widehat{\mathbb{C}}^{\maltese}$ in $\mathcal{C}([0,1]^d)$.*

Remark 3.1. When $X_1,\ldots,X_d$ are continuous, $C^{\maltese} = C$ is the unique copula of H and $\widehat{C}_n^{\maltese}$ is asymptotically equivalent to the empirical copula $\widehat{C}_n$ by Remark 2.1. Rüschendorf [16] showed that under suitable regularity conditions on C, $\widehat{\mathbb{C}}_n^{\maltese} \rightsquigarrow \widehat{\mathbb{C}}$ as $n \to \infty$, where $\widehat{\mathbb{C}}$ is defined in terms of a C-Brownian bridge $\mathbb{B}_C$, for all $u_1,\ldots,u_d \in [0,1]$, by

$$\widehat{\mathbb{C}}(u_1,\ldots,u_d) = \mathbb{B}_C(u_1,\ldots,u_d) - \sum_{j=1}^{d} \dot{C}_j(u_1,\ldots,u_d)\mathbb{B}_C(1,\ldots,1,u_j,1,\ldots,1).$$

This result has since been refined in various ways; see Segers [18] and references therein.

4. Proof of Theorem 3.1

The proof of the main result is quite involved. It rests on a series of steps and propositions that are described below. All proofs may be found in Appendix B.

Because $C^{\maltese}$ is a copula of H, it can be assumed without loss of generality that the sample $\mathcal{X}$ from H arises from a random sample $\mathcal{V} = \{(V_{11},\ldots,V_{1d}),\ldots,(V_{n1},\ldots,V_{nd})\}$ from $C^{\maltese}$, that is, for every $i \in \{1,\ldots,n\}$, one has $X_{i1} = F_1^{-1}(V_{i1}),\ldots,X_{id} = F_d^{-1}(V_{id})$. If B_n denotes the empirical distribution function of this latent sample $\mathcal{V}$, it is well known that as $n \to \infty$, the corresponding empirical process $\mathbb{B}_n = \sqrt{n}(B_n - C^{\maltese})$ converges weakly in $\ell^\infty([0,1]^d)$ to the $C^{\maltese}$-Brownian bridge $\mathbb{B}_{C^{\maltese}}$ (van der Vaart and Wellner [19]).

The first step consists of considering the case where the margins of H are known. In contrast to the continuous case, the variables $F_1(X_1), \dots, F_d(X_d)$ are not uniform and their joint distribution function D is not a copula. Observe that $C^{\maltese} = \mathfrak{M}_H(D)$ and introduce $C_n^{\maltese} = \mathfrak{M}_H(D_n)$, where D_n denotes the empirical distribution function of the transformed data $(F_1(X_{11}), \dots, F_d(X_{1d})), \dots, (F_1(X_{n1}), \dots, F_d(X_{nd}))$. Note that $C_n^{\maltese}$ cannot be computed in practice, because it relies on the unknown marginal distribution functions. As is easily seen by differentiation, $C_n^{\maltese}$ is a continuous distribution function on $[0, 1]^d$ whose jth margin is given, for all $u \in [0, 1]$, by

$$C_{nj}^{\maltese}(u) = \lambda_{F_j}(u) D_{nj}\big(u^+\big) + \big\{1 - \lambda_{F_j}(u)\big\} D_{nj}\big(u^-\big).$$

Because its margins are not uniform, $C_n^{\maltese}$ is not a copula. The following proposition shows that the empirical process $\mathbb{C}_n^{\maltese} = \sqrt{n}(C_n^{\maltese} - C^{\maltese})$ converges. Its proof rests on the fact that $\mathfrak{M}_H$ is a continuous linear contraction. This is because the weights $\lambda_{H,S}$ are non-negative and add up to 1, so that for any $g, g^* \in \ell^\infty([0, 1]^d)$, one has $\|\mathfrak{M}_H(g) - \mathfrak{M}_H(g^*)\| \le \|g - g^*\|$.

Proposition 4.1. *As $n \to \infty$, $\mathbb{C}_n^{\maltese} \rightsquigarrow \mathbb{C}^{\maltese} = \mathfrak{M}_H(\mathbb{B}_{C^{\maltese}})$ in $\mathcal{C}([0, 1]^d)$.*

Next, the process $\widehat{\mathbb{C}}_n^{\maltese}$ in which margins are unknown can be written in the form

$$\widehat{\mathbb{C}}_n^{\maltese} = \widetilde{\mathbb{C}}_n^{\maltese} + \widetilde{\mathbb{D}}_n, \tag{4.1}$$

where the summands are defined, for all $u_1, \dots, u_d \in [0, 1]$, by

$$\widetilde{\mathbb{C}}_n^{\maltese}(u_1, \dots, u_d) = \sqrt{n}\big[H_n^{\maltese}\big\{F_{n1}^{\maltese-1}(u_1), \dots, F_{nd}^{\maltese-1}(u_d)\big\} - H^{\maltese}\big\{F_{n1}^{\maltese-1}(u_1), \dots, F_{nd}^{\maltese-1}(u_d)\big\}\big]$$

and

$$\widetilde{\mathbb{D}}_n(u_1, \dots, u_d) = \sqrt{n}\big[C^{\maltese}\big\{F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \dots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\big\} - C^{\maltese}(u_1, \dots, u_d)\big].$$

The next proposition shows that $\widetilde{\mathbb{C}}_n^{\maltese}$ has the same asymptotic behavior as $\mathbb{C}_n^{\maltese}$.

Proposition 4.2. *As $n \to \infty$, $\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\| \xrightarrow{\text{p}} 0$.*

Next, one needs to determine the limit of the second summand in (4.1). The following result first shows that $\widetilde{\mathbb{D}}_n$ has the same asymptotic behavior as that of the auxiliary process $\mathbb{D}_n$ defined, for all $u_1, \dots, u_d \in [0, 1]$, by

$$\mathbb{D}_n(u_1, \dots, u_d) = \sqrt{n}\left[C^{\maltese}\left\{u_1 - \frac{\mathbb{C}_{n1}^{\maltese}(u_1)}{\sqrt{n}}, \dots, u_d - \frac{\mathbb{C}_{nd}^{\maltese}(u_d)}{\sqrt{n}}\right\} - C^{\maltese}(u_1, \dots, u_d)\right],$$

where $\mathbb{C}_{n1}^{\maltese}, \dots, \mathbb{C}_{nd}^{\maltese}$ are the margins of $\mathbb{C}_n^{\maltese}$.

Proposition 4.3. *As $n \to \infty$, $\|\mathbb{D}_n - \widetilde{\mathbb{D}}_n\| \xrightarrow{\text{p}} 0$.*

Finally, fix an arbitrary compact subset K of $\mathcal{O}$ and consider the mapping $\mathfrak{D}_K \colon \mathcal{C}([0,1]^d) \to \mathcal{C}(K)$ defined, for all $g \in \mathcal{C}([0,1]^d)$ and $(u_1, \ldots, u_d) \in K$, by

$$\mathfrak{D}_K(g)(u_1, \ldots, u_d) = -\sum_{j=1}^{d} \dot{C}_j^{\maltese}(u_1, \ldots, u_d) g(1, \ldots, 1, u_j, 1, \ldots, 1).$$

This mapping is clearly linear and continuous because for any $g, g^* \in \mathcal{C}([0,1]^d)$,

$$\left\|\mathfrak{D}_K(g) - \mathfrak{D}_K\left(g^*\right)\right\| \le \sum_{j=1}^{d} \dot{C}_j^{\maltese}(u_1, \ldots, u_d) \left\|g - g^*\right\| \le d\left\|g - g^*\right\|.$$

For, when they exist, the partial derivatives of any copula take values in $[0,1]$. The Continuous Mapping theorem then implies that, as $n \to \infty$, $\mathfrak{D}_K(\mathbb{C}_n^{\maltese}) \rightsquigarrow \mathfrak{D}_K(\mathbb{C}^{\maltese})$ in $\mathcal{C}(K)$. As shown next, the difference between $\mathbb{D}_n$ and $\mathfrak{D}_K(\mathbb{C}_n^{\maltese})$ is asymptotically negligible.

Proposition 4.4. *As $n \to \infty$, $\|\mathbb{D}_n - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})\|_K \xrightarrow{\mathrm{p}} 0$ for any compact $K \subset \mathcal{O}$.*

To complete the proof of Theorem 3.1, let K be any compact subset of $\mathcal{O}$. Combining Propositions 4.1–4.4, one finds that, as $n \to \infty$,

$$\left\|\widehat{\mathbb{C}}_n^{\maltese} - \mathbb{C}_n^{\maltese} - \mathfrak{D}_K\left(\mathbb{C}_n^{\maltese}\right)\right\|_K \xrightarrow{\mathrm{p}} 0.$$

The Continuous Mapping theorem can then be invoked together with Proposition 4.1 to conclude that $\widehat{\mathbb{C}}_n^{\maltese} \rightsquigarrow \widehat{\mathbb{C}}^{\maltese} = \mathbb{C}^{\maltese} + \mathfrak{D}_K(\mathbb{C}^{\maltese})$. To establish Corollary 3.1, first note that when $C^{\maltese} = \Pi$, $\dot{C}_j^{\maltese}$ is continuous on $[0,1]^d$ for all $j \in \{1, \ldots, d\}$. One can then define $\mathfrak{D}$ as $\mathfrak{D}_K$ with $K = [0,1]^d$ and use the following result to conclude.

