Accepted Manuscript

Weak convergence of the weighted empirical beta copula process

Betina Berghaus, Johan Segers

Please cite this article as: B. Berghaus, J. Segers, Weak convergence of the weighted empirical beta copula process, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.03.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Weak convergence of the weighted empirical beta copula process

Betina Berghaus[a], Johan Segers[b]

[a]*Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstraße 150, 44780 Bochum, Germany*
[b]*Université catholique de Louvain, ISBA, Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium*

Abstract

The empirical copula has proved to be useful in the construction and understanding of many statistical procedures related to dependence within random vectors. The empirical beta copula is a smoothed version of the empirical copula that enjoys better finite-sample properties. At the core lie fundamental results on the weak convergence of the empirical copula and empirical beta copula processes. Their scope of application can be increased by considering weighted versions of these processes. In this paper we show weak convergence for the weighted empirical beta copula process. The weak convergence result for the weighted empirical beta copula process is stronger than the one for the empirical copula and its use is more straightforward. The simplicity of its application is illustrated for weighted Cramér–von Mises tests for independence and for the estimation of the Pickands dependence function of an extreme-value copula.

Keywords: Copula, Empirical beta copula, Empirical copula, Pickands dependence function, Weighted weak convergence

1. Introduction

In many statistical questions related to multivariate dependence, a crucial role is played by the copula. A basic nonparametric copula estimator is the empirical copula, dating back to [12, 31] and defined as the empirical distribution function of the vectors of component-wise ranks. The asymptotic behavior of the empirical copula has been established under various assumptions on the true copula and the serial dependence of the observed random vectors; see, e.g., [8, 16, 17, 33]. The upshot is that the empirical copula process converges weakly to a centered Gaussian field with covariance function depending on the true copula and the serial dependence of the observations.

Recently, Berghaus et al. [3] investigated the weak convergence of the weighted empirical copula process. They showed that the empirical copula process divided by a weight function, that can be zero on parts of the boundary of the unit cube, still converges weakly to a Gaussian field. As illustrated in the latter reference, this stronger result allows for additional applications of the continuous mapping theorem or the functional delta method. However, this result is only valid for a clipped version of the process. Since the empirical copula itself is not a copula, weak convergence fails on the upper boundaries of the unit cube; see Remark 2.3 in [3].

The empirical beta copula [34] arises as a particular case of the empirical Bernstein copula [see, e.g., 26, 32] if the degrees of the Bernstein polynomials are set to the sample size. In the numerical experiments in [34], the empirical beta copula exhibited a better performance than the empirical copula, both in terms of bias and variance.

In contrast to the empirical copula, the empirical beta copula is a genuine copula, a property that it shares with the checkerboard copula, whose limit is derived in [18] and which is very close to the empirical copula if the margins are continuous. Since the empirical beta copula is itself a copula, it is possible to prove weighted weak convergence for the empirical beta copula process on the whole unit cube. This is the main result of the paper. Weak convergence on the whole unit cube rather than on a subset thereof is quite handy since it allows for a direct application of, e.g., the continuous mapping theorem. In particular, there is no longer any need to treat the boundary regions separately.

We consider two applications. First, we modify the Cramér–von Mises test statistic for independence in [20] by using the empirical beta copula and, more importantly, adding a weight function in the integral, emphasizing the tails. The asymptotic distribution of the statistic under the null hypothesis is an easy corollary of our main result. More interestingly, the inclusion of a weight function leads to a markedly better power against difficult alternatives such as the t copula with zero correlation parameter, with favorable comparisons even to the novel statistics introduced recently by Belalia et al. [1]. As a second application we consider the Capéraà–Fougères–Genest estimator [9] of the Pickands dependence function of a multivariate extreme-value copula. Under weak dependence, replacing the empirical copula by the empirical beta copula yields a more accurate estimator. Its asymptotic distribution is again an immediate consequence of our main result.

The paper is organized as follows. In Section 2 we introduce the various empirical copula processes and we state the main result of the paper, the weighted convergence of the empirical beta copula process on the whole unit cube. We illustrate the ease of application of the main result to the analysis of weighted Cramér–von Mises tests of independence (Section 3) and nonparametric estimation of multivariate extreme-value copulas (Section 4). The proofs are deferred to Section 5, whereas a number of technical arguments are worked out in detail in Section 6.

2. Notation and main result

Let $(X_n)_n$ be a strictly stationary time series whose d-variate stationary distribution function F has continuous marginal distribution functions $F_1, \ldots, F_d$ and copula C. Writing $X_i = (X_{i,1}, \ldots, X_{i,d})$, we have, for $\boldsymbol{x} \in \mathbb{R}^d$,

$$\Pr(X_{i,j} \leqslant x_j) = F_j(x_j), \quad \Pr(X_i \leqslant \boldsymbol{x}) = F(\boldsymbol{x}) = C\{F_1(x_1), \ldots, F_d(x_d)\}.$$

For vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, the inequality $\boldsymbol{x} \leqslant \boldsymbol{y}$ means that $x_j \leqslant y_j$ for all $j \in \{1, \ldots, d\}$. Similar conventions apply for other inequalities and for minima and maxima, denoted by the operators $\wedge$ and $\vee$, respectively. Given the sample $X_1, \ldots, X_n$, the aim is to estimate C and functionals thereof.

Although the copula C captures the instantaneous (cross-sectional) dependence, the setting is still general enough to include questions about serial dependence. For instance, if $(Y_n)_n$ is a univariate, strictly stationary time series, then the d-variate time series of lagged values $X_n = (Y_n, \ldots, Y_{n-d+1})$ is strictly stationary too and the instantaneous dependence within the series $(X_n)_n$ corresponds to serial dependence within the original series $(Y_n)_n$ up to lag $d-1$.

For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$, let $R_{i,j}$ denote the (maximal) rank of $X_{i,j}$ among $X_{1,j}, \ldots, X_{n,j}$. For convenience, we omit the sample size n in the notation for ranks. The random vectors $\hat{U}_i = (\hat{U}_{i,1}, \ldots, \hat{U}_{i,d})$, with $\hat{U}_{i,j} = n^{-1} R_{i,j}$ and $i \in \{1, \ldots, n\}$, are called pseudo-observations from C. Letting $\mathbf{1}_A$ denote the indicator of the event A, the empirical copula is defined, for all $\boldsymbol{u} \in [0,1]^d$, by

$$\hat{C}_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\hat{U}_i \leqslant \boldsymbol{u}\}}.$$

Under mixing conditions on the sequence $(X_n)_n$ and smoothness conditions on C, Bücher and Volgushev [8] showed that, as $n \to \infty$,

$$\hat{\mathbb{C}}_n = \sqrt{n} (\hat{C}_n - C) \rightsquigarrow \mathbb{C}_C, \tag{1}$$

in the metric space $\ell^\infty([0,1]^d) = \{f : [0,1]^d \to \mathbb{R} : \sup_{\boldsymbol{u} \in [0,1]^d} |f(\boldsymbol{u})| < \infty\}$ equipped with the supremum distance. The arrow $\rightsquigarrow$ in (1) denotes weak convergence in metric spaces as exposed in [36]. The limit process in (1) is given, for all $\boldsymbol{u} \in [0,1]^d$, by

$$\mathbb{C}_C(\boldsymbol{u}) = \alpha_C(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u}) \, \alpha_C(1, \ldots, 1, u_j, 1, \ldots, 1),$$

where $\dot{C}_j(\boldsymbol{u}) = \partial C(\boldsymbol{u})/\partial u_j$ and where α_C is a tight, centered Gaussian process on $[0,1]^d$ with covariance function defined, for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, by

$$\operatorname{cov}\{\alpha_C(\boldsymbol{u}), \alpha_C(\boldsymbol{v})\} = \sum_{i \in \mathbb{Z}} \operatorname{cov}(\mathbf{1}_{\{U_0 \leqslant \boldsymbol{u}\}}, \mathbf{1}_{\{U_i \leqslant \boldsymbol{v}\}}), \tag{2}$$

2

where $\boldsymbol{U}_i = (U_{i,1}, \ldots, U_{i,d})$ and $U_{i,j} = F_j(X_{i,j})$. Since F_j is continuous, the random variables $U_{i,j}$ are uniformly distributed on $[0, 1]$. The joint distribution function of $\boldsymbol{U}_i$ is C. The margins $F_1, \ldots, F_d$ being unknown, we cannot observe the $\boldsymbol{U}_i$, and this is why we use the $\hat{\boldsymbol{U}}_i$ instead. In the case of serial independence, weak convergence of $\hat{\mathbb{C}}_n$ has been investigated by many authors, see the survey by Bücher and Volgushev [8]; the series in (2) simplifies to $\operatorname{cov}(\mathbf{1}_{\{U_0 \leqslant u\}}, \mathbf{1}_{\{U_0 \leqslant v\}}) = C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})$ so that α_C is a C-Brownian bridge. In the stationary case, convergence of the series in (2) is a consequence of the mixing conditions imposed on $(X_n)_n$.

Weak convergence in (1) is helpful for deriving asymptotic properties of estimators and test statistics based upon the empirical copula, such as estimators of Kendall's tau or Spearman's rho or such as Kolmogorov–Smirnov and Cramér–von Mises statistics for testing independence. However, as argued by Berghaus et al. [3], sometimes weak convergence with respect to a stronger metric is required, i.e., a weighted supremum norm. Examples mentioned in the cited article include nonparametric estimators of the Pickands dependence function of an extreme-value copula and bivariate rank statistics with unbounded score functions such as the van der Waerden rank (auto-)correlation. This motivates the study of the weighted empirical copula process $\hat{\mathbb{C}}_n / g^\omega$, with $\omega \in (0, 1/2)$ and a suitable weight function g on $[0, 1]^d$. The limit of the empirical copula process is zero almost surely as soon as one of its arguments is zero or if all arguments but at most one are equal to one. We can thus hope to obtain weak convergence with respect to a weight function that vanishes at such points. A possible function with this property is defined, for all $\boldsymbol{u} \in [0, 1]^d$, by

$$g(\boldsymbol{u}) = \bigwedge_{j=1}^{d} \left\{ u_j \wedge \bigvee_{k \neq j} (1 - u_k) \right\}. \tag{3}$$

Note that $g(\boldsymbol{u})$ is small as soon as there exists j such that either u_j is small or else all other u_k are close to 1. The trajectories of the processes $\hat{\mathbb{C}}_n / g^\omega$ are not bounded on the unit cube, hence the processes cannot converge weakly in $\ell^\infty([0, 1]^d)$. A solution is to restrict the domain from $[0, 1]^d$ to sets of the form $[c/n, 1 - c/n]^d$ for $c \in (0, 1)$, or, more generally, to $\{\boldsymbol{v} \in [0, 1]^d : g(\boldsymbol{v}) \geqslant c/n\}$. Relying on such a workaround, Theorem 2.2 in [3] states weak convergence of the weighted empirical copula process $\hat{\mathbb{C}}_n / g^\omega$ to $\mathbb{C}_C / g^\omega$. Note that $g(\boldsymbol{v}) = 0$ if and only if $v_j = 0$ for some j or if there exists j such that $v_k = 1$ for all $k \neq j$, and that $\mathbb{C}_C(\boldsymbol{v}) = 0$ almost surely for such $\boldsymbol{v}$, too.

The empirical copula is a piecewise constant function whereas the estimation target is continuous. It is natural to consider smoothed versions of the empirical copula. Segers et al. [34] defined the empirical beta copula, for all $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0, 1]^d$, as

$$C_n^\beta(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} F_{n,R_{i,j}}(u_j), \tag{4}$$

where $F_{n,r}$ is the distribution function of the beta distribution $\mathcal{B}(r, n+1-r)$, i.e., $F_{n,r}(u) = \sum_{s=r}^{n} \binom{n}{s} u^s (1-u)^{n-s}$, for $u \in [0, 1]$ and $r \in \{1, \ldots, n\}$. Note that

$$C_n^\beta(\boldsymbol{u}) = \int_{[0,1]^d} \hat{C}_n(\boldsymbol{w}) \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}), \tag{5}$$

where $\mu_{n,\boldsymbol{u}}$ is the law of the random vector $(S_1/n, \ldots, S_d/n)$, with $S_1, \ldots, S_d$ being independent binomial random variables, $S_j \sim \operatorname{Bin}(n, u_j)$. In the absence of ties, the rank vector $(R_{1,j}, \ldots, R_{n,j})$ of the jth coordinate sample is a permutation of $(1, \ldots, n)$. As a consequence, the empirical beta copula can be shown to be a genuine copula, unlike the empirical copula.