Proposition 4.5. *When $C^{\maltese} = \Pi$, $\|\mathbb{D}_n - \mathfrak{D}(\mathbb{C}_n^{\maltese})\| \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$.*

Remark 4.1. Although the process $\widehat{\mathbb{C}}_n^{\maltese}$ fails to converge on $\mathcal{C}([0,1]^d)$ in general, the sequence $\|\widehat{\mathbb{C}}_n^{\maltese}\|$ is tight. Indeed, the definition of $\mathbb{D}_n$ and the Lipschitz property of $C^{\maltese}$ imply that $\|\mathbb{D}_n\| \le \|\mathbb{C}_{n1}^{\maltese}\| + \cdots + \|\mathbb{C}_{nd}^{\maltese}\| \le d\|\mathbb{C}_n^{\maltese}\|$. From (4.1) and the triangle inequality,

$$\left\|\widehat{\mathbb{C}}_n^{\maltese}\right\| \le (d+1)\left\|\mathbb{C}_n^{\maltese}\right\| + \left\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\right\| + \|\mathbb{D}_n - \widetilde{\mathbb{D}}_n\|. \tag{4.2}$$

The result thus follows because the three summands form tight sequences. Indeed, $\mathbb{C}_n^{\maltese}$ converges weakly in $\mathcal{C}([0,1]^d)$ by Proposition 4.1 and the other two terms converge in probability to 0 by Propositions 4.2 and 4.3, respectively. It is further of interest to observe that because $\|\dot{C}_j^{\maltese}\|_{\mathcal{O}} \le 1$ for all $j \in \{1, \ldots, d\}$, one has $\|\widehat{\mathbb{C}}^{\maltese}\|_{\mathcal{O}} \le (d+1)\|\mathbb{C}^{\maltese}\|$.

Finally, note that $\widehat{C}_n^{\maltese}$ is a uniformly consistent estimator of $C^{\maltese}$. This follows immediately from (4.2), the Continuous Mapping theorem and Slutsky's lemma.

Corollary 4.1. *As $n \to \infty$, $\|\widehat{C}_n^{\maltese} - C^{\maltese}\| \xrightarrow{\mathrm{p}} 0$.*

5. Applications

Theorem 3.1 characterizes the weak limit of the empirical process $\widehat{\mathbb{C}}_n^{\maltese}$ in $\mathcal{C}(K)$ for any compact subset K of $\mathcal{O}$. To illustrate the usefulness of this result for inference, a few initial examples of application are provided below. They pertain to classical statistics for monotone trend and tests of independence, respectively.

5.1. Tests of monotone trend

Kendall's tau and Spearman's rho are two classical measures of monotone trend for two-way cross-classifications of ordinal or interval data. As described, for example, in Agresti [1], powerful tests of independence can be based on these statistics. Both of them are functions of (mid-) ranks that can be expressed as functionals of $\widehat{C}_n^{\maltese}$ (Nešlehová [14]).

Given a random sample $\mathcal{X} = \{(X_{11}, X_{12}), \ldots, (X_{n1}, X_{n2})\}$ from a bivariate distribution function H, let R_{ij} denote the component-wise mid-rank of X_{ij} for $i \in \{1, \ldots, n\}$ and $j \in \{1, 2\}$. Let also a_n and b_n, respectively, represent the number of strictly concordant and discordant pairs in the sample. The non-normalized versions of Kendall's and Spearman's coefficients then satisfy

$$\tau_n = \frac{a_n - b_n}{\binom{n}{2}} = \frac{n-1}{n}\left\{-1 + 4\int_0^1\int_0^1 \widehat{C}_n^{\maltese}(u,v)\,\mathrm{d}\widehat{C}_n^{\maltese}(u,v)\right\},$$

$$\rho_n = \frac{12}{n^3}\sum_{i=1}^n\left(R_{i1} - \frac{n+1}{2}\right)\left(R_{i2} - \frac{n+1}{2}\right) = 12\int_{[0,1]^2}\left\{\widehat{C}_n^{\maltese}(u,v) - uv\right\}\mathrm{d}\Pi(u,v).$$

It is immediate from Corollary 4.1 that τ_n and ρ_n are consistent estimators of

$$\tau = -1 + 4\int_0^1\int_0^1 C^{\maltese}(u,v)\,\mathrm{d}C^{\maltese}(u,v), \qquad \rho = 12\int_{[0,1]^2}\left\{C^{\maltese}(u,v) - uv\right\}\mathrm{d}\Pi(u,v).$$

It is well known that τ_n is a U-statistic and hence asymptotically Gaussian (Lee [11]). Its limiting behavior can also be deduced from Theorem 3.1. To see this, first call on Hoeffding's identity (Nelsen [13], Corollary 5.1.2) to write

$$\int_{[0,1]^2} C^{\maltese}(u,v)\,\mathrm{d}\widehat{C}_n^{\maltese}(u,v) = \int_{[0,1]^2}\widehat{C}_n^{\maltese}(u,v)\,\mathrm{d}C^{\maltese}(u,v).$$

Given that $\widehat{C}_n^{\maltese}$ and $C^{\maltese}$ are absolutely continuous with respect to the Lebesgue measure, the fact that the complement of $\mathcal{O}$ in $[0,1]^d$ has Lebesgue measure 0 then implies that

$$\begin{aligned}&\sqrt{n}\left\{\int_{[0,1]^2}\widehat{C}_n^{\maltese}(u,v)\,\mathrm{d}\widehat{C}_n^{\maltese}(u,v) - \int_{[0,1]^2} C^{\maltese}(u,v)\,\mathrm{d}C^{\maltese}(u,v)\right\}\\ &\quad = \int_{\mathcal{O}}\widehat{\mathbb{C}}_n^{\maltese}(u,v)\,\mathrm{d}\widehat{C}_n^{\maltese}(u,v) + \int_{\mathcal{O}}\widehat{\mathbb{C}}_n^{\maltese}(u,v)\,\mathrm{d}C^{\maltese}(u,v).\end{aligned}$$

The following representation for the limit of $\sqrt{n}(\tau_n - \tau)$ can be deduced from this relation. Details are provided in Appendix C.

Proposition 5.1. *In dimension $d = 2$, $\sqrt{n}(\tau_n - \tau)$ converges weakly, as $n \to \infty$, to the centred Gaussian random variable*

$$\mathcal{T}_2 = 8 \int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}(u, v)\, \mathrm{d}C^{\maltese}(u, v).$$

Similarly, the asymptotic normality of $\sqrt{n}(\rho_n - \rho)$ can be deduced from the theory of U-statistics; see, for example, Quessy [15]. The latter paper also considers several d-variate extensions of ρ_n which mimic the multivariate versions of these coefficients for continuous data proposed by Schmid and Schmidt [17]. Recently, we proposed alternative estimators of ρ in the multivariate case and showed that they lead to powerful tests of independence and a graphical tool for visualizing dependence in discrete data (Genest, Nešlehová and Rémillard [9]). In particular, we considered

$$\rho_{nd} = \varrho_d \left[-\frac{1}{2^d} + \frac{1}{n} \sum_{i=1}^{n} \left\{ \prod_{j=1}^{d} \left(\frac{2n+1}{2n} - \frac{R_{ij}}{n} \right) \right\} \right],$$

where $\varrho_d = 2^d(d+1)/\{2^d - (d+1)\}$ and for each $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$, R_{ij} denotes the mid-rank of X_{ij} among $X_{1j}, \ldots, X_{nj}$. The latter reduces to ρ_n in the bivariate case and can be rewritten as

$$\rho_{nd} = \varrho_d \int_{[0,1]^d} \left\{ \widehat{C}_n^{\maltese}(u_1, \ldots, u_d) - \Pi(u_1, \ldots, u_d) \right\} \mathrm{d}\Pi(u_1, \ldots, u_d).$$

Furthermore, it is a consistent estimator of

$$\rho_d = \varrho_d \int_{[0,1]^d} \left\{ C^{\maltese}(u_1, \ldots, u_d) - \Pi(u_1, \ldots, u_d) \right\} \mathrm{d}\Pi(u_1, \ldots, u_d).$$

The asymptotic normality of $\sqrt{n}(\rho_{nd} - \rho_d)$, established by Genest, Nešlehová and Rémillard [9], can be shown alternatively using Theorem 3.1. A detailed proof of the following result is given in Appendix C.

Proposition 5.2. *In arbitrary dimension $d \geq 2$, $\sqrt{n}(\rho_{nd} - \rho_d)$ converges weakly, as $n \to \infty$, to the centred Gaussian random variable*

$$\mathcal{R}_d = \varrho_d \int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}(u_1, \ldots, u_d)\, \mathrm{d}\Pi(u_1, \ldots, u_d).$$

5.2. Tests of independence

When dealing with contingency tables that are sparse or whose dimension varies with the sample size, Theorem 3.1 can be used to construct consistent and powerful tests of independence. This is because random variables $X_1, \dots, X_d$ are mutually independent if and only if $C^{\maltese} = \Pi$. To test the null hypothesis $\mathcal{H}_0$ of mutual independence between $X_1, \dots, X_d$, one could consider, for example, the Cramér–von Mises statistic

$$S_n = n \int_{[0,1]^d} \{\widehat{C}_n^{\maltese}(u_1, \dots, u_d) - \Pi(u_1, \dots, u_d)\}^2 \, d\Pi(u_1, \dots, u_d).$$

Note that when $X_1, \dots, X_d$ are continuous, S_n is equivalent to the statistic suggested by Deheuvels [4] and later studied by Genest and Rémillard [10]. The limiting distribution of S_n under $\mathcal{H}_0$ is easily deduced from Corollary 3.1 when the variables are integer-valued or increasing transformations thereof. In fact, a straightforward adaptation of the proof of Proposition 5.2 yields the following result.

Proposition 5.3. *Under $\mathcal{H}_0$ one has, as $n \to \infty$, $S_n \rightsquigarrow S$, where*

$$S = \int_{[0,1]^d} \{\widehat{\mathbb{C}}^{\maltese}(u_1, \dots, u_d)\}^2 \, d\Pi(u_1, \dots, u_d).$$

If $\mathcal{H}_0$ does not hold, then, as $n \to \infty$,

$$\frac{S_n}{n} \xrightarrow{\text{p}} \int_{[0,1]^d} \{C^{\maltese}(u_1, \dots, u_d) - \Pi(u_1, \dots, u_d)\}^2 \, d\Pi(u_1, \dots, u_d) > 0.$$

In particular, Proposition 5.3 implies that a test based on S_n is consistent against any alternative, that is, when $\mathcal{H}_0$ fails then, as $n \to \infty$, $\Pr(S_n > \varepsilon) \to 1$ for all $\varepsilon > 0$.

Unfortunately, the limiting null distribution of S_n depends on the margins of H which are generally unknown. To carry out the test, one must thus resort to resampling techniques, such as the multiplier bootstrap (van der Vaart and Wellner [19]). An illustration of how this can be done is presented below in the case $d = 2$.

Algorithm 5.1. *Given a random sample $\mathcal{X} = \{(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})\}$ from a bivariate distribution function H, define, for $i \in \{1, \dots, n\}$ and $j \in \{1, 2\}$,*

$$V_{nj,i}(u) = \lambda_{F_{nj}}(u)\mathbf{1}\{X_{ij} \le A_{nj}(k_j)\} + \{1 - \lambda_{F_{nj}}(u)\}\mathbf{1}\{X_{ij} \le A_{nj}(k_j - 1)\},$$

whenever $F_{nj}\{A_{nj}(k_j - 1)\} < u \le F_{nj}\{A_{nj}(k_j)\}$ for some $k_j \in \{0, \dots, n_j\}$. The test based on S_n can now be carried out as follows.

Step 1: *For each $m \in \{1, \dots, M\}$, generate an independent random sample $\xi_1^{(m)}, \dots, \xi_n^{(m)}$ of size n from a univariate distribution with mean zero and variance* 1, *and set $\bar{\xi}^{(m)} = (\xi_1^{(m)} + \dots + \xi_n^{(m)})/n$.*

Step 2: *For each $m \in \{1, \dots, M\}$, define the process $\mathfrak{C}_n^{(m)}$ at each $u, v \in [0, 1]$ by*

$$\mathfrak{C}_n^{(m)}(u, v) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\xi_i^{(m)} - \bar{\xi}^{(m)}) \{V_{n1,i}(u) - u\}\{V_{n2,i}(v) - v\}$$

and compute

$$S_n^{(m)} = \int_0^1 \int_0^1 \{\mathfrak{C}_n^{(m)}(u, v)\}^2 \, \mathrm{d}v \, \mathrm{d}u.$$

Step 3: *Estimate the* p*-value for the test by*

$$\frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\big(S_n^{(m)} > S_n\big).$$

An efficient implementation of this procedure is described in a companion paper in preparation, in which the validity of the multiplier bootstrap is established in this specific context. Here, the finite-sample properties of this test are merely illustrated through a small simulation study involving:

- five copulas: independence, Clayton (Cl) and Gaussian (Ga) with $\tau \in \{0.1, 0.2\}$;
- four margins: Binomial(3, 0.5), Poisson(1), Poisson(20), Geometric(0.5), respectively, denoted by F_1, F_2, F_3 and F_4;
- three statistics: S_n, the standard χ^2, and a modified version available in R in which the p-value is computed by a Monte Carlo method;
- sample size $n = 100$ and nominal level $\alpha = 5\%$;
- $M = 1000$ multiplier replicates and $N = 1000$ repetitions of the simulation.

The results of the study are displayed in Table 1 below. The test based on S_n maintains its nominal level very well in every scenario. In contrast, the standard χ^2 statistic performs rather poorly except when one of the margins is F_1. Resorting to the Monte Carlo χ^2 statistic improves the level, but the test is still slightly liberal in some cases.

The power of the test based on S_n is way better than that of its two competitors in columns 1–7 and 9. In columns 8 and 10, χ^2 is slightly better when $\tau = 0.1$. Note however that in these cases, the level of the χ^2 statistic is completely off. For a more thorough simulation study, see Murphy [12].

6. Conclusion

This paper considered the empirical multilinear copula process $\widehat{\mathbb{C}}_n^{\maltese}$ based on count data. Its convergence was established in $\mathcal{C}(K)$ for any compact $K \subset \mathcal{O}$, where $\mathcal{O}$ is an open subset of $[0, 1]^d$ avoiding the points at which the first order partial derivatives of $C^{\maltese}$ do not exist. The convergence of $\widehat{\mathbb{C}}_n^{\maltese}$ in $\mathcal{C}(K)$ is sufficient to deduce the asymptotic behavior of simple functionals thereof that are commonly used in statistical inference. This was demonstrated in Section 5 using two standard measures of association based on mid-ranks. While these specific results could have been obtained using the theory of U-statistics, knowledge of the limiting behavior of $\widehat{\mathbb{C}}_n^{\maltese}$ will be essential in other situations. The new consistent test of independence studied in Section 5 provides an example.

It is natural to ask whether the present findings can be extended to the empirical multilinear copula process based on arbitrary discontinuous data. Such an extension may well be possible,

Table 1. Percentage of rejection of the null hypothesis $\mathcal{H}_0$ of mutual independence for the three tests considered in the simulation study under various conditions

			Distribution of X_1									
			F_1	F_1	F_2	F_1	F_2	F_3	F_1	F_2	F_3	F_4
			Distribution of X_2									
			F_1	F_2	F_2	F_3	F_3	F_3	F_4	F_4	F_4	F_4
τ	C	Test										
0	Π	S_n	4.6	5.1	4.9	5.2	5.0	4.8	4.5	4.8	4.9	5.0
		χ^2	4.5	5.2	9.6	4.6	13.9	14.6	5.8	11.6	17.0	14.1
		χ^2-MC	4.6	5.4	7.1	5.3	7.1	4.4	5.5	6.1	6.1	5.7
0.1	Cl	S_n	26.6	25.4	22.3	29.2	27.0	29.3	22.0	19.1	22.5	18.2
		χ^2	17.5	9.4	14.6	8.3	10.2	23.7	6.2	16.2	11.7	22.3
		χ^2-MC	17.6	9.8	9.1	9.9	4.3	9.1	6.5	7.2	3.6	6.3
	Ga	S_n	27.1	26.2	25.7	28.7	27.7	29.3	25.5	25.0	26.2	23.7
		χ^2	11.5	9.9	21.9	6.0	15.8	17.0	8.7	27.4	20.0	34.2
		χ^2-MC	11.9	10.6	15.7	6.5	7.7	7.2	8.4	14.5	7.8	13.1
0.2	Cl	S_n	72.2	69.2	68.8	76.6	75.4	81.0	62.1	59.9	65.1	52.9
		χ^2	56.3	34.1	32.7	27.6	16.0	42.6	18.3	32.0	16.8	36.5
		χ^2-MC	56.2	33.4	22.8	30.1	7.3	22.7	18.4	16.7	4.5	13.0
	Ga	S_n	73.5	74.4	74.7	78.2	78.0	82.0	72.4	72.5	75.4	68.2
		χ^2	41.9	33.8	52.1	14.8	29.8	31.3	25.3	56.7	33.7	66.4
		χ^2-MC	43.2	33.4	39.5	16.9	15.2	13.3	26.7	34.2	14.1	36.7

given that the estimator $\widehat{C}_n^{\maltese}$ is defined in general. We are currently investigating this issue. Once this task has been completed, the process $\widehat{\mathbb{C}}_n^{\maltese}$ will provide a solid foundation for inference in copula models with arbitrary margins.

Appendix A: Proofs from Section 2

Proof of Proposition 2.2. For all $x_1, \dots, x_d \in \mathbb{R}$, one has

$$H^{\maltese}(x_1, \dots, x_d) = \int_{[0,1]^d} H(x_1 + u_1, \dots, x_d + u_d)\,\mathrm{d}u_1 \cdots \mathrm{d}u_d.$$

If $(x_1, \dots, x_d) \in [k_1 - 1, k_1) \times \cdots \times [k_d - 1, k_d)$ for some $k_1, \dots, k_d \in \mathbb{N}$, one can replace each $x_j + u_j$ by $k_j - 1$ or by k_j, according as $0 < u_j < k_j - x_j$ or $k_j - x_j \leq u_j < 1$ because H is supported on $\mathbb{N}^d$. After straightforward simplification, it follows that

$$H^{\maltese}(x_1, \ldots, x_d) = \sum_{S \subseteq \{1,\ldots,d\}} H(k_S) \biggl\{ \prod_{\ell \notin S} (k_\ell - x_\ell) \biggr\} \biggl\{ \prod_{\ell \in S} (x_\ell - k_\ell + 1) \biggr\},$$

where $k_S = (k_{S_1}, \ldots, k_{S_d})$ and $k_{S_j} = k_j$ if $j \in S$ and $k_{S_j} = k_j - 1$ otherwise.

If $F^{\maltese}$ is a generic margin of $H^{\maltese}$, then $F^{\maltese}$ is a linear interpolation of F, that is, $F^{\maltese}(x) = 0$ for $x < -1$ while

$$F^{\maltese}(x) = F(k-1) + \Delta F(k)(x - k + 1) \tag{A.1}$$

when $x \in [k-1, k)$ for some $k \in \mathbb{N}$. Thus when $u \in (F(k-1), F(k)]$, one has

$$F^{\maltese -1}(u) = k - 1 + \frac{u - F(k-1)}{\Delta F(k)}. \tag{A.2}$$

If $u = 0$, one can set $F^{\maltese -1}(0) = -1$ for convenience, because the support of X is bounded below by 0 by hypothesis. It is then immediate that $H^{\maltese}\{F_1^{\maltese -1}(u_1), \ldots, F_d^{\maltese -1}(u_d)\}$ yields the formula for $C^{\maltese}$ given in Proposition 2.1. □

Appendix B: Proofs from Section 4

The following elementary result is used in the sequel.