Under a smoothness condition on C, it follows from Theorem 3.6(ii) in [34] that weak convergence in $\ell^\infty([0, 1]^d)$ of the empirical copula process $\hat{\mathbb{C}}_n$ in (1) to a limit process $\mathbb{C}$ with continuous trajectories is sufficient to conclude the weak convergence of the empirical beta copula process: in the space $\ell^\infty([0, 1]^d)$, we have, as $n \to \infty$,

$$\mathbb{C}_n^\beta = \sqrt{n}\,(C_n^\beta - C) = \hat{\mathbb{C}}_n + o_p(1) \rightsquigarrow \mathbb{C}. \tag{6}$$

The asymptotic distribution of the empirical beta copula is thus the same as the one of the empirical copula. Still, for finite samples, numerical experiments in [34] revealed the empirical beta copula to be more accurate in terms of both bias and variance.

Our aim is to extend the convergence statement in (6) for weighted versions $\mathbb{C}_n^\beta / g^\omega$, with g as in (3) and for suitable exponents $\omega > 0$. As the empirical beta copula is a genuine copula, the zero-set of $\mathbb{C}_n^\beta$ includes the zero-set of g, and

on this set we implicitly define $\mathbb{C}_n^\beta/g^\omega$ to be zero. With this convention, the sample paths of $\mathbb{C}_n^\beta/g^\omega$ are bounded on $[0, 1]^d$; see Lemma 8 below. We can therefore hope to prove weak convergence of $\mathbb{C}_n^\beta/g^\omega \rightsquigarrow \mathbb{C}_C/g^\omega$ in $\ell^\infty([0, 1]^d)$ without having to exclude those border regions of $[0, 1]^d$ where g is small, as was necessary in [3].

The analysis of $\mathbb{C}_n^\beta/g^\omega$ will be based on the one of $\hat{\mathbb{C}}_n/g^\omega$ via (5). We will therefore need the same smoothness condition on C as imposed in Berghaus et al. [3, Condition 2.1], combining Conditions 2.1 and 4.1 in [33]. Condition 1 below is satisfied by many copula families: in [33, Section 5], part (i) of the condition is verified for strict Archimedean copulas with continuously differentiable generators, whereas both parts of the condition are verified for the non-singular bivariate Gaussian copula and for bivariate extreme-value copulas with twice continuously differentiable Pickands dependence function and a growth condition on the latter's second derivative near the boundary points of its domain.

Condition 1. (i) For every $j \in \{1, \ldots, d\}$, the first-order partial derivative $\dot{C}_j(\boldsymbol{u})\partial C(\boldsymbol{u})/\partial u_j$ exists and is continuous on $V_j = \{\boldsymbol{u} \in [0, 1]^d : u_j \in (0, 1)\}$.

(ii) For every $j_2, j_2 \in \{1, \ldots, d\}$, the second-order partial derivative $\ddot{C}_{j_1 j_2}(\boldsymbol{u}) = \partial^2 C(\boldsymbol{u})/\partial u_{j_1} \partial u_{j_2}$ exists and is continuous on $V_{j_1} \cap V_{j_2}$. Moreover, there exists a constant $K > 0$ such that, for all $j_1, j_2 \in \{1, \ldots, d\}$, we have

$$\left| \ddot{C}_{j_1 j_2}(\boldsymbol{u}) \right| \leqslant K \min \left\{ \frac{1}{u_{j_1}(1 - u_{j_1})}, \frac{1}{u_{j_2}(1 - u_{j_2})} \right\} \tag{7}$$

for all $\boldsymbol{u} \in V_{j_1} \cap V_{j_2}$.

The alpha-mixing coefficients of the sequence $(X_n)_n$ are defined as

$$\alpha(k) = \sup\{|\Pr(A \cap B) - \Pr(A)\Pr(B)| : A \in \sigma(X_j, j \leqslant i), B \in \sigma(X_{j+k}, j \geqslant i), i \in \mathbb{Z}\},$$

for $k \in \{1, 2, \ldots\}$. The sequence $(X_n)_n$ is said to be strongly mixing or alpha-mixing if $\alpha(k) \to 0$ as $k \to \infty$. Now we can state the main result.

Theorem 2. *Suppose that* $X_1, X_2, \ldots$ *is a strictly stationary, alpha-mixing sequence with* $\alpha(k) = O(a^k)$, *as* $k \to \infty$, *for some* $a \in (0, 1)$. *Assume that within each variable, ties do not occur with probability one. If the copula C satisfies Condition 1, then, for any* $\omega \in [0, 1/2)$, *we have, in* $\ell^\infty([0, 1]^d)$, *as* $n \to \infty$, $\mathbb{C}_n^\beta/g^\omega \rightsquigarrow \mathbb{C}_C/g^\omega$.

Remark 3. The tie-excluding assumption is needed to ensure that the empirical beta copula is a genuine copula almost surely. The assumption implies that the d stationary marginal distributions are continuous. For iid sequences, continuity of the margins is also sufficient. In the strictly stationary case, ties may occur with positive probability even if the margins are continuous; for instance, take a Markov chain where the current state is repeated with positive probability.

Remark 4. The result also holds under weaker assumptions on the serial dependence. In [3] it is shown that weak convergence of the weighted empirical copula process is still valid under more general assumptions on the marginal empirical processes and quantile processes and an assumption on the multivariate empirical process. In this case, however, the range of ω is smaller; see Theorem 4.5 in [3].

3. Application: Weighted Cramér–von Mises tests for independence

Testing for independence is a classical subject which still attracts interest today. One approach consists of comparing the multivariate empirical cumulative distribution function to the product of empirical cumulative distribution functions. Integrating out the difference with respect to the sample distribution yields a Cramér–von Mises style test statistic going back to Hoeffding [25] and Blum et al. [4]. To achieve better power, one may, in the spirit of the Anderson–Darling goodness-of-fit test statistic, introduce a weight function in the integral that tends to infinity near (parts of) the boundary of the domain; see De Wet [11].

Deheuvels [13, 14] was perhaps the first to reformulate the question in the copula framework: for continuous variables, the problem consists in testing whether the true copula, C, is equal to the independence copula, $\Pi(\boldsymbol{u}) = u_1 \times \cdots \times u_d$. The empirical copula process $\sqrt{n}(\hat{C}_n - \Pi)$, for which he proposed an ingenious combinatorial transformation, can thus be taken as a basis for the construction of test statistics. Genest and Rémillard [20] relied on his ideas to test

the white noise hypothesis and considered Cramér–von Mises statistics based on the empirical copula process. Genest et al. [19] studied the power of such statistics against local alternatives, while Kojadinovic and Holmes [27] developed an extension to the case of testing for independence between random vectors. For the latter problem, Fan et al. [15] proposed an alternative approach based on empirical characteristic functions.

Recently, Belalia et al. [1] proposed to use the Bernstein empirical copula [26, 32] rather than the empirical copula in the Cramér–von Mises test statistic. Moreover, they constructed new test statistics based on the Bernstein copula density estimator by Bouezmarni et al. [5]. Recall that the empirical beta copula arises from the Bernstein empirical copula by a specific choice of the degree of the Bernstein polynomials.

A situation of particular interest is when the true copula differs from the independence copula mainly in the tails. For instance, the bivariate t copula with zero correlation parameter has both Spearman's rho and Kendall's tau equal to zero. Still, the common value of its coefficients of upper and lower tail dependence is positive and depends on the degrees-of-freedom parameter. In their numerical experiments, Belalia et al. [1] found that for such alternatives, the power of the Cramér–von Mises test based on both the empirical copula and the Bernstein empirical copula is particularly weak. Their test statistics based on the Bernstein copula density estimator performed much better.

To increase the power of the Cramér–von Mises statistic against such difficult alternatives, a natural approach is to follow De Wet [11] and introduce a weight function emphasizing the tails. For $\gamma \in [0, 2)$, we propose the weighted Cramér–von Mises statistic

$$T_{n,\gamma} = n \int_{[0,1]^d} \frac{\{C_n^\beta(\boldsymbol{u}) - C(\boldsymbol{u})\}^2}{\{g(\boldsymbol{u})\}^\gamma} \, \mathrm{d}\boldsymbol{u}. \tag{8}$$

We are mostly interested in the case where $C(\boldsymbol{u}) = \Pi(\boldsymbol{u}) = \prod_{j=1}^d u_j$, the independence copula. If $\gamma = 0$, the weight function disappears and we are back to the original Cramér–von Mises statistic, but with the empirical beta copula replacing the empirical copula.

Corollary 5. *Under the assumptions of Theorem 2, we have, for every $\gamma \in [0, 2)$, the weak convergence*

$$T_{n,\gamma} \rightsquigarrow T_\gamma = \int_{[0,1]^d} \frac{\{\mathbb{C}_C(\boldsymbol{u})\}^2}{\{g(\boldsymbol{u})\}^\gamma} \, \mathrm{d}\boldsymbol{u}$$

as $n \to \infty$. This is particularly true in case of independent random sampling from a d-variate distribution with continuous margins and independent components ($C = \Pi$).

Proof. We have

$$T_{n,\gamma} = \int_{[0,1]^d} \left[\frac{\mathbb{C}_n^\beta(\boldsymbol{u})}{\{g(\boldsymbol{u})\}^{\gamma/4}} \right]^2 \frac{1}{\{g(\boldsymbol{u})\}^{\gamma/2}} \, \mathrm{d}\boldsymbol{u}.$$

By Theorem 2 applied to $\omega = \gamma/4 \in [0, 1/2)$, the first part of the integrand converges weakly, in $\ell^\infty([0, 1]^d)$, to the stochastic process $(\mathbb{C}_C/g^{\gamma/4})^2$. Further, since $\gamma/2 \in [0, 1)$, the integral $\int_{[0,1]^d} \{g(\boldsymbol{u})\}^{-\gamma/2} \, \mathrm{d}\boldsymbol{u}$ is finite. The linear functional that sends a measurable function $f \in \ell^\infty([0, 1]^d)$ to the scalar $\int_{[0,1]^d} f(\boldsymbol{u}) \, \{g(\boldsymbol{u})\}^{-\gamma/2} \, \mathrm{d}\boldsymbol{u}$ is therefore bounded. The conclusion follows from the continuous mapping theorem. $\square$

A comprehensive simulation study comparing the performance of the weighted Cramér–von Mises statistic against all competitors and for a wide range of tuning parameters and data-generating processes is out of this paper's scope. We limit ourselves to the case identified as the most difficult one in Belalia et al. [1], the bivariate t copula with zero correlation parameter. We copy the settings in their Section 5: the degrees-of-freedom parameter is $\nu = 2$ and we consider independent random samples of size $n \in \{100, 200, 400, 500\}$. We compare the power of our statistic $T_{n,\gamma}$ with the powers of their statistics T_n, δ_n, I_n at the $\alpha = 5\%$ significance level based on 1 000 replications.

We implemented our estimator in the statistical software environment R [30] using the package copula [28]. The critical values were computed by a Monte Carlo approximation based on 10 000 random samples from the uniform distribution on the unit square. Note that computing the critical values in this manner ensures that the proposed test holds its level. For the statistics in [1], we copied the relevant values from their Tables 4, 5, and 6. Their statistics depend on the degree, k, of the Bernstein polynomials, which they selected in $\{5, 10, \ldots, 30\}$. Note that for $\gamma = 0$ and $k = n$, our statistic $T_{n,\gamma}$ coincides with their statistic T_n. Their statistics δ_n and I_n are based on the Bernstein copula density estimator in [5].