Lemma B.1. *If G is a cumulative distribution function, then for all $u \in (0,1)$ and $x \in \mathbb{R}$, one has $u \le G(x) \Leftrightarrow G^{-1}(u) \le x \Leftrightarrow G \circ G^{-1}(u) \le G(x)$.*

Proof of Proposition 4.1. First note that for fixed values of $k_1, \ldots, k_d \in \mathbb{N}$, one has

$$\begin{aligned} D_n\bigl\{F_1(k_1), \ldots, F_d(k_d)\bigr\} &= \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\bigl\{F_j(X_{ij}) \le F_j(k_j)\bigr\} \\ &= \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\bigl\{F_j \circ F_j^{-1}(V_{ij}) \le F_j(k_j)\bigr\}. \end{aligned}$$

In view of Lemma B.1, it follows that

$$D_n\bigl\{F_1(k_1), \ldots, F_d(k_d)\bigr\} = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\bigl\{V_{ij} \le F_j(k_j)\bigr\} = B_n\bigl\{F_1(k_1), \ldots, F_d(k_d)\bigr\}.$$

From the definition of $\mathfrak{M}_H$, one then has $\mathfrak{M}_H(D_n) = \mathfrak{M}_H(B_n)$ and hence $C_n^{\maltese} = \mathfrak{M}_H(B_n)$. The linearity of $\mathfrak{M}_H$ and the fact that $\mathfrak{M}_H(C^{\maltese}) = C^{\maltese}$ further imply that $\mathbb{C}_n^{\maltese} = \mathfrak{M}_H(\mathbb{B}_n)$ from which it also follows that

$$\bigl\|\mathbb{C}_n^{\maltese}\bigr\| = \bigl\|\mathfrak{M}_H(\mathbb{B}_n)\bigr\| \le \|\mathbb{B}_n\| \tag{B.1}$$

because the operator $\mathfrak{M}_H$ is a contraction. Given that $\mathfrak{M}_H$ is a continuous mapping and that $\mathbb{B}_n \rightsquigarrow \mathbb{B}_{C^\maltese}$ as $n \to \infty$, the Continuous Mapping theorem yields the conclusion. □

The following auxiliary results are needed for the proof of Proposition 4.2.

Lemma B.2. *For all* $u_1, \ldots, u_d \in [0, 1]$,

$$C_n^\maltese(u_1, \ldots, u_d) = H_n^\maltese\bigl\{F_1^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1}(u_d)\bigr\}.$$

Proof. First note that the functions on both sides of the above identity are continuous on $[0, 1]^d$. This is the case for $C_n^\maltese$, as explained in Section 4. To see why this is true for the other one, fix arbitrary $u_1, \ldots, u_d \in [0, 1)$ and observe that

$$\begin{aligned}&\bigl|H_n^\maltese\bigl\{F_1^{\maltese-1}(u_1+), \ldots, F_d^{\maltese-1}(u_d+)\bigr\} - H_n^\maltese\bigl\{F_1^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1}(u_d)\bigr\}\bigr| \\ &\quad \le \sum_{j=1}^d \bigl|F_{nj}^\maltese \circ F_j^{\maltese-1}(u_j+) - F_{nj}^\maltese \circ F_j^{\maltese-1}(u_j)\bigr|.\end{aligned}$$

Now each of the summands on the right-hand side must vanish. For, even if u_j is a point of discontinuity of $F_j^{\maltese-1}$ for some $j \in \{1, \ldots, d\}$, the fact that $F_j^\maltese$ is continuous implies that $F_j^\maltese \circ F_j^{\maltese-1}(u_j) = F_j^\maltese \circ F_j^{\maltese-1}(u_j+) = u_j$. Now for arbitrary $x, y \in \mathbb{R}$, one has

$$F_j^\maltese(x) = F_j^\maltese(y) \quad \Rightarrow \quad F_{nj}^\maltese(x) = F_{nj}^\maltese(y), \tag{B.2}$$

because F_{nj} can only jump where F_j does. Hence $F_{nj}^\maltese \circ F_j^{\maltese-1}(u_j) = F_{nj}^\maltese \circ F_j^{\maltese-1}(u_j+)$.

Therefore, it suffices to look at the case where $u_1, \ldots, u_d \in (0, 1)$. Suppose that for each $j \in \{1, \ldots, d\}$, $u_j \in (F_j(k_j - 1), F_j(k_j)]$ for some $k_j \in \mathbb{N}$. It then follows from (A.2) that, for all $j \in \{1, \ldots, d\}$,

$$F_j^{\maltese-1}(u_j) = k_j - 1 + \frac{u_j - F_j(k_j - 1)}{\Delta F_j(k_j)},$$

and hence

$$\begin{aligned}k_j - F_j^{\maltese-1}(u_j) &= \frac{F_j(k_j) - u_j}{\Delta F_j(k_j)}, \\ F_j^{\maltese-1}(u_j) - k_j + 1 &= \frac{u_j - F_j(k_j - 1)}{\Delta F_j(k_j)}.\end{aligned}$$

Consequently,

$$H_n^\maltese\bigl\{F_1^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1}(u_d)\bigr\} = \sum_{S \subset \{1,\ldots,d\}} \lambda_{H,S}(u_1, \ldots, u_d) H_n(k_{S_1}, \ldots, k_{S_d}).$$

Now in view of Lemma B.1, one has

$$H_n(k_1,\dots,k_d) = \frac{1}{n}\sum_{i=1}^n \prod_{j=1}^d \mathbf{1}(X_{ij} \le k_j) = \frac{1}{n}\sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\bigl\{F_j^{-1}(V_{ij}) \le k_j\bigr\}$$
$$= \frac{1}{n}\sum_{i=1}^n \prod_{j=1}^d \mathbf{1}\bigl\{V_{ij} \le F_j(k_j)\bigr\} = B_n\bigl\{F_1(k_1),\dots,F_d(k_d)\bigr\}.$$

Therefore,

$$H_n^{\maltese}\bigl\{F_1^{\maltese-1}(u_1),\dots,F_d^{\maltese-1}(u_d)\bigr\} = \sum_{S\subset\{1,\dots,d\}} \lambda_{H,S}(u_1,\dots,u_d) B_n\bigl\{F_1(k_{S_1}),\dots,F_d(k_{S_d})\bigr\},$$

which is $\mathfrak{M}_H(B_n)(u_1,\dots,u_d)$. From the proof of Proposition 4.1, $\mathfrak{M}_H(B_n) = C_n^{\maltese}$. □

Lemma B.3. *For arbitrary $n \in \mathbb{N}$, $G_n = F_n^{\maltese} \circ F^{\maltese-1}$ is a continuous distribution function on $[0,1]$ and $G_n^{-1} = F^{\maltese} \circ F_n^{\maltese-1}$.*

Proof. As G_n is the convolution of two non-decreasing functions, it is non-decreasing. Furthermore, $G_n(0) = 0$ and $G_n(1) = 1$ by construction. Proceeding as in the proof of Lemma B.2, one can show that G_n is indeed continuous. Turning to G_n^{-1}, fix $u \in [0,1]$ and observe that for any $x \in \mathbb{R}$ such that $F_n^{\maltese}(x) \ge u$, one has $F^{\maltese} \circ F_n^{\maltese-1}(u) \le F^{\maltese}(x)$ because $F^{\maltese}$ is non-decreasing. Now suppose that $y \in \mathbb{R}$ is such that for all $x \in \mathbb{R}$, $F_n^{\maltese}(x) \ge u \Rightarrow y \le F^{\maltese}(x)$. By virtue of Lemma B.1, this is equivalent to saying that for all $x \in \mathbb{R}$, $F_n^{\maltese}(x) \ge u \Rightarrow F^{\maltese-1}(y) \le x$. This implies that $F^{\maltese-1}(y) \le F_n^{\maltese-1}(u)$. Applying Lemma B.1 once again, one can see that $y \le F^{\maltese} \circ F_n^{\maltese-1}(u)$. Consequently,

$$F^{\maltese} \circ F_n^{\maltese-1}(u) = \inf\bigl\{F^{\maltese}(x)\colon\ F_n^{\maltese}(x) \ge u\bigr\}.$$

Next, $F^{\maltese} \circ F^{\maltese-1}(u) = u$ by continuity of $F^{\maltese}$. Hence, for all $x \in \mathbb{R}$, $F^{\maltese} \circ F^{\maltese-1} \circ F^{\maltese}(x) = F^{\maltese}(x)$. Invoking implication (B.2), one deduces that $F_n^{\maltese} \circ F^{\maltese-1} \circ F^{\maltese}(x) = F_n^{\maltese}(x)$, which implies

$$\inf\bigl\{F^{\maltese}(x)\colon\ F_n^{\maltese}(x) \ge u\bigr\} = \inf\bigl\{F^{\maltese}(x)\colon\ F_n^{\maltese} \circ F^{\maltese-1} \circ F^{\maltese}(x) \ge u\bigr\}$$
$$= \inf\bigl\{v\colon\ F_n^{\maltese} \circ F^{\maltese-1}(v) \ge u\bigr\} = \inf\bigl\{v\colon\ G_n(v) \ge u\bigr\}.$$

In other words, $F^{\maltese} \circ F_n^{\maltese-1} = G_n^{-1}$. □

Proof of Proposition 4.2. First note that in view of Lemma B.2 and Proposition 2.2, one has, for all $u_1,\dots,u_d \in [0,1]$,

$$\mathbb{C}_n^{\maltese}(u_1,\dots,u_d) = \sqrt{n}\bigl[H_n^{\maltese}\bigl\{F_1^{\maltese-1}(u_1),\dots,F_d^{\maltese-1}(u_d)\bigr\} - H^{\maltese}\bigl\{F_1^{\maltese-1}(u_1),\dots,F_d^{\maltese-1}(u_d)\bigr\}\bigr].$$

Next observe that, for all $u_1, \ldots, u_d \in [0, 1]$,

$$\widetilde{\mathbb{C}}_n^{\maltese}(u_1, \ldots, u_d) = \mathbb{C}_n^{\maltese}\{F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\}. \tag{B.3}$$

Indeed, one can write

$$\begin{aligned}
&\mathbb{C}_n^{\maltese}\{F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\} \\
&\quad = \sqrt{n}\big[H_n^{\maltese}\{F_1^{\maltese-1} \circ F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1} \circ F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\} \\
&\qquad - H^{\maltese}\{F_1^{\maltese-1} \circ F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1} \circ F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\}\big].
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\big|H_n^{\maltese}\{F_1^{\maltese-1} \circ F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1} \circ F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\} \\
&\qquad - H_n^{\maltese}\{F_{n1}^{\maltese-1}(u_1), \ldots, F_{nd}^{\maltese-1}(u_d)\}\big| \\
&\quad \le \sum_{j=1}^{d} \big|F_{nj}^{\maltese} \circ F_j^{\maltese-1} \circ F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j) - F_{nj}^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\big|.
\end{aligned}$$

Now the right-hand side is zero by Lemma B.3 and the fact that for all $j \in \{1, \ldots, d\}$ and $u_j \in [0, 1]$, $F_{nj}^{\maltese} \circ F_{nj}^{\maltese-1}(u_j) = u_j$ because $F_{nj}^{\maltese}$ is a continuous distribution function. As $F_1^{\maltese}, \ldots, F_d^{\maltese}$ are also continuous distribution functions, one has

$$\begin{aligned}
&\big|H^{\maltese}\{F_1^{\maltese-1} \circ F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1), \ldots, F_d^{\maltese-1} \circ F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\} \\
&\qquad - H^{\maltese}\{F_{n1}^{\maltese-1}(u_1), \ldots, F_{nd}^{\maltese-1}(u_d)\}\big| \\
&\quad \le \sum_{j=1}^{d} \big|F_j^{\maltese} \circ F_j^{\maltese-1} \circ F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j) - F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\big| = 0.
\end{aligned}$$

Therefore, identity (B.3) holds and one can write

$$\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\| = \|\mathbb{C}_n^{\maltese} - \mathbb{C}_n^{\maltese}\{F_1^{\maltese} \circ F_{n1}^{\maltese-1}, \ldots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}\}\|.$$

Next, using (A.1) and (A.2) applied to F and F_n, respectively, a direct calculation yields

$$\begin{aligned}
&\sqrt{n}\{u_j - F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\} \\
&\quad = \mathbb{B}_{nj}\{F_j(k_j - 1)\}\left\{\frac{F_{nj}(k_j) - u_j}{\Delta F_{nj}(k_j)}\right\} + \mathbb{B}_{nj}\{F_j(k_j)\}\left\{\frac{u_j - F_{nj}(k_j - 1)}{\Delta F_{nj}(k_j)}\right\},
\end{aligned}$$

whenever $u_j \in (F_{nj}(k_j - 1), F_{nj}(k_j)]$ for some $k_j \in \mathbb{N}$. It follows that

$$\sup_{u_j \in [0,1]} \big|F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j) - u_j\big| \le \frac{1}{\sqrt{n}} \|\mathbb{B}_{nj}\| \le \frac{1}{\sqrt{n}} \|\mathbb{B}_n\|. \tag{B.4}$$

As $n \to \infty$, $\|\mathbb{B}_n\| \rightsquigarrow \|\mathbb{B}_{C^{\maltese}}\|$ and hence $\|\mathbb{B}_n\|/\sqrt{n} \xrightarrow{\mathrm{p}} 0$. Now for arbitrary $\varepsilon > 0$, one has

$$P^*\bigl(\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\| > \varepsilon\bigr) = P^*\bigl\{\|\mathbb{C}_n^{\maltese} - \mathbb{C}_n^{\maltese}\bigl(F_1^{\maltese} \circ F_{n1}^{\maltese-1}, \dots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}\bigr)\| > \varepsilon\bigr\},$$

where P^* denotes outer probability. Given $\delta > 0$, the right-hand side is the same as

$$P^*\left\{\|\mathbb{C}_n^{\maltese} - \mathbb{C}_n^{\maltese}\bigl(F_1^{\maltese} \circ F_{n1}^{\maltese-1}, \dots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}\bigr)\| > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} < \delta\right\}$$
$$+ P^*\left\{\|\mathbb{C}_n^{\maltese} - \mathbb{C}_n^{\maltese}\bigl(F_1^{\maltese} \circ F_{n1}^{\maltese-1}, \dots, F_d^{\maltese} \circ F_{nd}^{\maltese-1}\bigr)\| > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} \ge \delta\right\},$$

and in view of (B.4), the latter is bounded above by

$$P^*\bigl\{\omega_n\bigl(\mathbb{C}_n^{\maltese}, \delta\bigr) > \varepsilon\bigr\} + P^*\left(\frac{\|\mathbb{B}_n\|}{\sqrt{n}} \ge \delta\right),$$

where

$$\omega_n\bigl(\mathbb{C}_n^{\maltese}, \delta\bigr) = \sup_{\substack{u_j, v_j \in [0,1]:\ |u_j - v_j| < \delta, \\ j \in \{1,\dots,d\}}} \bigl|\mathbb{C}_n^{\maltese}(u_1, \dots, u_d) - \mathbb{C}_n^{\maltese}(v_1, \dots, v_d)\bigr|.$$

Therefore,

$$\limsup_{n\to\infty} P^*\bigl(\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\| > \varepsilon\bigr) \le \limsup_{n\to\infty} P^*\bigl\{\omega_n\bigl(\mathbb{C}_n^{\maltese}, \delta\bigr) > \varepsilon\bigr\}.$$

Finally, recall that $\mathbb{C}_n^{\maltese}$ converges weakly in $\mathcal{C}([0,1]^d)$ to a measurable random element $\mathfrak{M}_H(\mathbb{B}_{C^{\maltese}})$. Because $\mathcal{C}([0,1]^d)$ is complete and separable, Theorem 11.5.4. in Dudley [5] implies that $\mathfrak{M}_H(\mathbb{B}_{C^{\maltese}})$ is tight. It then follows from Lemma 1.3.8. and Theorem 1.5.7. in van der Vaart and Wellner [19] that the sequence $\mathbb{C}_n^{\maltese}$ is asymptotically tight and hence asymptotically uniformly equicontinuous in probability, viz.

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} P^*\bigl\{\omega_n\bigl(\mathbb{C}_n^{\maltese}, \delta\bigr) > \varepsilon\bigr\} = 0.$$

This means that as $n \to \infty$, $P^*(\|\mathbb{C}_n^{\maltese} - \widetilde{\mathbb{C}}_n^{\maltese}\| > \varepsilon) \to 0$ for all $\epsilon > 0$. □

Proof of Proposition 4.3. For fixed $j \in \{1, \dots, d\}$ and $u_j \in [0,1]$, first write $F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)$ in the form $u_j - \sqrt{n}\{u_j - F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\}/\sqrt{n}$. Then

$$\|\mathbb{D}_n - \widetilde{\mathbb{D}}_n\| = \sqrt{n} \sup_{u_1,\dots,u_d \in [0,1]} \left| C^{\maltese}\left\{u_1 - \frac{\mathbb{C}_{n1}^{\maltese}(u_1)}{\sqrt{n}}, \dots, u_d - \frac{\mathbb{C}_{nd}^{\maltese}(u_d)}{\sqrt{n}}\right\} - C^{\maltese}\left[u_1 - \frac{\sqrt{n}\{u_1 - F_1^{\maltese} \circ F_{n1}^{\maltese-1}(u_1)\}}{\sqrt{n}}, \dots, u_d - \frac{\sqrt{n}\{u_d - F_d^{\maltese} \circ F_{nd}^{\maltese-1}(u_d)\}}{\sqrt{n}}\right]\right|.$$

The Lipschitz property of copulas further implies that

$$\|\mathbb{D}_n - \widetilde{\mathbb{D}}_n\| \le \sum_{j=1}^{d} \sup_{u_j \in [0,1]} \bigl|\mathbb{C}_{nj}^{\maltese}(u_j) - \sqrt{n}\bigl\{u_j - F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\bigr\}\bigr|.$$

For any $j \in \{1, \dots, d\}$, one can now call upon Proposition 4.2 with $d = 1$ and $H = F_j$ to conclude that, as $n \to \infty$, $\sup_{u_j \in [0,1]} |\mathbb{C}_{nj}^{\maltese}(u_j) - \sqrt{n}\{u_j - F_j^{\maltese} \circ F_{nj}^{\maltese-1}(u_j)\}| \xrightarrow{\mathrm{p}} 0$. □

The proof of Proposition 4.4 relies on the following lemma.

Lemma B.4. *Let $u_1, \dots, u_d \in [0, 1]$ and $v_1, \dots, v_d \in [0, 1]$ be such that for each $j \in \{1, \dots, d\}$, $u_j, v_j \in (F_j(k_j - 1), F_j(k_j))$ for some $k_j \in \mathbb{N}$. Then*

$$C^{\maltese}(v_1, \dots, v_d) - C^{\maltese}(u_1, \dots, u_d) = \sum_{m=1}^{d} (v_m - u_m)\dot{C}_m^{\maltese}(w_{m1}, \dots, w_{md}),$$

where w_{mj} equals u_j or v_j according as $j < m$ or $j \ge m$, respectively.