5

	k	T_n	δ_n	I_n	γ	$T_{n,\gamma}$
$n = 100$	5	0.056	0.114	0.094	0.25	0.102
	10	0.064	0.130	0.168	0.50	0.091
	15	0.066	0.166	0.254	0.75	0.138
	20	0.070	0.132	0.270	1.00	0.179
	25	0.070	0.102	0.284	1.25	0.216
	30	0.068	0.114	0.294	1.50	0.292
					1.75	0.401
$n = 200$	5	0.076	0.176	0.094	0.25	0.123
	10	0.080	0.222	0.308	0.50	0.161
	15	0.088	0.226	0.442	0.75	0.233
	20	0.094	0.210	0.466	1.00	0.335
	25	0.096	0.176	0.472	1.25	0.428
	30	0.086	0.148	0.458	1.50	0.605
					1.75	0.705
$n = 400$	5	0.044	0.366	0.230	0.25	0.278
	10	0.038	0.492	0.588	0.50	0.427
	15	0.048	0.472	0.702	0.75	0.555
	20	0.044	0.432	0.762	1.00	0.777
	25	0.048	0.382	0.772	1.25	0.864
	30	0.050	0.354	0.780	1.50	0.930
					1.75	0.964
$n = 500$	5	0.072	0.398	0.192	0.25	0.406
	10	0.096	0.542	0.688	0.50	0.588
	15	0.100	0.552	0.746	0.75	0.773
	20	0.110	0.506	0.806	1.00	0.883
	25	0.106	0.476	0.824	1.25	0.966
	30	0.096	0.458	0.824	1.50	0.986
					1.75	0.992

Table 1: Testing the independence hypothesis when the true copula is equal to the t copula with zero correlation parameter and degrees-of-freedom parameter $\nu = 2$. Powers based on $1\,000$ random samples of sizes $n \in \{100, 200, 400, 500\}$ at significance level $\alpha = 5\%$. Comparison between, on the one hand, the statistics T_n, δ_n, I_n in Belalia et al. [1] with degree k of the Bernstein polynomials and, on the other hand, the weighted Cramér–von Mises statistic $T_{n,\gamma}$ in Eq. (8) with weight parameter γ. The values in the columns headed T_n, δ_n and I_n have been copied from Tables 4–6 in [1].

The results are presented in Table 1. The unweighted Cramér–von Mises statistic T_n does a poor job in detecting the alternative. The novel statistics δ_n and I_n in [1] are more powerful, especially the statistic I_n, which is a Cramér–von Mises statistic based on the Bernstein copula density estimator. For the weighted Cramér–von Mises statistic $T_{n,\gamma}$, the power increases with γ. For the largest considered value, $\gamma = 1.75$, the power is higher than the one of T_n, δ_n and I_n for any value of k considered.

4. Application: Nonparametric estimation of a Pickands dependence function

A d-variate copula C is a multivariate extreme-value copula if and only if it can be written as

$$C(\boldsymbol{u}) = \exp\left\{\left(\sum_{j=1}^{d} \ln u_j\right) A\left(\frac{\ln u_1}{\sum_{j=1}^{d} \ln u_j}, \ldots, \frac{\ln u_{d-1}}{\sum_{j=1}^{d} \ln u_j}\right)\right\},$$

6

for $\boldsymbol{u} \in (0, 1]^d \setminus \{(1, \ldots, 1)\}$. The function $A : \Delta_{d-1} \to [1/d, 1]$ is called the Pickands dependence function [after 29], its domain being the unit simplex $\Delta_{d-1} = \{\boldsymbol{t} = (t_1, \ldots, t_{d-1}) \in [0, 1]^{d-1} : t_1 + \cdots + t_{d-1} \leqslant 1\}$.

Writing $t_d = t_d(\boldsymbol{t}) = 1 - t_1 - \cdots - t_{d-1}$ for $\boldsymbol{t} \in \Delta_{d-1}$, we have $C(u^{t_1}, \ldots, u^{t_d}) = u^{A(\boldsymbol{t})}$ for all $u \in (0, 1)$, and thus

$$\ln\{A(\boldsymbol{t})\} = -\gamma + \int_0^1 \{C(u^{t_1}, \ldots, u^{t_d}) - \mathbf{1}_{[e^{-1}, 1]}(u)\} \frac{\mathrm{d}u}{u \ln u},$$

where $\gamma = 0.5772156649\ldots$ is the Euler–Mascheroni constant. The rank-based Capéraà–Fougères–Genest (CFG) estimator, $\hat{A}_n^{\mathrm{CFG}}(\boldsymbol{t})$, arises by replacing C in the above formula by the empirical copula, $\hat{C}_n$; see [9] for the original estimator and see [21, ?] for the rank-based versions in dimensions two and higher, respectively. We now propose to replace C by the empirical beta copula (4) instead, which gives the estimator

$$\ln\{\hat{A}_{n,\beta}^{\mathrm{CFG}}(\boldsymbol{t})\} = -\gamma + \int_0^1 \{C_n^{\beta}(u^{t_1}, \ldots, u^{t_d}) - \mathbf{1}_{[e^{-1}, 1]}(u)\} \frac{\mathrm{d}u}{u \ln u}. \tag{9}$$

The technique could also be used for other estimators based upon the empirical copula [2, 6].

For the CFG-estimator on usually employs the endpoint-corrected version

$$\ln\{\hat{A}_{n,\mathrm{c}}^{\mathrm{CFG}}(\boldsymbol{t})\} = \ln\{\hat{A}_n^{\mathrm{CFG}}(\boldsymbol{t})\} - \sum_{j=1}^d t_j \ln\{\hat{A}_n^{\mathrm{CFG}}(\boldsymbol{e}_j)\}, \tag{10}$$

where $\boldsymbol{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the jth canonical unit vector in $\mathbb{R}^d$. For the estimator based on the empirical beta copula the endpoint correction is immaterial, since C_n^{β} is a copula itself and thus $\ln \hat{A}_{n,\beta}^{\mathrm{CFG}}(\boldsymbol{e}_j) = 0$ for all $j \in \{1, \ldots, d\}$.

Thanks to Theorem 2, the limit of the beta CFG estimator can be derived from Theorem 2 by a straightforward application of the continuous mapping theorem. The result does not require serial independence and can be extended to higher dimensions.

Corollary 6. *Let C be a d-variate extreme-value copula with Pickands dependence function $A : \Delta_{d-1} \to [1/d, 1]$. Under the assumptions of Theorem 2 we have, as $n \to \infty$,*

$$\sqrt{n}\left\{\hat{A}_{n,\beta}^{\mathrm{CFG}}(\cdot) - A(\cdot)\right\} \rightsquigarrow \mathbb{A}(\cdot)$$

in $\ell^{\infty}(\Delta_{d-1})$, where, for $\boldsymbol{t} \in \Delta_{d-1}$, we define $\mathbb{A}(\boldsymbol{t}) = A(\boldsymbol{t}) \int_0^1 \mathbb{C}_C(u^{t_1}, \ldots, u^{t_d}) (u \ln u)^{-1} \mathrm{d}u$.

Proof. Let $\omega \in (0, 1/2)$. We have

$$\sqrt{n}\left[\ln\{\hat{A}_{n,\beta}^{\mathrm{CFG}}(\boldsymbol{t})\} - \ln\{A(\boldsymbol{t})\}\right] = \int_0^1 \mathbb{C}_n^{\beta}(u^{t_1}, \ldots, u^{t_d}) \frac{\mathrm{d}u}{u \ln u} = \int_0^1 \frac{\mathbb{C}_n^{\beta}(u^{t_1}, \ldots, u^{t_d})}{\{g(u^{t_1}, \ldots, u^{t_d})\}^{\omega}} \{g(u^{t_1}, \ldots, u^{t_d})\}^{\omega} \frac{\mathrm{d}u}{u \ln u}.$$

The integral $\int_0^1 \{g(u^{t_1}, \ldots, u^{t_d})\}^{\omega} (u \ln u)^{-1} \mathrm{d}u$ is bounded, uniformly in $\boldsymbol{t} \in \Delta_{d-1}$. Therefore, the linear map that sends a measurable function $f \in \ell^{\infty}([0, 1]^d)$ to the bounded function

$$\boldsymbol{t} \mapsto \int_0^1 f(u^{t_1}, \ldots, u^{t_d}) \{g(u^{t_1}, \ldots, u^{t_d})\}^{\omega} (u \ln u)^{-1} \mathrm{d}u$$

is continuous. By Theorem 2 and the continuous mapping theorem, we find, as $n \to \infty$,

$$\sqrt{n}\left[\ln\{\hat{A}_{n,\beta}^{\mathrm{CFG}}(\cdot)\} - \ln\{A(\cdot)\}\right] \rightsquigarrow \left(\int_0^1 \mathbb{C}_C(u^{t_1}, \ldots, u^{t_d}) \frac{\mathrm{d}u}{u \ln u}\right)_{\boldsymbol{t} \in \Delta_{d-1}}$$

in $\ell^{\infty}(\Delta_{d-1})$. Finally, the result follows by an application of the functional delta method. $\qquad\square$

7

We compare the finite-sample performance of the endpoint-corrected CFG estimator with the variant based on the empirical beta copula. As performance criterion for an estimator $\hat{A}$, we use the integrated mean squared error,

$$\int_{\Delta_{d-1}} \mathrm{E}\left[\{\hat{A}(t) - A(t)\}^2\right] \, \mathrm{d}t = \mathrm{E}\left[\{\hat{A}(T) - A(T)\}^2\right],$$

where the random variable T is uniformly distributed on Δ_{d-1} and is independent of the sample from which $\hat{A}$ was computed. We approximate the integrated mean squared error through a Monte Carlo procedure: for a large integer M, we generate M random samples of size n from a given copula and we calculate

$$\frac{1}{M} \sum_{m=1}^{M} \left\{\hat{A}_n^{(m)}(T^{(m)}) - A(T^{(m)})\right\}^2,$$

where $\hat{A}_n^{(m)}$ denotes the estimator based upon sample number m, and where the random variables $T^{(1)}, \ldots, T^{(m)}$ are uniformly distributed on Δ_{d-1} and are independent of each other and of the copula samples. The approximation error is $O_p(1/\sqrt{M})$, aggregating both the sampling error and the integration. A similar trick was used in [34] and is more efficient then first estimating the pointwise mean squared error through a Monte Carlo procedure and then integrating this out via numerical integration.

We considered the following data-generating processes:

(M1) Independent random sampling from the bivariate Gumbel copula [24], which has Pickands dependence function $A(t) = \{t^{1/\alpha} + (1-t)^{1/\alpha}\}^\alpha$ for $t \in [0, 1]$ and with parameter $\alpha \in [0, 1]$, for which Kendall's tau is $\tau = 1 - \alpha$. We also considered independent random samples from the bivariate Galambos, Hüsler–Reiss and t-EV copula families, yielding similar results as for the bivariate Gumbel copula, not shown to save space. See, e.g., [22] for the definitions of these copulas.

(M2) Independent random sampling from a special case of the trivariate asymmetric logistic extreme-value copula [35], with Pickands dependence function $A(t_1, t_2) = \sum_{(i,j) \in \{(1,2),(2,3),(3,1)\}} \{(\theta t_i)^{1/\alpha} + (\phi t_j)^{1/\alpha}\}^\alpha + 1 - \theta - \phi$ for $(t_1, t_2) \in \Delta_2$ and $t_3 = 1 - t_1 - t_2$. As in [23, Section 5], we set $\phi = 0.3$ and $\theta = 0.6$, and α varies between 0 and 1.

(M3) Sampling from the strictly stationary bivariate moving maximum process $(U_{t1}, U_{t2})_{t \in \mathbb{Z}}$ given by

$$U_{t1} = \max\left\{W_{t-1,1}^{1/a}, W_{t1}^{1/(1-a)}\right\} \quad \text{and} \quad U_{t2} = \max\left\{W_{t-1,2}^{1/b}, W_{t2}^{1/(1-b)}\right\},$$

where $a, b \in [0, 1]$ are two parameters and where $(W_{t1}, W_{t2})_{t \in \mathbb{Z}}$ is an iid sequence of bivariate random vectors whose common distribution is an extreme value-copula with some Pickands dependence function B. By Eq. (8.1) in [7], the stationary distribution of (U_{t1}, U_{t2}) is an extreme-value copula too, and its Pickands dependence function can be easily calculated to be

$$A(t) = \{a(1-t) + bt\} B\left\{\frac{bt}{a(1-t) + bt}\right\} + \{(1-a)(1-t) + (1-b)t\} B\left\{\frac{(1-b)t}{(1-a)(1-t) + (1-b)t}\right\},$$

for $t \in [0, 1]$. We let $B(t) = \{t^{1/\alpha} + (1-t)^{1/\alpha}\}^\alpha$ for $t \in [0, 1]$ and $\alpha \in [0, 1]$ (the bivariate Gumbel copula as above, with Kendall's tau $\tau = 1 - \alpha$) and we set $a = 0.1$ and $b = 0.7$, so that A is asymmetric.