Proof. First, write $C^{\maltese}(v_1, \dots, v_d) - C^{\maltese}(u_1, \dots, u_d)$ in the alternative form

$$\sum_{m=1}^{d} \bigl\{C^{\maltese}(w_{m1}, \dots, w_{md}) - C^{\maltese}(w_{(m+1)1}, \dots, w_{(m+1)d})\bigr\}.$$

It must then be shown that for all $m \in \{1, \dots, d\}$, one has

$$\begin{aligned} &C^{\maltese}(w_{m1}, \dots, w_{md}) - C^{\maltese}(w_{(m+1)1}, \dots, w_{(m+1)d}) \\ &\quad = (v_m - u_m)\dot{C}_m^{\maltese}(w_{m1}, \dots, w_{md}). \end{aligned} \tag{B.5}$$

To this end, observe that on the left-hand side of (B.5), $C^{\maltese}$ is evaluated at two vectors whose components are identical, except in position m. Let $w_1, \dots, w_{m-1}, w_{m+1}, \dots, w_d$ be the matching components and note that $w_{mm} = v_m$ while $w_{(m+1)m} = u_m$. Given $S \subset \{1, \dots, d\}$, let s_m be the size of $S \cap \{m\}$. From the definition of $\lambda_{H,S}$, one has

$$\begin{aligned} \lambda_{H,S}(w_{m1}, \dots, w_{md}) &= \lambda_{H,S}(w_1, \dots, w_{m-1}, v_m, w_{m+1}, \dots, w_d) \\ &= \left\{\frac{v_m - F_m(k_m - 1)}{\Delta F_m(k_m)}\right\}^{s_m} \times \left\{\frac{F_m(k_m) - v_m}{\Delta F_m(k_m)}\right\}^{1-s_m} \\ &\quad \times \left\{\prod_{\substack{\ell \notin S \\ \ell \ne m}} \frac{F_\ell(k_\ell) - w_\ell}{\Delta F_\ell(k_\ell)}\right\} \times \left\{\prod_{\substack{\ell \in S \\ \ell \ne m}} \frac{w_\ell - F_\ell(k_\ell - 1)}{\Delta F_\ell(k_\ell)}\right\} \end{aligned}$$

and

$$\lambda_{H,S}(w_{(m+1)1},\ldots,w_{(m+1)d}) = \lambda_{H,S}(w_1,\ldots,w_{m-1},u_m,w_{m+1},\ldots,w_d)$$
$$= \left\{\frac{u_m - F_m(k_m-1)}{\Delta F_m(k_m)}\right\}^{s_m} \times \left\{\frac{F_m(k_m)-u_m}{\Delta F_m(k_m)}\right\}^{1-s_m}$$
$$\times \left\{\prod_{\substack{\ell\notin S\\ \ell\neq m}} \frac{F_\ell(k_\ell)-w_\ell}{\Delta F_\ell(k_\ell)}\right\} \times \left\{\prod_{\substack{\ell\in S\\ \ell\neq m}} \frac{w_\ell - F_\ell(k_\ell-1)}{\Delta F_\ell(k_\ell)}\right\}.$$

Consequently, their difference is equal to

$$(v_m-u_m)\frac{(-1)^{1-s_m}}{\Delta F_m(k_m)} \times \left\{\prod_{\substack{\ell\notin S\\ \ell\neq m}} \frac{F_\ell(k_\ell)-w_\ell}{\Delta F_\ell(k_\ell)}\right\} \times \left\{\prod_{\substack{\ell\in S\\ \ell\neq m}} \frac{w_\ell - F_\ell(k_\ell-1)}{\Delta F_\ell(k_\ell)}\right\}.$$

It then follows from the definition of $C^{\maltese}$ that

$$C^{\maltese}(w_{m1},\ldots,w_{md}) - C^{\maltese}(w_{(m+1)1},\ldots,w_{(m+1)d})$$
$$= \sum_{S\subset\{1,\ldots,d\}} H(k_S)\bigl\{\lambda_{H,S}(w_{m1},\ldots,w_{md}) - \lambda_{H,S}(w_{(m+1)1},\ldots,w_{(m+1)d})\bigr\}$$
$$= (v_m-u_m)\sum_{S\subset\{1,\ldots,d\}} H(k_S)\frac{(-1)^{1-s_m}}{\Delta F_m(k_m)}\left\{\prod_{\substack{\ell\notin S\\ \ell\neq m}} \frac{F_\ell(k_\ell)-w_\ell}{\Delta F_\ell(k_\ell)}\right\} \times \left\{\prod_{\substack{\ell\in S\\ \ell\neq m}} \frac{w_\ell - F_\ell(k_\ell-1)}{\Delta F_\ell(k_\ell)}\right\}$$
$$= (v_m-u_m)\dot{C}^{\maltese}_m(w_{m1},\ldots,w_{md}).$$

This completes the argument. □

Proof of Proposition 4.4. Recall from the definition of $\mathcal{O}$ that because K is compact, it can be covered by finitely many open cubes of the form

$$\mathcal{O}_\ell = \bigl(F_1(k_{1\ell}-1), F_1(k_{\ell 1})\bigr) \times\cdots\times \bigl(F_d(k_{d\ell}-1), F_d(k_{d\ell})\bigr),$$

where $k_{1\ell},\ldots,k_{d\ell}\in\mathbb{N}$ for $\ell\in\{1,\ldots,L\}$. Given that the sets $\mathcal{O}_1,\ldots,\mathcal{O}_L$ are mutually disjoint, $K_\ell = K\cap\mathcal{O}_\ell$ is compact for each $\ell\in\{1,\ldots,L\}$. Therefore, $K = K_1\cup\cdots\cup K_L$ is a union of finitely many disjoint compact sets. For arbitrary $\delta>0$, let

$$K_{\ell,\delta} = \bigcup_{(x_1,\ldots,x_d)\in K_\ell} \bigl\{(u_1,\ldots,u_d)\in\mathbb{R}^d\colon |u_1-x_1|+\cdots+|u_d-x_d|<\delta\bigr\}.$$

Because $K_1,\ldots,K_L$ are compact, there exists $\delta_0>0$ such that $K_{\ell,\delta_0}\subset\mathcal{O}_\ell$ for all $\ell\in\{1,\ldots,L\}$. Now fix $\delta^*<\delta_0$ and let K^* denote the closure of $K_{\delta^*} = K_{1,\delta^*}\cup\cdots\cup K_{d,\delta^*}$, which is compact.

Then for all $\delta \in (0, \delta^*)$, one has $K \subset K_\delta \subset K^* \subset \mathcal{O}$. For fixed $\delta \in (0, \delta^*)$ and $\varepsilon > 0$, write

$$\begin{aligned} &P^*\{\|\mathbb{D}_n - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})\|_K > \varepsilon\} \\ &\quad = P^*\left\{\|\mathbb{D}_n - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})\|_K > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} < \frac{\delta}{d}\right\} \\ &\qquad + P^*\left\{\|\mathbb{D}_n - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})\|_K > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} \geq \frac{\delta}{d}\right\}. \end{aligned}$$

When the event $\{\|\mathbb{B}_n\|/\sqrt{n} < \delta/d\}$ holds and $(u_1, \dots, u_d) \in K_\ell$ for some $\ell \in \{1, \dots, L\}$,

$$(v_1, \dots, v_d) = \left(u_1 - \frac{\mathbb{C}_{n1}^{\maltese}(u_1)}{\sqrt{n}}, \dots, u_d - \frac{\mathbb{C}_{nd}^{\maltese}(u_d)}{\sqrt{n}}\right) \in K_{\ell,\delta}$$

because, for all $j \in \{1, \dots, d\}$, $\|\mathbb{C}_{nj}^{\maltese}\| \leq \|\mathbb{C}_n^{\maltese}\| \leq \|\mathbb{B}_n\|$ by (B.1). From Lemma B.4,

$$\begin{aligned} &|\mathbb{D}_n(u_1, \dots, u_d) - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})(u_1, \dots, u_d)| \\ &\quad = \left|\sum_{j=1}^d \mathbb{C}_{nj}^{\maltese}(u_j)\{\dot{C}_j^{\maltese}(u_1, \dots, u_d) - \dot{C}_j^{\maltese}(u_1, \dots, u_{j-1}, v_j, \dots, v_d)\}\right| \\ &\quad \leq \|\mathbb{B}_n\| \sum_{j=1}^d |\dot{C}_j^{\maltese}(u_1, \dots, u_d) - \dot{C}_j^{\maltese}(u_1, \dots, u_{j-1}, v_j, \dots, v_d)|. \end{aligned}$$

Consequently, $\|\mathbb{D}_n - \mathfrak{D}_K(\mathbb{C}_n^{\maltese})\|_K$ is bounded above by

$$\begin{aligned} &\|\mathbb{B}_n\| \sum_{j=1}^d \sup_{(u_1,\dots,u_d)\in K} |\dot{C}_j^{\maltese}(u_1, \dots, u_d) - \dot{C}_j^{\maltese}(u_1, \dots, u_{j-1}, v_j, \dots, v_d)| \\ &\quad \leq \|\mathbb{B}_n\| \sum_{j=1}^d \sup_{\substack{(u_1,\dots,u_d)\in K, \\ (w_1,\dots,w_d)\in K_\delta, \\ \sum_{m=1}^d |u_m - w_m| < \delta}} |\dot{C}_j^{\maltese}(u_1, \dots, u_d) - \dot{C}_j^{\maltese}(w_1, \dots, w_d)| \\ &\quad \leq \|\mathbb{B}_n\| \omega(\delta), \end{aligned}$$

where

$$\omega(\delta) = \sum_{j=1}^d \sup_{\substack{(u_1,\dots,u_d),(w_1,\dots,w_d)\in K^*, \\ \sum_{m=1}^d |u_m - w_m| < \delta}} |\dot{C}_j^{\maltese}(u_1, \dots, u_d) - \dot{C}_j^{\maltese}(w_1, \dots, w_d)|.$$

This observation implies that

$$\begin{aligned}
&\limsup_{n\to\infty} P^*\left\{\left\|\mathbb{D}_n - \mathfrak{D}_K\left(\mathbb{C}_n^{\maltese}\right)\right\|_K > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} < \frac{\delta}{d}\right\} \\
&\quad\le \limsup_{n\to\infty} P^*\left\{\|\mathbb{B}_n\|\omega(\delta) > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} < \frac{\delta}{d}\right\} \\
&\quad\le \limsup_{n\to\infty} P^*\left\{\|\mathbb{B}_n\|\omega(\delta) > \varepsilon\right\} = P^*\left\{\|\mathbb{B}_{C^{\maltese}}\|\omega(\delta) > \varepsilon\right\},
\end{aligned}$$

where the equality is justified by the fact that $\|\mathbb{B}_n\| \rightsquigarrow \|\mathbb{B}_{C^{\maltese}}\|$ as $n \to \infty$. Now $\omega(\delta) \to 0$ as $\delta \to 0$ because $\dot{C}_j^{\maltese}$ is absolutely continuous on K^* for all $j \in \{1, \ldots, d\}$. Therefore, $P^*\{\|\mathbb{B}_{C^{\maltese}}\|\omega(\delta) > \varepsilon\} \to 0$ as $\delta \to 0$. Finally, observe that

$$\limsup_{n\to\infty} P^*\left(\left\|\mathbb{D}_n - \mathfrak{D}_K\left(\mathbb{C}_n^{\maltese}\right)\right\|_K > \varepsilon, \frac{\|\mathbb{B}_n\|}{\sqrt{n}} \ge \frac{\delta}{d}\right) \le \limsup_{n\to\infty} P^*\left(\frac{\|\mathbb{B}_n\|}{\sqrt{n}} \ge \frac{\delta}{d}\right) = 0$$

because $\|\mathbb{B}_n\|/\sqrt{n} \xrightarrow{\text{p}} 0$ as $n \to \infty$. As $\varepsilon > 0$ is arbitrary, one can conclude. □

The proof of Proposition 4.5 relies on the following lemma.

Lemma B.5. *Let G be the distribution function of a uniform random variable on* $(0, 1)$. *Then for every $j \in \{1, \ldots, d\}$ and as $n \to \infty$,*

$$Y_{nj} = \sqrt{n} \sup_{0\le u\le 1}\left| G\left\{u - \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right\} - u + \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right| \xrightarrow{\text{p}} 0.$$

Proof. Fix $j \in \{1, \ldots, d\}$ and write

$$\begin{aligned}
Y_{nj} = \sqrt{n} \sup_{0\le u\le 1}\Bigg[&\left\{\frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}} - u\right\}\mathbf{1}\left\{u < \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right\} \\
&+ \left\{-\frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}} - (1-u)\right\}\mathbf{1}\left\{1 - u < -\frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right\}\Bigg].
\end{aligned}$$

Observe that if $\|\mathbb{C}_{nj}^{\maltese}\| \le M$ for some constant $M > 0$, then as $n \to \infty$,

$$Y_{nj} \le \sup_{0\le u\le M/\sqrt{n}}\left|\mathbb{C}_{nj}^{\maltese}(u)\right| + \sup_{1-M/\sqrt{n}\le u\le 1}\left|\mathbb{C}_{nj}^{\maltese}(u)\right| \xrightarrow{\text{p}} 0$$

because $\mathbb{C}_{nj}^{\maltese} \rightsquigarrow \mathbb{C}_j^{\maltese}$ in $\mathcal{C}([0,1])$ and $\mathbb{C}_j^{\maltese}(0) = \mathbb{C}_j^{\maltese}(1) = 0$. Now fix $\varepsilon > 0$ and invoke the tightness of $\mathbb{C}_{nj}^{\maltese}$ to find $M > 0$ such that $\Pr(\|\mathbb{C}_{nj}^{\maltese}\| > M) < \varepsilon/2$ for all $n \in \mathbb{N}$. Thus,

$$\Pr(Y_{nj} > \varepsilon) \le \Pr\left(\left\|\mathbb{C}_{nj}^{\maltese}\right\| > M\right) + \Pr\left(\sup_{0\le u\le M/\sqrt{n}}\left|\mathbb{C}_{nj}^{\maltese}(u)\right| + \sup_{1-M/\sqrt{n}\le u\le 1}\left|\mathbb{C}_{nj}^{\maltese}(u)\right| > \varepsilon\right).$$

If n is large enough, the right-hand side of the above inequality is at most ε. □

Proof of Proposition 4.5. For all $u_1, \dots, u_d \in [0,1]$, let

$$\mathbb{D}_n^*(u_1, \dots, u_d) = \sqrt{n}\left[\prod_{j=1}^d \left\{u_j - \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right\} - \prod_{j=1}^d u_j\right].$$

Then

$$\|\mathbb{D}_n - \mathbb{D}_n^*\| \le \sum_{j=1}^d \sqrt{n} \sup_{0\le u\le 1}\left|G\left\{u - \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right\} - u + \frac{\mathbb{C}_{nj}^{\maltese}(u)}{\sqrt{n}}\right|$$

because $|\prod_{j=1}^d a_j - \prod_{j=1}^d b_j| \le \sum_{j=1}^d |a_j - b_j|$ for all $a_1, \dots, a_d, b_1, \dots, b_d \in (0,1)$. Lemma B.5 thus implies that $\|\mathbb{D}_n - \mathbb{D}_n^*\| \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$. Now by the multinomial formula,

$$\|\mathbb{D}_n^* - \mathfrak{D}(\mathbb{C}_n^{\maltese})\| \le \sqrt{n} \sum_{S\subset\{1,\dots,d\},|S|\ge 2} \prod_{j\in S} \frac{\|\mathbb{C}_{nj}^{\maltese}\|}{\sqrt{n}} \xrightarrow{\mathrm{p}} 0.$$

□

Appendix C: Proofs from Section 5

The proofs of Propositions 5.1 and 5.2 have much in common. They both rely on the following straightforward consequence of Proposition 6.3.9 in Brockwell and Davis [2].

Lemma C.1. *Let Z_n be a sequence of random variables. Suppose that for all $\delta, \epsilon > 0$, there exists a sequence $Y_{n,\delta,\epsilon}$ of random variables such that for all $n \in \mathbb{N}$, $\Pr(|Z_n - Y_{n,\delta,\epsilon}| > \delta) < \epsilon$ and $Y_{n,\delta,\epsilon} \rightsquigarrow Y_{\delta,\epsilon}$ as $n \to \infty$. Further assume that there exists a random variable Z such that for all $\delta, \epsilon > 0$, $\Pr(|Z - Y_{\delta,\epsilon}| > \delta) < \epsilon$. Then $Z_n \rightsquigarrow Z$ as $n \to \infty$.*

The convergence of Spearman's rho is presented first.

Proof of Proposition 5.2. Because the complement of $\mathcal{O}$ in $[0,1]^d$ has Lebesgue measure zero, it suffices to show that

$$Z_n = \int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}\, \mathrm{d}\Pi \rightsquigarrow Z = \int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}\, \mathrm{d}\Pi.$$

Given δ, $\epsilon > 0$, call on Remark 4.1 to pick $M > 0$ such that $\Pr(\|\widehat{\mathbb{C}}^{\maltese}\|_{\mathcal{O}} > M) < \epsilon$ and $\Pr(\|\widehat{\mathbb{C}}_n^{\maltese}\| > M) < \epsilon$ for all $n \in \mathbb{N}$. Then choose a compact set $K = K_{\delta,\epsilon} \subset \mathcal{O}$ such that $\Pi(\mathcal{O} \setminus K) < \delta/M$. Now define

$$Y_{n,\delta,\epsilon} = \int_K \widehat{\mathbb{C}}_n^{\maltese}\, \mathrm{d}\Pi, \qquad Y_{\delta,\epsilon} = \int_K \widehat{\mathbb{C}}^{\maltese}\, \mathrm{d}\Pi.$$

Theorem 3.1 implies that $Y_{n,\delta,\epsilon} \rightsquigarrow Y_{\delta,\epsilon}$ as $n \to \infty$. Furthermore,

$$|Z_n - Y_{n,\delta,\epsilon}| = \left| \int_{\mathcal{O}\setminus K} \widehat{\mathbb{C}}_n^{\maltese} \, \mathrm{d}\Pi \right| \le \|\widehat{\mathbb{C}}_n^{\maltese}\| \Pi(\mathcal{O} \setminus K) < \frac{\delta}{M} \|\widehat{\mathbb{C}}_n^{\maltese}\|,$$

while

$$|Z - Y_{\delta,\epsilon}| = \left| \int_{\mathcal{O}\setminus K} \widehat{\mathbb{C}}^{\maltese} \, \mathrm{d}\Pi \right| \le \|\widehat{\mathbb{C}}^{\maltese}\|_{\mathcal{O}} \Pi(\mathcal{O} \setminus K) < \frac{\delta}{M} \|\widehat{\mathbb{C}}^{\maltese}\|_{\mathcal{O}}.$$

For all $n \in \mathbb{N}$, one then has $\Pr(|Z_n - Y_{n,\delta,\epsilon}| > \delta) \le \Pr(\|\widehat{\mathbb{C}}_n^{\maltese}\|\delta/M > \delta) < \epsilon$ and similarly $\Pr(|Z - Y_{\delta,\epsilon}| > \delta) < \epsilon$. The conclusion is then a consequence of Lemma C.1. □

The following lemma, needed for the proof of Proposition 5.1, is excerpted from Genest, Nešlehová and Rémillard [9].

Lemma C.2. *Let H be a distribution function on $\mathbb{R}^d$ and denote by H_n its empirical counterpart corresponding to a random sample of size n. If the sequence of processes $\mathbb{G}_n$ is tight with respect to the uniform norm on the space $\mathcal{C}_b(\mathbb{R}^d)$ of bounded and continuous functions on $\mathbb{R}^d$, then, as $n \to \infty$, $R_n = \int \mathbb{G}_n \, \mathrm{d}H_n - \int \mathbb{G}_n \, \mathrm{d}H \xrightarrow{\mathrm{p}} 0$.*

Proof of Proposition 5.1. Observe that

$$\sqrt{n}(\tau_n - \tau) = 4 \int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}(u,v) \, \mathrm{d}\widehat{C}_n^{\maltese}(u,v) + 4 \int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v).$$

First, it will be shown that, as $n \to \infty$,

$$Z_n = \int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v) \rightsquigarrow Z = \int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v).$$

To see this, fix arbitrary $\delta, \epsilon > 0$ and use Remark 4.1 to pick $M > 0$ such that $\Pr(\|\widehat{\mathbb{C}}^{\maltese}\|_{\mathcal{O}} > M) < \epsilon$ and $\Pr(\|\widehat{\mathbb{C}}_n^{\maltese}\| > M) < \epsilon$ for all $n \in \mathbb{N}$. Then choose a compact set $K = K_{\delta,\epsilon} \subset \mathcal{O}$ such that $C^{\maltese}(\mathcal{O} \setminus K) < \delta/M$. Setting

$$Y_{n,\delta,\epsilon} = \int_K \widehat{\mathbb{C}}_n^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v),$$
$$Y_{\delta,\epsilon} = \int_K \widehat{\mathbb{C}}^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v),$$

one can invoke Theorem 3.1 to deduce that $Y_{n,\delta,\epsilon} \rightsquigarrow Y_{\delta,\epsilon}$ as $n \to \infty$. The rest of the argument rests on Lemma C.1, in analogy to the proof of Proposition 5.2.