The results are shown in Figure 1. Each plot is based on 10,000 samples of size $n \in \{20, 50, 100\}$. For weak dependence (small τ, large α), the beta variant (9) is the more efficient one, whereas for strong dependence (large τ, small α), it is the usual CFG estimator (10) which is more accurate.

In order to gain a better understanding, we have also traced some trajectories of estimated Pickands dependence functions for independent random samples of the bivariate Gumbel copula at $\tau \in \{0.3, 0.9\}$ and $n \in \{20, 50, 100\}$; see Figure 2. For each trajectory of the CFG estimator, there is a corresponding trajectory of the new estimator that is based on the same sample. For large τ, the true extreme-value copula C is close to the Fréchet–Hoeffding upper bound, $M(u_1, u_2) = \max(u_1, u_2)$. As a result, C is strongly curved around the main diagonal $u_1 = u_2$, and this implies a strong curvature of the Pickands dependence function A around $t = 1/2$. The empirical beta copula can be seen as a smoothed version of the empirical copula with an implicit bandwidth of the order $1/\sqrt{n}$ [34, p. 47]. For smaller n, oversmoothing occurs, producing a negative bias for the empirical beta copula around the diagonal $u_1 = u_2$ and thus a positive bias for the beta variant of the CFG estimator around $t = 1/2$.
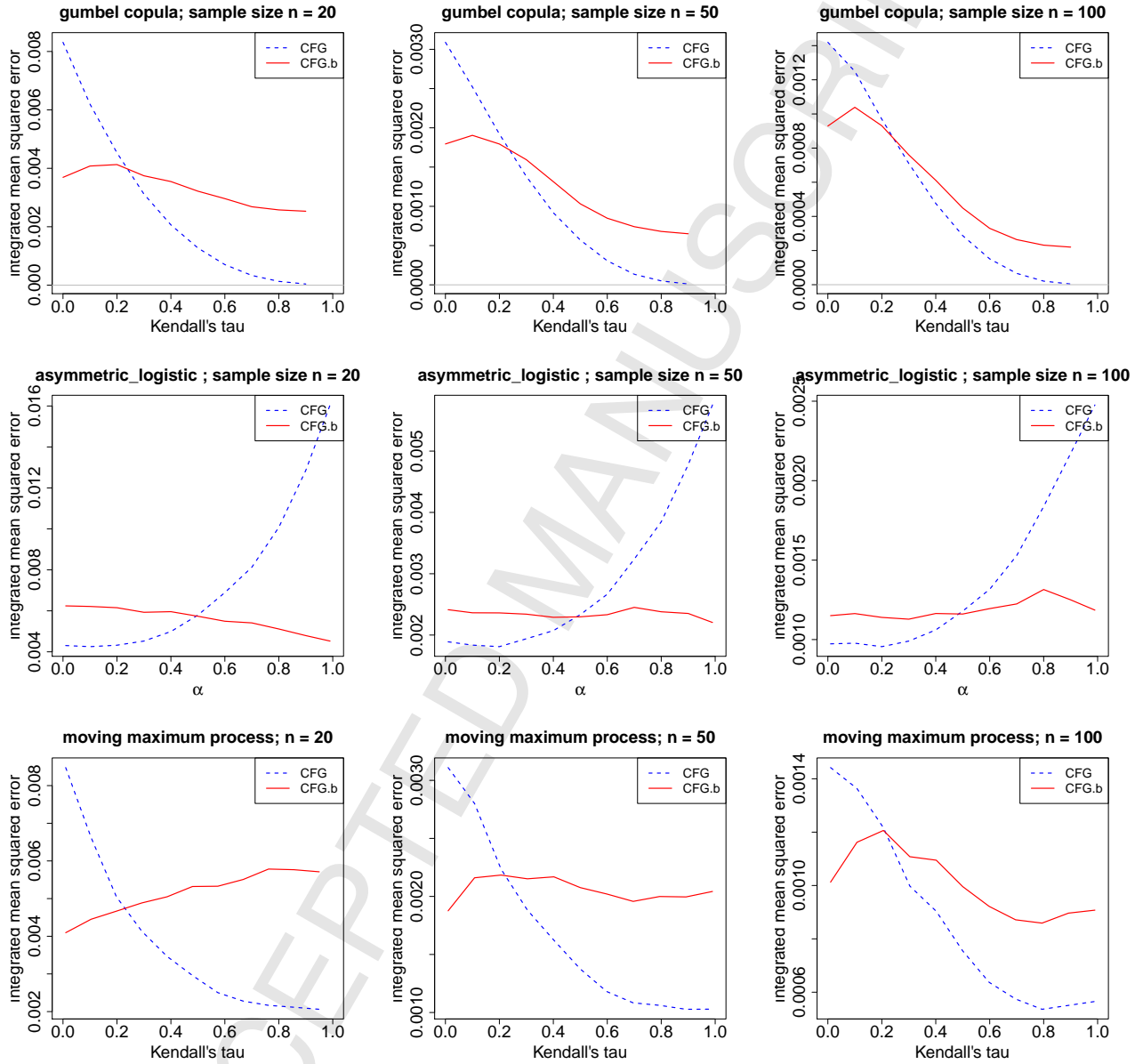
Figure 1: Integrated mean squared error (vertical axis) of the endpoint-corrected CFG-estimator (10) (dashed, blue) and the empirical beta variant (9) (solid, red) based on samples of the data-generating process (M1), (M2) and (M3) (top to bottom) for various choices of the parameter α or $\tau = 1 - \alpha$. Each point is based on 10 000 random samples of size $n \in \{20, 50, 100\}$ (left to right).
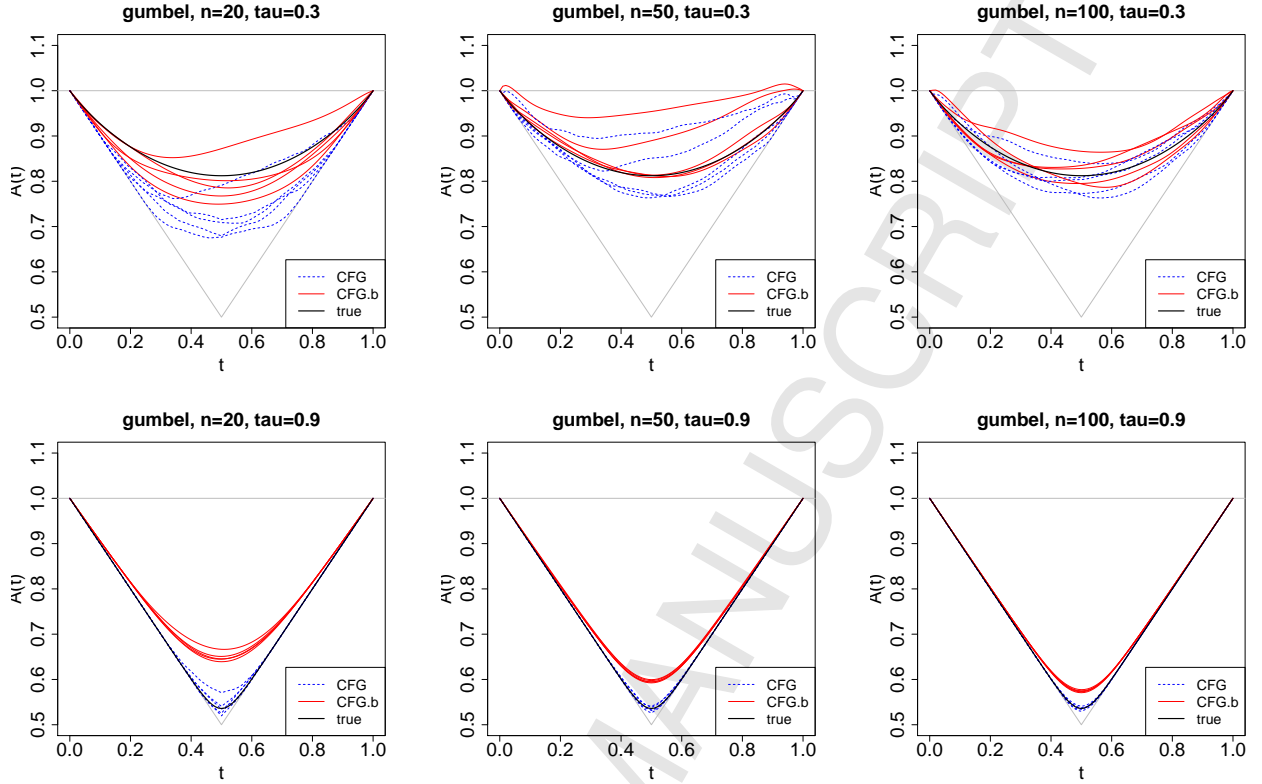
9

Figure 2: Plots of trajectories of the CFG-estimator (10) (dashed blue) and the empirical beta variant (9) (solid red) of the Pickands dependence function (solid black) based on samples from the Gumbel copula with Kendall's $\tau \in \{0.3, 0.9\}$ (top to bottom) and $n \in \{20, 50, 100\}$ (left to right).

5. Proof of Theorem 2

Recall the empirical copula process $\hat{\mathbb{C}}_n = \sqrt{n}(\hat{C}_n - C)$ and the empirical beta copula process $\hat{\mathbb{C}}_n^\beta = \sqrt{n}(C_n^\beta - C)$. The link between the empirical copula $\hat{C}_n$ and the empirical beta copula C_n^β is given in (5). In the derivation of the limit of the weighted empirical beta copula process, the following decomposition plays a central role:

$$\frac{\mathbb{C}_n^\beta(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} = \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} \int_{[0,1]^d} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w})$$
$$+ \int_{[0,1]^d} \left\{ \frac{\hat{\mathbb{C}}_n(\boldsymbol{w})}{g(\boldsymbol{w})^\omega} - \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} \right\} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) + \int_{[0,1]^d} \sqrt{n} \, \frac{C(\boldsymbol{w}) - C(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}). \quad (11)$$

It is reasonable to assume that the last two terms on the right-hand side vanish as $n \to \infty$. Indeed, the measure $\mu_{n,\boldsymbol{u}}$ concentrates around its mean $\boldsymbol{u}$, if the sample size grows, and both integrands are small if $\boldsymbol{w}$ is close to $\boldsymbol{u}$. By the same reasoning, the integral in the first term should be close to 1. The decomposition can thus be used to obtain weak convergence of $\mathbb{C}_n^\beta/g^\omega$ on the interior of the unit cube. The boundary of the unit cube has to be treated separately.