Secondly, to establish that, as $n \to \infty$,

$$\int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}(u,v) \, \mathrm{d}\widehat{C}_n^{\maltese}(u,v) \rightsquigarrow \int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}(u,v) \, \mathrm{d}C^{\maltese}(u,v), \tag{C.1}$$

use a change of variables and the definition of $H_n^{\maltese}$ to write

$$\begin{aligned}\int_{[0,1]^2} \widehat{\mathbb{C}}_n^{\maltese}(u,v)\,\mathrm{d}\widehat{C}_n^{\maltese}(u,v) &= \int_{\mathbb{R}^2} \widehat{\mathbb{C}}_n^{\maltese}\bigl\{F_{n1}^{\maltese}(x_1), F_{n2}^{\maltese}(x_2)\bigr\}\,\mathrm{d}H_n^{\maltese}(x_1,x_2)\\ &= \int_{\mathbb{R}^2} \mathbb{G}_n(x_1,x_2)\,\mathrm{d}H_n(x_1,x_2),\end{aligned}$$

where, for all $x_1, x_2 \in \mathbb{R}$,

$$\mathbb{G}_n(x_1,x_2) = \int_{[0,1]^2} \widehat{\mathbb{C}}_n^{\maltese}\bigl\{F_{n1}^{\maltese}(x_1+u-1), F_{n2}^{\maltese}(x_2+v-1)\bigr\}\,\mathrm{d}v\,\mathrm{d}u.$$

It is clear that $\|\mathbb{G}_n\| \le \|\widehat{\mathbb{C}}_n^{\maltese}\|$ and hence, by virtue of Remark 4.1, the sequence of processes $\mathbb{G}_n$ is tight on $\mathcal{C}_b(\mathbb{R}^2)$. Lemma C.2 thus implies that, as $n \to \infty$,

$$\left|\int_{\mathbb{R}^2} \mathbb{G}_n(x_1,x_2)\,\mathrm{d}H_n(x_1,x_2) - \int_{\mathbb{R}^2} \mathbb{G}_n(x_1,x_2)\,\mathrm{d}H(x_1,x_2)\right| \xrightarrow{\mathrm{p}} 0.$$

Undoing the change of variables and using the definitions of $H^{\maltese}$ and $C^{\maltese}$, one finds

$$\begin{aligned}\int_{\mathbb{R}^2} \mathbb{G}_n(x_1,x_2)\,\mathrm{d}H(x_1,x_2) &= \int_{\mathbb{R}^2} \widehat{\mathbb{C}}_n^{\maltese}\bigl\{F_{n1}^{\maltese}(x_1), F_{n2}^{\maltese}(x_2)\bigr\}\,\mathrm{d}H^{\maltese}(x_1,x_2)\\ &= \int_{[0,1]^2} \widehat{\mathbb{C}}_n^{\maltese}\bigl\{F_{n1}^{\maltese}\circ F_1^{\maltese-1}(u), F_{n2}^{\maltese}\circ F_2^{\maltese-1}(v)\bigr\}\,\mathrm{d}C^{\maltese}(u,v)\\ &= \int_{\mathcal{O}} \widehat{\mathbb{C}}_n^{\maltese}\bigl\{F_{n1}^{\maltese}\circ F_1^{\maltese-1}(u), F_{n2}^{\maltese}\circ F_2^{\maltese-1}(v)\bigr\}\,\mathrm{d}C^{\maltese}(u,v).\end{aligned}$$

Claim (C.1) is established if one can show that, as $n \to \infty$,

$$\bigl\|\widehat{\mathbb{C}}_n^{\maltese}\bigl(F_{n1}^{\maltese}\circ F_1^{\maltese-1}, F_{n2}^{\maltese}\circ F_2^{\maltese-1}\bigr) - \widehat{\mathbb{C}}_n^{\maltese}\bigr\|_K \xrightarrow{\mathrm{p}} 0 \tag{C.2}$$

for any fixed compact set $K \subset \mathcal{O}$. Given such a set, one can proceed exactly as in the proof of Proposition 4.4 to find $\delta^* > 0$ and a compact set $K^* \subset \mathcal{O}$ such that for all $\delta \in (0, \delta^*)$, $K \subset K_\delta \subset K^*$.

Next, fix $\epsilon > 0$ and $\delta \in (0, \delta^*)$ and recall that $\|\mathbb{C}_{nj}^{\maltese}\| \le \|\mathbb{B}_n\|$ for $j = 1, 2$. As in the proof of Proposition 4.4, one has that when $\{\|\mathbb{B}_n\|/\sqrt{n} < \delta/2\}$ holds,

$$\bigl(F_{n1}^{\maltese}\circ F_1^{\maltese-1}(u), F_{n2}^{\maltese}\circ F_2^{\maltese-1}(v)\bigr) = \left(u + \frac{\mathbb{C}_{n1}^{\maltese}(u)}{\sqrt{n}}, v + \frac{\mathbb{C}_{n2}^{\maltese}(v)}{\sqrt{n}}\right) \in K_\delta$$

whenever $(u, v) \in K$. Therefore,

$$P^*\bigl\{\bigl\|\widehat{\mathbb{C}}_n^{\maltese}\bigl(F_{n1}^{\maltese}\circ F_1^{\maltese-1}, F_{n2}^{\maltese}\circ F_2^{\maltese-1}\bigr) - \widehat{\mathbb{C}}_n^{\maltese}\bigr\|_K > \epsilon\bigr\}$$

is bounded above by

$$P^*\Big\{\sup_{\substack{(u,v),(u^*,v^*)\in K^*\\|u-u^*|+|v-v^*|<\delta}} \big|\widehat{\mathbb{C}}_n^{\maltese}(u^*,v^*) - \widehat{\mathbb{C}}_n^{\maltese}(u,v)\big| > \epsilon\Big\} + P^*\big(\|\mathbb{B}_n\|/\sqrt{n} \ge \delta/2\big).$$

Given that $\widehat{\mathbb{C}}_n^{\maltese}$ converges to $\widehat{\mathbb{C}}^{\maltese}$ on $\mathcal{C}(K^*)$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P^*\Big\{\sup_{\substack{(u,v),(u^*,v^*)\in K^*\\|u-u^*|+|v-v^*|<\delta}} \big|\widehat{\mathbb{C}}_n^{\maltese}(u^*,v^*) - \widehat{\mathbb{C}}_n^{\maltese}(u,v)\big| > \epsilon\Big\} = 0.$$

Claim (C.2) now readily follows from the fact that $\|\mathbb{B}_n\|/\sqrt{n} \xrightarrow{\mathrm{p}} 0$, as $n\to\infty$. In conclusion, $\sqrt{n}(\tau_n - \tau) \rightsquigarrow \mathcal{T}_2 = 8\int_{\mathcal{O}} \widehat{\mathbb{C}}^{\maltese}(u,v)\,\mathrm{d}C^{\maltese}(u,v)$ as $n\to\infty$, as claimed. □

Acknowledgments

The authors thank the Editors and referees for many helpful comments. They acknowledge grants from the Canada Research Chairs Program, the Natural Sciences and Engineering Research Council and the Fonds de recherche du Québec – Nature et technologies.

References

[1] Agresti, A. (2007). *An Introduction to Categorical Data Analysis*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley-Interscience. MR2293447

[2] Brockwell, P.J. and Davis, R.A. (1991). *Time Series*: *Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459

[3] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad*. *Roy*. *Belg*. *Bull*. *Cl*. *Sci*. (5) **65** 274–292. MR0573609

[4] Deheuvels, P. (1979). Non parametric tests of independence. In *Statistique non paramétrique asymptotique* (J.-P. Raoult, ed.) 95–107. Berlin: Springer.

[5] Dudley, R.M. (2002). *Real Analysis and Probability*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358

[6] Fermanian, J.-D., Radulović, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. MR2093613

[7] Genest, C. and Nešlehová, J. (2007). A primer on copulas for count data. *Astin Bull*. **37** 475–515. MR2422797

[8] Genest, C., Nešlehová, J. and Ruppert, M. (2011). Discussion: Statistical models and methods for dependence in insurance data. *J*. *Korean Statist*. *Soc*. **40** 141–148. MR2830507

[9] Genest, C., Nešlehová, J.G. and Rémillard, B. (2013). On the estimation of Spearman's rho and related tests of independence for possibly discontinuous multivariate data. *J*. *Multivariate Anal*. **117** 214–228. MR3053544

[10] Genest, C. and Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *TEST* **13** 335–370. MR2154005

[11] Lee, A.J. (1990). *U-Statistics*: *Theory and Practice*. *Statistics*: *Textbooks and Monographs* **110**. New York: Dekker Inc. MR1075417

[12] Murphy, O.A. (2013). Copula-based tests of independence for bivariate discrete data. Master's thesis, McGill Univ.

[13] Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2197664

[14] Nešlehová, J. (2007). On rank correlation measures for non-continuous random variables. *J. Multivariate Anal*. **98** 544–567. MR2293014

[15] Quessy, J.-F. (2009). Tests of multivariate independence for ordinal data. *Comm. Statist. Theory Methods* **38** 3510–3531. MR2589791

[16] Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist*. **4** 912–923. MR0420794

[17] Schmid, F. and Schmidt, R. (2007). Multivariate extensions of Spearman's rho and related statistics. *Statist. Probab. Lett*. **77** 407–416. MR2339046

[18] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. MR2948900

[19] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671

Received July 2012 and revised February 2013