The case $\omega = 0$ corresponds to the unweighted case, so we assume henceforth that $\omega \in (0, 1/2)$. Fix a scalar γ such that $1/\{2(1 - \omega)\} < \gamma < 1$. Consider the abbreviations $\{g \geqslant n^{-\gamma}\} = \{\boldsymbol{v} \in [0,1]^d : g(\boldsymbol{v}) \geqslant n^{-\gamma}\}$ and similarly $\{g < n^{-\gamma}\}$. By Lemma 8, we have, as $n \to \infty$,

$$\mathbb{C}_n^\beta/g^\omega = \mathbb{C}_n^\beta/g^\omega \, \mathbf{1}_{\{g \geqslant n^{-\gamma}\}} + \mathbb{C}_n^\beta/g^\omega \, \mathbf{1}_{\{g < n^{-\gamma}\}} = \mathbb{C}_n^\beta/g^\omega \, \mathbf{1}_{\{g \geqslant n^{-\gamma}\}} + o(1), \quad \text{a.s.}$$

The three terms on the right-hand side of (11) are treated in Lemmas 9, 10 and 11. We find, as $n \to \infty$,

$$\mathbb{C}_n^\beta/g^\omega = \hat{\mathbb{C}}_n/g^\omega \, \mathbf{1}_{\{g \geqslant n^{-\gamma}\}} \{1 + o(1)\} + o_p(1). \quad (12)$$

10

Recall $\boldsymbol{U}_i = (U_{i,1}, \ldots, U_{i,d})$ with $U_{i,j} = F_j(X_{i,j})$. The empirical distribution function and the empirical process associated to the unobservable sample $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n$ are

$$C_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{U_i \leqslant \boldsymbol{u}\}}, \quad \alpha_n(\boldsymbol{u}) = \sqrt{n} \{C_n(\boldsymbol{u}) - C(\boldsymbol{u})\},$$

respectively, for $\boldsymbol{u} \in [0, 1]^d$. Consider the process defined, for all $\boldsymbol{u} \in [0, 1]^d$, by

$$\bar{\mathbb{C}}_n(\boldsymbol{u}) = \alpha_n(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u}) \, \alpha_n(1, \ldots, 1, u_j, 1, \ldots, 1),$$

with u_j appearing at the jth coordinate. Note the slight but convenient abuse of notation in the definition of $\bar{\mathbb{C}}_n$: if $\boldsymbol{u}$ is such that $u_j \in \{0, 1\}$, then $\alpha_n(1, \ldots, 1, u_j, 1, \ldots, 1) = 0$ almost surely, so that the fact that for such $\boldsymbol{u}$, the partial derivative $\dot{C}_j(\boldsymbol{u})$ has been left undefined in Condition 1 plays no role.

By (12) above and by Theorem 2.2 in [3] (see also Remark 7 below), as $n \to \infty$,

$$\mathbb{C}_n^{\beta}/g^{\omega} = \{\bar{\mathbb{C}}_n/g^{\omega} + o_p(1)\} \, \mathbf{1}_{\{g \geqslant n^{-\gamma}\}}\{1 + o(1)\} + o_p(1) = \bar{\mathbb{C}}_n/g^{\omega} \, \mathbf{1}_{\{g \geqslant n^{-\gamma}\}} + o_p(1).$$

In view of Lemma 4.9 in [3], the indicator function can be omitted, and, applying Theorem 2.2 in the same reference again, we deduce that, as $n \to \infty$,

$$\mathbb{C}_n^{\beta}/g^{\omega} = \bar{\mathbb{C}}_n/g^{\omega} + o_p(1) \rightsquigarrow \mathbb{C}_C/g^{\omega},$$

as required. This completes the proof of Theorem 2.

Remark 7. Some of the results in [3] have to be adapted to the present situation.

a) In the latter reference, the pseudo-observations are defined as $\hat{U}_{i,j} = R_{i,j}/(n + 1)$ rather than $R_{i,j}/n$. However, this does not affect the asymptotics, since the difference of the two empirical copulas is at most d/n, almost surely. For $\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}$, this modification makes a difference of the order $O_p(n^{\gamma\omega+1/2-1}) = o_p(1)$, as $n \to \infty$.

b) In Theorem 2.2 in [3], the approximation of $\hat{\mathbb{C}}_n$ by $\bar{\mathbb{C}}_n$ is stated on the interior of the set $[c/n, 1 - c/n]^d$ for any $c \in (0, 1)$. But it can be seen in the proof of the latter statement that the result can be easily extended to the set $\{g \geqslant c/n\}$. See Section 6.5 below for details.

6. Auxiliary results

Throughout and unless otherwise stated, we assume the conditions of Theorem 2.

6.1. Negligibility of the boundary regions

Lemma 8. *For $\gamma > 1/\{2(1 - \omega)\}$, we have, as $n \to \infty$,*

$$\sup_{\boldsymbol{u} \in \{g \leqslant n^{-\gamma}\}} |\mathbb{C}_n^{\beta}(\boldsymbol{u})/g(\boldsymbol{u})^{\omega}| = o(1), \quad a.s.$$

Proof. Let $\gamma > 1/\{2(1 - \omega)\}$ and $\boldsymbol{u} \in \{g \leqslant n^{-\gamma}\}$. Without loss of generality, we only need to consider the cases $g(\boldsymbol{u}) = u_1$ and $g(\boldsymbol{u}) = 1 - u_1$. The remaining cases can be treated analogously.

Start with the case $g(\boldsymbol{u}) = u_1 \leqslant n^{-\gamma}$. Since C_n^{β} is a copula almost surely, we have $C_n^{\beta}(\boldsymbol{u}) \leqslant u_1$. This in turn gives us

$$|\mathbb{C}_n^{\beta}(\boldsymbol{u})/g(\boldsymbol{u})^{\omega}| \leqslant \sqrt{n} \, u_1^{-\omega} |C_n^{\beta}(\boldsymbol{u}) + C(\boldsymbol{u})| \leqslant 2 \sqrt{n} \, u_1^{1-\omega} \leqslant 2n^{1/2+\gamma\omega-\gamma}, \quad a.s.,$$

an upper bound which vanishes as $n \to \infty$ by the choice of γ.

Now suppose that $g(\boldsymbol{u}) = 1 - u_1 \leqslant n^{-\gamma}$. By the definition of $g(\boldsymbol{u})$, we can assume without loss of generality that $1 - u_j \leqslant 1 - u_1$ for all $j \in \{3, \ldots, d\}$. Again, we will use the fact that C_n^{β} is a copula almost surely. Note that $C_n^{\beta}(1, u_2, 1, \ldots, 1) = u_2$. Hence, by the Lipschitz continuity of copulas we obtain, almost surely,

$$|\mathbb{C}_n^{\beta}(\boldsymbol{u})/g(\boldsymbol{u})^{\omega}| \leqslant \sqrt{n}(1 - u_1)^{-\omega}\{|C_n^{\beta}(\boldsymbol{u}) - u_2| + |C(\boldsymbol{u}) - u_2|\} \leqslant 2 \sqrt{n}(1 - u_1)^{-\omega} \sum_{j \neq 2}(1 - u_j)$$
$$\leqslant 2 \sqrt{n} \sum_{j \neq 2}(1 - u_j)^{1-\omega} \leqslant 2(d - 1)n^{1/2+\gamma\omega-\gamma} = o(1)$$

as $n \to \infty$. The upper bounds do not depend on $\boldsymbol{u} \in \{g \leqslant n^{-\gamma}\}$, whence the uniformity in $\boldsymbol{u}$. $\qquad\square$

11

6.2. The three terms in the decomposition (11)

The following lemma is to be compared with Proposition 3.5 in [34]. There, a pointwise approximation rate of $O(1/n)$ was established. Here, we state a rate which is slightly slower, $O\{\ln(n)/n\}$, but uniformly in $\boldsymbol{u}$.

Lemma 9. *If C satisfies Condition 1, then, as $n \to \infty$,*

$$\sup_{\boldsymbol{u} \in [0,1]^d} \left| \int_{[0,1]^d} \{C(\boldsymbol{w}) - C(\boldsymbol{u})\} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \right| = O\{\ln(n)/n\}.$$

Proof. Put $\varepsilon_n = \ln(n)/n$. First, we show that we can ignore those $\boldsymbol{u}$ for which $u_j \leqslant \varepsilon_n$ for some $j \in \{1, \ldots, d\}$. Indeed, for such $\boldsymbol{u}$, the absolute value in the statement is bounded by

$$\int_{[0,1]^d} w_j \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) + u_j = 2u_j \leqslant 2\varepsilon_n.$$

Let $\boldsymbol{u} \in [\varepsilon_n, 1]^d$. We show how to reduce the analysis to the case where $\boldsymbol{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$. Let $J = J(\boldsymbol{u})$ denote the set of indices $j \in \{1, \ldots, d\}$ such that $u_j > 1 - \varepsilon_n$ and suppose that J is not empty. Consider the vector $\boldsymbol{e} \in \{0, 1\}^d$ which has components $e_j = 1$ for $j \in J$ and $e_j = 0$ otherwise. For $\boldsymbol{v} \in [0, 1]^d$, the vector $\boldsymbol{v} \vee \boldsymbol{e}$ has components $(\boldsymbol{v} \vee \boldsymbol{e})_j$ equal to v_j if $j \notin J$ and to 1 if $j \in J$. Recall that copulas are Lipschitz continuous with respect to the L^1 norm with Lipschitz constant 1. It follows that

$$\left| \int_{[0,1]^d} \{C(\boldsymbol{w}) - C(\boldsymbol{u})\} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \right| \leqslant \left| \int_{[0,1]^d} \{C(\boldsymbol{w} \vee \boldsymbol{e}) - C(\boldsymbol{u} \vee \boldsymbol{e})\} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \right|$$
$$+ \int_{[0,1]^d} |C(\boldsymbol{w}) - C(\boldsymbol{w} \vee \boldsymbol{e})| \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) + |C(\boldsymbol{u} \vee \boldsymbol{e}) - C(\boldsymbol{u})|.$$

The first integral on the right-hand side does not depend on the variables w_j for $j \in J$. It can therefore be reduced to an integral as in the statement of the lemma with respect to the variables in the set $\{1, \ldots, d\} \setminus J$. The copula of those variables is a multivariate margin of the original copula and Condition 1 applies to it as well. By construction, all remaining u_j are in the interval $[\varepsilon_n, 1 - \varepsilon_n]$, as required.

Next, note that we have

$$|C(\boldsymbol{w}) - C(\boldsymbol{w} \vee \boldsymbol{e})| \leqslant \sum_{j \in J} |w_j - 1| \leqslant \sum_{j \in J} (|w_j - u_j| + \varepsilon_n).$$

By the Cauchy–Schwarz inequality,

$$\int_{[0,1]^d} |w_j - u_j| \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \leqslant \{n^{-1} u_j (1 - u_j)\}^{1/2} \leqslant n^{-1/2} \varepsilon_n^{1/2} \leqslant \varepsilon_n.$$

Hence $\int_{[0,1]^d} |C(\boldsymbol{w}) - C(\boldsymbol{w} \vee \boldsymbol{e})| \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \leqslant 2d\varepsilon_n$. Finally,

$$|C(\boldsymbol{u} \vee \boldsymbol{e}) - C(\boldsymbol{u})| \leqslant \sum_{j \in J} (1 - u_j) \leqslant d\varepsilon_n.$$

It remains to consider the case $\boldsymbol{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$. As in the proof of Proposition 3.4 in [34], we have

$$\int_{[0,1]^d} \{C(\boldsymbol{w}) - C(\boldsymbol{u})\} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) = \sum_{j=1}^d \int_0^1 \left[\int_{[0,1]^d} (w_j - u_j)\{\dot{C}_j(\boldsymbol{u} + t(\boldsymbol{w} - \boldsymbol{u})) - \dot{C}_j(\boldsymbol{u})\} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \right] \mathrm{d}t.$$

It is sufficient to show that the absolute value of the integral in square brackets is $O(\varepsilon_n)$, uniformly in $j \in \{1, \ldots, d\}$ and $t \in (0, 1)$ and $\boldsymbol{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$.

The integral over $[0, 1]^d$ can be reduced to an integral over $(0, 1)^d$: indeed, the integrand is bounded in absolute value by 1 (recall $0 \leqslant \dot{C}_j \leqslant 1$), and the mass on the boundary is $\mu_{n,\boldsymbol{u}}([0, 1]^d \setminus (0, 1)^d) = \Pr[\exists_j : S_j \in \{0, n\}] \leqslant 2d(1 - \varepsilon_n)^n \leqslant 2d \exp(-n\varepsilon_n) = 2d/n = o(\varepsilon_n)$ as $n \to \infty$.

12

In view of the second part of Condition 1, we have

$$\int_{(0,1)^d} (w_j - u_j)\{\dot{C}_j(\mathbf{u} + t(\mathbf{w} - \mathbf{u})) - \dot{C}_j(\mathbf{u})\}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w})$$

$$= t \sum_{k=1}^{d} \int_0^1 \left[\int_{(0,1)^d} (w_j - u_j)(w_k - u_k)\ddot{C}_{jk}\{\mathbf{u} + st(\mathbf{w} - \mathbf{u})\}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \right]\, \mathrm{d}s.$$

It is sufficient to show that the absolute value of the integral in square brackets is $O(\varepsilon_n)$, uniformly in $j, k \in \{1, \ldots, d\}$ and $s, t \in (0, 1)$ and $\mathbf{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$.

We apply the bound in (7) to $\ddot{C}_{jk}\{\mathbf{u} + st(\mathbf{w} - \mathbf{u})\}$. We have $\min(1/a, 1/b) \leqslant (ab)^{-1/2}$, and the latter is a convex function of $(a, b) \in (0, \infty)^2$. The point $\mathbf{u} + st(\mathbf{w} - \mathbf{u})$ is located on the line segment connecting $\mathbf{u}$ and $\mathbf{w}$. Therefore,

$$\left| \ddot{C}_{jk}\{\mathbf{u} + st(\mathbf{w} - \mathbf{u})\} \right| \leqslant K \left[\frac{1}{\{u_j(1 - u_j)u_k(1 - u_k)\}^{1/2}} + \frac{1}{\{w_j(1 - w_j)w_k(1 - w_k)\}^{1/2}} \right].$$

We obtain

$$\left| \int_{(0,1)^d} (w_j - u_j)(w_k - u_k)\ddot{C}_{jk}\{\mathbf{u} + st(\mathbf{w} - \mathbf{u})\}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \right|$$

$$\leqslant K \int_{(0,1)^d} \left[\frac{|(w_j - u_j)(w_k - u_k)|}{\{u_j(1 - u_j)u_k(1 - u_k)\}^{1/2}} + \frac{|(w_j - u_j)(w_k - u_k)|}{\{w_j(1 - w_j)w_k(1 - w_k)\}^{1/2}} \right]\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}).$$

First, by the Cauchy–Schwarz inequality and the fact that $\mathrm{E}[(S_i/n - u_i)^2] = n^{-1}u_i(1 - u_i)$ for all $i \in \{1, \ldots, d\}$, we have

$$\int_{(0,1)^d} \frac{|(w_j - u_j)(w_k - u_k)|}{\{u_j(1 - u_j)u_k(1 - u_k)\}^{1/2}}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \leqslant \prod_{i \in \{j,k\}} \left\{ \int_{(0,1)^d} \frac{(w_i - u_i)^2}{u_i(1 - u_i)}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \right\}^{1/2} \leqslant n^{-1} \leqslant \varepsilon_n.$$

Second, again by Cauchy–Schwarz inequality,

$$\int_{(0,1)^d} \frac{|(w_j - u_j)(w_k - u_k)|}{\{w_j(1 - w_j)w_k(1 - w_k)\}^{1/2}}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \leqslant \prod_{i \in \{j,k\}} \left\{ \int_{(0,1)^d} \frac{(w_i - u_i)^2}{w_i(1 - w_i)}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \right\}^{1/2}.$$

Each of the two integrals ($i = j$ and $i = k$), and therefore their geometric mean, will be bounded by the same quantity. Note that $1/\{w_i(1 - w_i)\} = 1/w_i + 1/(1 - w_i)$ and that the integral involving $1/(1 - w_i)$ is equal to the one involving $1/w_i$ when u_i is replaced by $1 - u_i$, which we are allowed to do since $\mathbf{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$ anyway. Therefore, we can replace $w_i(1 - w_i)$ by w_i in the denominator at the cost of a factor two. Further,

$$\int_{(0,1)^d} \frac{(w_i - u_i)^2}{w_i}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) \leqslant \int_{[0,1]^d} \mathbf{1}_{(0,1]}(w_i) \frac{(w_i - u_i)^2}{w_i}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) = \int_{[0,1]^d} \mathbf{1}_{(0,1]}(w_i)(w_i - 2u_i + u_i^2/w_i)\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w})$$

$$= u_i - 2u_i \Pr(S_i/n > 0) + u_i^2 \mathrm{E}[n\, \mathbf{1}_{\{S_i/n > 0\}}/S_i] \leqslant -u_i + 2\Pr(S_i = 0) + nu_i^2 \mathrm{E}[\mathbf{1}_{\{S_i \geqslant 1\}}/S_i].$$

Recall that $u_i \in [\varepsilon_n, 1 - \varepsilon_n]$ and thus $\Pr(S_i = 0) \leqslant (1 - \varepsilon_n)^n \leqslant \exp(-n\varepsilon_n) = n^{-1} = o(\varepsilon_n)$. Further, the expectation of the reciprocal of a binomial random variable is treated in Lemma 14. Note that $n\varepsilon_n = \ln n \to \infty$. We find, as $n \to \infty$,

$$\sup_{\mathbf{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d} \max_{i \in \{1, \ldots, d\}} \int_{(0,1)^d} \frac{(w_i - u_i)^2}{w_i(1 - w_i)}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) = O(n^{-1}) = o(\varepsilon_n).$$

The proof is complete. $\qquad\square$

Lemma 10. *For any* $1/\{2(1 - \omega)\} < \gamma < 1$, *we have, as* $n \to \infty$,

$$\sup_{\mathbf{u} \in \{g \geqslant n^{-\gamma}\}} \left| \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega}\, \mathrm{d}\mu_{n,\mathbf{u}}(\mathbf{w}) - 1 \right| = O\{n^{-(1-\gamma)/2} \ln(n)\}.$$

13

Proof. Since $g(S_1/n, \ldots, S_d/n)$ is a random variable taking values in $[0, 1]$, we can write

$$\int_{[0,1]^d} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, d\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) = \frac{1}{g(\boldsymbol{u})^\omega} \mathbb{E}[g(S_1/n, \ldots, S_d/n)^\omega] = \frac{1}{g(\boldsymbol{u})^\omega} \int_0^1 \Pr\{g(S_1/n, \ldots, S_d/n) > t^{1/\omega}\} \, dt. \qquad (13)$$

Split the integral into two pieces, $\int_0^{a_{n,\pm}} + \int_{a_{n,\pm}}^1$, where $a_{n,\pm} = a_{n,\pm}(\boldsymbol{u}) = g(\boldsymbol{u})^\omega (1 \pm \varepsilon_n)^\omega$. Write $\varepsilon_n = n^{-(1-\gamma)/2} \ln n$. Recall that $\omega \in (0, 1/2)$.

On the one hand, we find

$$
\begin{aligned}
\int_{[0,1]^d} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, d\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) &\leqslant \frac{a_{n,+}}{g(\boldsymbol{u})^\omega} + \frac{1 - a_{n,+}}{g(\boldsymbol{u})^\omega} \Pr\{g(S_1/n, \ldots, S_d/n) > a_{n,+}^{1/\omega}\} \\
&\leqslant (1 + \varepsilon_n)^\omega + g(\boldsymbol{u})^{-\omega} \Pr\{g(S_1/n, \ldots, S_d/n) > g(\boldsymbol{u})(1 + \varepsilon_n)\} \\
&\leqslant 1 + \varepsilon_n + g(\boldsymbol{u})^{-\omega} 2d \exp\{-ng(\boldsymbol{u})h(1 + \varepsilon_n)\}
\end{aligned}
$$

where we used (18) in the last step. Since $h(1 + \varepsilon_n) \geqslant \varepsilon_n^2/3$ for $\varepsilon_n \in [0, 1]$ and since $g(\boldsymbol{u}) \geqslant n^{-\gamma}$, the upper bound is bounded by

$$1 + \varepsilon_n + 2dn^{\gamma\omega} \exp\{-(\ln n)^2/3\} = 1 + \varepsilon_n + o(\varepsilon_n)$$

as $n \to \infty$. On the other hand, restricting the integral in (13) to $[0, a_{n,-}]$, we have

$$
\begin{aligned}
\int_{[0,1]^d} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, d\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) &\geqslant \frac{a_{n,-}}{g(\boldsymbol{u})^\omega} \Pr\{g(S_1/n, \ldots, S_d/n) > a_{n,-}^{1/\omega}\} \\
&= (1 - \varepsilon_n)^\omega \Pr\{g(S_1/n, \ldots, S_d/n) > g(\boldsymbol{u})(1 - \varepsilon_n)\} \\
&\geqslant (1 - \varepsilon_n)^\omega [1 - 4d \exp\{-ng(\boldsymbol{u})h(1 + \varepsilon_n)\}],
\end{aligned}
$$

where we used (19) in the last step. Since $0 \leqslant \varepsilon_n \to 0$ and $g(\boldsymbol{u}) \geqslant n^{-\gamma}$, the lower bound is bounded from below by

$$(1 - \varepsilon_n)^\omega [1 - 4d \exp\{-ng(\boldsymbol{u})h(1 + \varepsilon_n)\}] \geqslant (1 - \varepsilon_n)[1 - 4d \exp\{-(\ln n)^2/3\}] \geqslant 1 - \varepsilon_n - o(\varepsilon_n)$$

as $n \to \infty$. $\qquad \square$

Lemma 11. *As $n \to \infty$, we have, for any $\gamma \in (1/\{2(1 - \omega)\}, 1)$*

$$\sup_{\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}} \left| \int_{[0,1]^d} \left\{ \frac{\hat{\mathbb{C}}_n(\boldsymbol{w})}{g(\boldsymbol{w})^\omega} - \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})} \right\} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, d\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \right| = o_p(1).$$

Proof. Let $\delta_n = 1/\ln(n)$. Write $\|\hat{\mathbb{C}}_n\|_\infty = \sup\{|\hat{\mathbb{C}}_n(\boldsymbol{v})| : \boldsymbol{v} \in [0, 1]^d\}$. We have $\|\hat{\mathbb{C}}_n\|_\infty = O_p(1)$ as $n \to \infty$ by weak convergence of $\hat{\mathbb{C}}_n$ in $\ell^\infty([0, 1]^d)$. We split the integral over $\boldsymbol{w} \in [0, 1]^d$ into two pieces: the integral over the domain

$$A_{n,\boldsymbol{u}} = \{\boldsymbol{w} \in [0, 1]^d : |\boldsymbol{w} - \boldsymbol{u}|_\infty > \delta_n\} \cup \{g < n^{-\gamma}(1 - \delta_n)\}$$

and the integral over its complement; here $|\boldsymbol{x}|_\infty = \max(|x_1|, \ldots, |x_d|)$.

For all $\boldsymbol{w} \in [0, 1]^d$ and all $\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}$, we have

$$R_n(\boldsymbol{u}, \boldsymbol{w}) = \left| \frac{\hat{\mathbb{C}}_n(\boldsymbol{w})}{g(\boldsymbol{w})^\omega} - \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} \right| \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \leqslant \frac{|\hat{\mathbb{C}}_n(\boldsymbol{w})|}{g(\boldsymbol{u})^\omega} + \frac{|\hat{\mathbb{C}}_n(\boldsymbol{u})|}{g(\boldsymbol{u})^{2\omega}} \leqslant 2\|\hat{\mathbb{C}}_n\|_\infty n^{2\gamma\omega}.$$

Moreover, for all $\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}$, using Chebyshev's inequality and the concentration inequality (19), we have

$$
\begin{aligned}
\mu_{n,\boldsymbol{u}}(A_{n,\boldsymbol{u}}) &\leqslant \sum_{j=1}^d \Pr\{|S_j/n - u_j| > \delta_n\} + \Pr\{g(S_1/n, \ldots, S_d/n) < g(\boldsymbol{u})(1 - \delta_n)\} \\
&\leqslant dn^{-1}\delta_n^{-2} + 4d \exp\{-n^{1-\gamma}h(1 + \delta_n)\}.
\end{aligned}
$$

14

Since $\omega \in (0, 1/2)$, $\gamma \in (0, 1)$, $\delta_n = 1/\ln(n)$ and $h(1 + \delta_n) \geqslant \delta_n^2/3$, it follows that, as $n \to \infty$,

$$\sup_{\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}} \int_{A_{n,\boldsymbol{u}}} R_n(\boldsymbol{u}, \boldsymbol{w}) \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \leqslant n^{2\gamma\omega} [n^{-1}\delta_n^{-2} + \exp\{-n^{1-\gamma}h(1 + \delta_n)\}] \, O_p(1) = o_p(1).$$

It remains to consider the integral over $\boldsymbol{w} \in [0, 1]^d \setminus A_{n,\boldsymbol{u}}$, i.e., $|\boldsymbol{w} - \boldsymbol{u}|_\infty \leqslant \delta_n$ and $g(\boldsymbol{w}) \geqslant n^{-\gamma}(1 - \delta_n) > n^{-1}$, at least for sufficiently large n. By Lemma 4.1 in [3], we have, as $n \to \infty$,

$$\sup_{\substack{\boldsymbol{u},\boldsymbol{w} \in \{g \geqslant n^{-1}\} \\ |\boldsymbol{u}-\boldsymbol{w}|_\infty \leqslant \delta_n}} \left| \frac{\hat{\mathbb{C}}_n(\boldsymbol{w})}{g(\boldsymbol{w})^\omega} - \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})} \right| = o_p(1). \tag{14}$$

In view of Lemma 10, we obtain that, as $n \to \infty$,

$$\sup_{\boldsymbol{u} \in \{g \geqslant n^{-\gamma}\}} \int_{[0,1]^d \setminus A_{n,\boldsymbol{u}}} R_n(\boldsymbol{u}, \boldsymbol{w}) \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) \leqslant o_p(1) \int_{[0,1]^d} \frac{g(\boldsymbol{w})^\omega}{g(\boldsymbol{u})^\omega} \, \mathrm{d}\mu_{n,\boldsymbol{u}}(\boldsymbol{w}) = o_p(1).$$

The stated limit relation follows by combining the assertions on the integral over $A_{n,\boldsymbol{u}}$ and the one over its complement.

Note that in Lemma 4.1 in [3], the supremum in (14) is taken over $[1/n, 1 - 1/n]^d$ instead of over $\{g > n^{-1}\}$. But it can be seen in the proof of that statement that the result can be extended to the set $\{g \geqslant n^{-1}\}$. Furthermore, in the latter reference, the pseudo-observations are defined as $\hat{U}_{i,j} = R_{i,j}/(n + 1)$. However, this does not affect the above proof, since the difference of the two empirical copulas is at most d/n, almost surely. This gives an additional error term on the event $\{g \geqslant n^{-1}\}$ which is of the order $O_p(n^{\omega+1/2-1}) = o_p(1)$, as $n \to \infty$. □

6.3. On the expectation of the reciprocal of a binomial random variable

Lemma 12. *Let $u \in (0, 1)$ and $n \geqslant 2$ be integer. If $S \sim \mathrm{Bin}(n, u)$ and $T \sim \mathrm{Bin}(n - 1, u)$, then*

$$\mathrm{E}\left[\frac{1}{S} \, \mathbf{1}_{\{S \geqslant 1\}}\right] = nu \, \mathrm{E}\left[\frac{1}{(1 + T)^2}\right] = nu \int_0^1 (1 - u + us)^{n-1}(-\ln s) \, \mathrm{d}s. \tag{15}$$

Proof. For $k \in \{1, \ldots, n\}$, we have

$$\frac{\Pr(S = k)}{\Pr(T + 1 = k)} = \frac{\binom{n}{k}u^k(1 - u)^{n-k}}{\binom{n-1}{k-1}u^{k-1}(1 - u)^{(n-1)-(k-1)}} = \frac{nu}{k}.$$

We obtain that

$$\mathrm{E}[\mathbf{1}_{\{S \geqslant 1\}}/S] = \sum_{k=1}^n \frac{1}{k} \Pr(S = k) = \sum_{k=1}^n \frac{1}{k} \frac{nu}{k} \Pr(T + 1 = k) = (nu) \, \mathrm{E}\{1/(1 + T)^2\}.$$

Now we apply a trick due to Chao and Strawderman [10]. We have

$$\frac{1}{(1 + T)^2} = \int_{t=0}^1 \frac{1}{t} \int_{s=0}^t s^T \, \mathrm{d}s \, \mathrm{d}t = \int_{s=0}^1 s^T \int_{t=s}^1 \frac{\mathrm{d}t}{t} \, \mathrm{d}s = \int_0^1 s^T(-\ln s) \, \mathrm{d}s.$$

Taking expectations and using Fubini's theorem, we obtain

$$\mathrm{E}\{1/(1 + T)^2\} = \int_0^1 \mathrm{E}(s^T)\,(-\ln s) \, \mathrm{d}s = \int_0^1 (1 - u + us)^{n-1}(-\ln s) \, \mathrm{d}s,$$

as required. □

Lemma 13. *Let $u_n \in (0, 1]$ and $S_n \sim \mathrm{Bin}(n, u_n)$. If $nu_n \to \infty$, then, as $n \to \infty$, $(nu_n^2) \, \mathrm{E}[\mathbf{1}_{\{S_n \geqslant 1\}}/S_n] = u_n + O(1/n)$.*

15

Proof. We start from (15):

$$(nu_n^2)\,\mathrm{E}[\mathbf{1}_{\{S_n \geqslant 1\}}/S_n] = n^2 u_n^3 \int_0^1 (1 - u_n + u_n s)^{n-1}(-\ln s)\,\mathrm{d}s.$$

We split the integral in two parts, cutting at $s = 1/2$.

First we consider the case $s \leqslant 1/2$. For some positive constant K, we have

$$n^2 u_n^3 \int_0^{1/2} (1 - u_n + u_n s)^{n-1}(-\ln s)\,\mathrm{d}s \leqslant n^2 u_n^3 (1 - u_n/2)^{n-1} \int_0^{1/2} (-\ln s)\,\mathrm{d}s$$
$$\leqslant K\,n^2 u_n^3 (1 - u_n/2)^n \leqslant K\,n^2 u_n^3 \exp(-nu_n/2).$$

For any $m > 0$, this expression is $O\{u_n(nu_n)^{-m}\} = o(1/n)$ as $n \to \infty$, hence by choosing $m = 1$ it is $O(1/n)$ as $n \to \infty$.

Second we consider the case $s \geqslant 1/2$. The substitution $s = 1 - v/(nu_n)$ yields

$$n^2 u_n^3 \int_{1/2}^1 (1 - u_n + u_n s)^{n-1}(-\ln s)\,\mathrm{d}s = u_n \int_0^{(nu_n/2)} (1 - v/n)^{n-1}[-(nu_n)\ln\{1 - v/(nu_n)\}]\,\mathrm{d}v. \tag{16}$$

We need to show that this integral is $u_n + O(1/n)$ as $n \to \infty$.

For facility of writing, put $k_n = nu_n$. Recall that $k_n \to \infty$ as $n \to \infty$ by assumption. The inequalities $x \leqslant -\ln(1 - x) \leqslant x/(1 - x)$, valid for all $x \in [0, 1)$, imply that, for all $v \in [0, k_n/2]$,

$$0 \leqslant -k_n \ln(1 - v/k_n) - v \leqslant \frac{v^2}{k_n - v} \leqslant \frac{2v^2}{k_n}.$$

As $(1 - v/n)^{n-1} \leqslant \{1 - k_n/(2n)\}^{-1}(1 - v/n)^{-n} \leqslant 2\exp(-v)$ for $v \in [0, k_n/2]$, we find

$$u_n \int_0^{k_n/2} (1 - v/n)^{n-1} \,|-k_n \ln(1 - v/k_n) - v|\,\mathrm{d}v \leqslant \frac{4u_n}{k_n} \int_0^{k_n/2} \exp(-v)\,v^2\,\mathrm{d}v = O(1/n),$$

as $n \to \infty$. As a consequence, replacing $-k_n \ln(1 - v/k_n)$ by v in (16) produces an error of the required order $O(1/n)$.

It remains to consider the integral

$$u_n \int_0^{k_n/2} (1 - v/n)^{n-1}\,v\,\mathrm{d}v.$$

Via the substitution $x = 1 - v/n$, this integral can be computed explicitly. After some routine calculations, we find it is equal to

$$u_n \frac{n}{n + 1}[1 - \{1 - k_n/(2n)\}^n (1 + k_n/2)].$$

Since $\{1 - k_n/(2n)\}^n \leqslant \exp(-k_n/2)$, the previous expression is $u_n + O(u_n/n) + O\{u_n \exp(-k_n/2)k_n\}$, as $n \to \infty$. The error term is $O(1/n)$, as required. $\qquad\square$

Lemma 14. *If $0 < u_n \leqslant 1$ is such that $nu_n \to \infty$ as $n \to \infty$, then*

$$\sup_{u_n \leqslant u \leqslant 1} |nu^2\,\mathrm{E}[\mathbf{1}_{\{S \geqslant 1\}}/S] - u| = O(n^{-1}),$$

as $n \to \infty$, where the expectation is taken for $S \sim \mathrm{Bin}(n, u)$.

Proof. The function sending $u \in [u_n, 1]$ to $|nu^2\,\mathrm{E}[\mathbf{1}_{\{S \geqslant 1\}}/S] - 1|$, with $S \sim \mathrm{Bin}(n, u)$, is continuous and therefore attains its supremum at some $v_n \in [u_n, 1]$. Since $nv_n \geqslant nu_n \to \infty$ as $n \to \infty$, we can apply Lemma 13 to find that the supremum is $O(1/n)$ as $n \to \infty$. $\qquad\square$

16

6.4. Inequalities for binomial random variables

If $S \sim \mathrm{Bin}(n, u)$ is a binomial random variable with success probability $u \in (0, 1)$, then Bennett's inequality states that

$$\Pr\left(\sqrt{n}\left|\tfrac{S}{n} - u\right| \geqslant \lambda \right) \leqslant 2 \exp\left\{-\frac{\lambda^2}{2u}\psi\left(\frac{\lambda}{\sqrt{nu}}\right)\right\} = 2 \exp\left\{-nu\, h\left(1 + \frac{\lambda}{\sqrt{nu}}\right)\right\}$$

for $\lambda > 0$, where $\psi(x) = 2\, h(1 + x)/x^2$ and $h(x) = x(\ln x - 1) + 1$; see, e.g., Proposition A.6.2 in [36]. Setting $\lambda = \sqrt{n}\, u\delta$, we find that, for all $\delta > 0$,

$$\Pr(|S/n - u| \geqslant u\delta) \leqslant 2 \exp\{-nu\, h(1 + \delta)\}. \tag{17}$$

Note that $h(1 + \delta) = \int_0^\delta \ln(1 + t)\, \mathrm{d}t \geqslant \int_0^\delta (t - t^2/2)\, \mathrm{d}t = \delta^2(1 - \delta/3)/2$ for $\delta \geqslant 0$ and thus $h(1 + \delta) \geqslant \delta^2/3$ for all $\delta \in [0, 1]$. We extend (17) to a vector of independent binomial random variables and in terms of the weight function g in (3).

Lemma 15. *If $S_1, \ldots, S_d$ are independent random variables with $S_j \sim \mathrm{Bin}(n, u_j)$ and $u_j \in (0, 1)$ for all $j \in \{1, \ldots, d\}$, then, for $\delta > 0$,*

$$\Pr\{g(S_1/n, \ldots, S_d/n) \geqslant g(\boldsymbol{u})(1 + \delta)\} \leqslant 2d \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}, \tag{18}$$

$$\Pr\{g(S_1/n, \ldots, S_d/n) \leqslant g(\boldsymbol{u})(1 - \delta)\} \leqslant 4d \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}, \tag{19}$$

with h as above; in particular, $h(1 + \delta) \geqslant \delta^2/3$ for all $\delta \in (0, 1]$.

Proof. Let us start with (19). The definition of the weight function g in (3) yields

$$\Pr\{g(S_1/n, \ldots, S_d/n) \leqslant g(\boldsymbol{u})(1 - \delta)\} \leqslant \sum_{j=1}^d \Big[\Pr\{S_j/n \leqslant g(\boldsymbol{u})(1 - \delta)\} + \Pr\Big\{\max_{k \neq j}(1 - S_k/n) \leqslant g(\boldsymbol{u})(1 - \delta)\Big\}\Big].$$

Let us first consider the first term on the right-hand side, i.e., $\Pr\{S_j/n \leqslant g(\boldsymbol{u})(1 - \delta)\}$. By definition of the weight function we have $g(\boldsymbol{u}) \leqslant u_j$. By Bennett's inequality (17),

$$\Pr\{S_j/n \leqslant g(\boldsymbol{u})(1 - \delta)\} \leqslant \Pr\{S_j/n \leqslant u_j(1 - \delta)\} \leqslant \Pr\{|S_j/n - u_j| \geqslant u_j\delta\}$$
$$\leqslant 2 \exp\{-nu_j h(1 + \delta)\} \leqslant 2 \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}.$$

Second, consider the term $\Pr\{\max_{k \neq j}(1 - S_k/n) \leqslant g(\boldsymbol{u})(1 - \delta)\}$. Suppose $j = 1$; the other cases can be treated exactly along the same lines. We have $g(\boldsymbol{u}) \leqslant \max_{k \neq 1}(1 - u_k)$. Assume without loss of generality that $\max_{k \neq 1}(1 - u_k) = 1 - u_2$. Then we obtain $g(\boldsymbol{u}) \leqslant 1 - u_2$ and, by Bennett's inequality (17) applied to $n - S_2 \sim \mathrm{Bin}(n, 1 - u_2)$,

$$\Pr\Big\{\max_{k \neq 1}(1 - S_k/n) \leqslant g(\boldsymbol{u})(1 - \delta)\Big\} \leqslant \Pr\Big\{\max_{k \neq 1}(1 - S_k/n) \leqslant (1 - u_2)(1 - \delta)\Big\}$$
$$\leqslant \Pr\{1 - S_2/n \leqslant (1 - u_2)(1 - \delta)\}$$
$$\leqslant \Pr\{|1 - S_2/n - (1 - u_2)| \geqslant (1 - u_2)\delta\}$$
$$\leqslant 2 \exp\{-n(1 - u_2)h(1 + \delta)\}$$
$$\leqslant 2 \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}.$$

Let us now show (18). First suppose $g(\boldsymbol{u}) = u_1$. Since $g(S_1/n, \ldots, S_d/n) \leqslant S_1/n$ we have, by Bennett's inequality (17),

$$\Pr\{g(S_1/n, \ldots, S_d/n) \geqslant g(\boldsymbol{u})(1 + \delta)\} \leqslant \Pr\{S_1/n \geqslant u_1(1 + \delta)\} \leqslant \Pr\{|S_1/n - u_1| \geqslant u_1\delta\}$$
$$\leqslant 2 \exp\{-nu_1 h(1 + \delta)\} = 2 \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}.$$

Finally, suppose that $g(\boldsymbol{u}) = 1 - u_1 \geqslant 1 - u_k$, for $k = 3, \ldots, d$. Note that $g(S_1/n, \ldots, S_d/n) \leqslant \max_{k \neq 2}(1 - S_j/n)$, which yields

$$\Pr\{g(S_1/n, \ldots, S_d/n) \geqslant g(\boldsymbol{u})(1 + \delta)\} \leqslant \Pr\Big\{\max_{k \neq 2}(1 - S_k/n) \geqslant (1 - u_1)(1 + \delta)\Big\} \leqslant \sum_{k \neq 2} \Pr\Big\{1 - S_k/n \geqslant (1 - u_1)(1 + \delta)\Big\}.$$

17

By Bennett's inequality (17) applied to $n - S_k \sim \text{Bin}(n, 1 - u_k)$ for every $k \neq 2$, we have, since $(1 - u_1)/(1 - u_k) \geqslant 1$,

$$\Pr\{1 - S_k/n \geqslant (1 - u_1)(1 + \delta)\} \leqslant \Pr\{|1 - S_k/n - (1 - u_k)| \geqslant (1 - u_1)(1 + \delta) - (1 - u_k)\}$$
$$\leqslant 2 \exp\left[-n(1 - u_k)h\left\{\tfrac{1-u_1}{1-u_k}(1 + \delta)\right\}\right].$$

For $a \geqslant 1$ and $\delta \geqslant 0$, a direct calculation shows that $h\{a(1+\delta)\} - a\,h(1+\delta) \geqslant h(a) \geqslant 0$ and thus $h\{a(1+\delta)\} \geqslant a\,h(1+\delta)$. Apply this inequality to $a = (1 - u_1)/(1 - u_k)$ to find

$$\Pr\{1 - S_k/n \geqslant (1 - u_1)(1 + \delta)\} \leqslant 2 \exp\{-n(1 - u_k)\tfrac{1-u_1}{1-u_k}h(1 + \delta)\}$$
$$= 2 \exp\{-n(1 - u_1)h(1 + \delta)\} = 2 \exp\{-ng(\boldsymbol{u})h(1 + \delta)\}. \qquad \square$$

6.5. Extensions of results in [3]

For any sequence $\delta_n > 0$ that converges to zero as $n \to \infty$, Lemma 4.10 in [3] can be extended to

$$\sup\left\{\left|\frac{\mathbb{C}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} - \frac{\mathbb{C}_n(\boldsymbol{u}')}{g(\boldsymbol{u}')^\omega}\right| \;:\; g(\boldsymbol{u}) \geqslant c/n,\, g(\boldsymbol{u}') \geqslant c/n,\, |\boldsymbol{u} - \boldsymbol{u}'| \leqslant \delta_n\right\} = o_p(1), \tag{20}$$

as $n \to \infty$. Here, $\mathbb{C}_n = \sqrt{n}\,(\tilde{C}_n - C)$ and $\tilde{C}_n$ is the empirical copula based on the generalized inverse function of the marginal empirical distribution functions; see beginning of Section 4.2 in [3]. Furthermore, Theorem 4.5 in the same reference can be extended to the conclusion that, as $n \to \infty$,

$$\sup\left\{\left|\frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g(\boldsymbol{u})^\omega}\right| \;:\; g(\boldsymbol{u}) \geqslant c/n\right\} = o_p(1). \tag{21}$$

Proof. Let us start with (20). The result is similar to the result in Lemma 4.10, in particular Eq. (4.1), in [3]. A look at the proof of the result shows that the restriction $\boldsymbol{u}, \boldsymbol{u}' \in [c/n, 1 - c/n]^d$ instead of $\boldsymbol{u}, \boldsymbol{u}' \in \{g \geqslant c/n\}$ is not needed. The proof of Eq. (4.1) in Lemma 4.10 in [3] is based on Lemma 4.7, 4.8 and Eqs. (4.8) and (4.9) which are all valid on sets of the form $N(c_{n1}, c_{n2}) = \{g \in (c_{n1}, c_{n2}]\}$. Hence, in the proof, all suprema can be taken over $\boldsymbol{u}, \boldsymbol{u}' \in \{g \geqslant c/n\}$ instead of $\boldsymbol{u}, \boldsymbol{u}' \in [c/n, 1 - c/n]^d$, which gives us exactly (20).

For the proof of (21) note that for any $\boldsymbol{u} \in \{g \geqslant c/n\}$ we can find $\boldsymbol{u}' \in \{g \geqslant n^{-1/2}\}$ such that $|\boldsymbol{u} - \boldsymbol{u}'| \leqslant dn^{-1/2}$. To find such a $\boldsymbol{u}'$ is all that it is needed to extend the proof of Theorem 4.5 in [3] to obtain (21). $\qquad \square$

Acknowledgments

References

References

[1] M. Belalia, T. Bouezmarni, F.C. Lemyre, A. Taamouti, Testing independence based on Bernstein empirical copula and copula density, J. Nonparam. Statist. 29 (2017) 346–380.

[2] B. Berghaus, A. Bücher, H. Dette, Minimum distance estimators of the Pickands dependence function and related tests of multivariate extreme-value dependence, J. Soc. Franc. Stat. 154 (2013) 116–137.

[3] B. Berghaus, A. Bücher, S. Volgushev, Weak convergence of the empirical copula process with respect to weighted metrics, Bernoulli 23 (2017) 743–772.

[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] T. Bouezmarni, J. Rombouts, A. Taamouti, Asymptotic properties of the Bernstein density copula estimator for α-mixing data, J. Multivariate Anal. 101 (2010) 1–10.

[6] A. Bücher, H. Dette, S. Volgushev, New estimators of the Pickands dependence function and a test for extreme-value dependence, Ann. Statist. 39 (2011) 1963–2006.

[7] A. Bücher, J. Segers, Extreme value copula estimation based on block maxima of a multivariate stationary time series, Extremes 17 (2014) 495–528.

[8] A. Bücher, S. Volgushev, Empirical and sequential empirical copula processes under serial dependence, J. Multivariate Anal. 119 (2013) 61–70.

[9] P. Capéraà, A.-L. Fougères, C. Genest, A nonparametric estimation procedure for bivariate extreme value copulas, Biometrika 84 (1997) 567–577.

[10] M.T. Chao, W. E. Strawderman, Negative moments of positive random variables, J. Amer. Statist. Assoc. 67 (1972) 429–431.

[11] T. De Wet, Cramér–von Mises tests for independence, J. Multivariate Anal. 10 (1980) 38–50.

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

[13] P. Deheuvels, An asymptotic decomposition for multivariate distribution-free tests of independence, J. Multivariate Anal. 11 (1981) 102–113.

[14] P. Deheuvels, A nonparametric test of independence, Publ. Inst. Statist. Univ. Paris 26 (1981) 29–50.

[15] Y. Fan, P. Lafaye de Micheaux, S. Penev, D. Salopek, Multivariate nonparametric test of independence, J. Multivariate Anal. 153 (2017) 189–210.

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

[17] P. Gänßler, W. Stute, Seminar on Empirical Processes, Birkhäuser Verlag, Basel, 1987.

[18] C. Genest, J.G. Nešlehová, B. Rémillard, Asymptotic behavior of the empirical multilinear copula process under broad conditions, J. Multivariate Anal. 159 (2017) 82–110.

[19] C. Genest, J.-F. Quessy, B. Rémillard, Local efficiency of a Cramér–von Mises test of independence, J. Multivariate Anal. 97 (2006) 274–294.

[20] C. Genest, B. Rémillard, Tests of independence and randomness based on the empirical copula process, Test 13 (2004) 335–369.

[21] C. Genest, J. Segers, Rank-based inference for bivariate extreme-value copulas, Ann. Statist. 37 (2009) 2990–3022.

[22] G. Gudendorf, J. Segers, Extreme-value copulas, In: Copula Theory and its Applications, pp. 127–145, Springer, Heidelberg, 2010.

[23] G. Gudendorf, J. Segers, Nonparametric estimation of multivariate extreme-value copulas J. Stat. Plann. Inf. 142 (2012) 3073–3085.

[24] E. Gumbel, Bivariate logistic distributions, J. Amer. Statist. Assoc. 56 (1961) 335–349.

[25] W. Hoeffding, A non-parametric test of independence, Ann. Math. Statist. 19 (1948) 546–557.

[26] P. Janssen, J. Swanepoel, N. Veraverbeke, Large sample behavior of the Bernstein copula estimator, J. Stat. Plann. Inf. 142 (2012) 1189–1197.

[27] I. Kojadinovic, M. Holmes, Tests of independence among continuous random vectors based on Cramér–von Mises functionals of the empirical copula process, J. Multivariate Anal. 100 (2009) 1137–1154.

[28] I. Kojadinovic, J. Yan, Modeling multivariate distributions with continuous margins using the copula R package, J. Stat. Software 34 (2010) 1–20.

[29] I. Pickands III, Multivariate extreme value distributions, In: Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981), Volume 49, pp. 859–878, 894–902, 1981.

[30] R Core Team R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017.

[31] L. Rüschendorf, Asymptotic distributions of multivariate rank order statistics, Ann. Statist. 4 (19976) 912–923.

[32] A. Sancetta, S. Satchell, The Bernstein copula and its applications to modeling and approximations of multivariate distributions, Econom. Theory 20 (2004) 535–562.

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

[34] J. Segers, M. Sibuya, H. Tsukahara, The empirical beta copula, J. Multivariate Anal. 155 (2017) 35–51.

[35] J.A. Tawn, Modelling multivariate extreme value distributions, Biometrika 77 (1990) 245–253.

[36] A.W. van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes, Springer, New York, 1996.