

## DISCUSSION PAPER

2017/15

---

Weak convergence of  
the weighted empirical beta copula process

---

Berghaus, B. and J. Segers

# Weak convergence of the weighted empirical beta copula process

Betina Berghaus<sup>a</sup>, Johan Segers<sup>b</sup>

<sup>a</sup>*Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstr. 150, 44780 Bochum, Germany*

<sup>b</sup>*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 the estimation of the Pickands dependence function of an extreme-value copula.

KEY WORDS: Copula, Empirical beta copula, Empirical copula, weighted weak convergence, Pickands dependence function

---

## 1. Introduction

In many statistical questions related to multivariate dependence, a crucial role is played by the copula function. A basic nonparametric copula estimator is the empirical copula, which dates back to [Rüschendorf \(1976\)](#) and [Deheuvels \(1979\)](#) and which is 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., [Gaenssler and Stute, 1987](#); [Fermanian et al., 2004](#); [Segers, 2012](#); [Bücher and Volgushev, 2013](#)). 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. \(2017\)](#) 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

---

*Email addresses:* [betina.berghaus@rub.de](mailto:betina.berghaus@rub.de) (Betina Berghaus), [johan.segers@uclouvain.be](mailto:johan.segers@uclouvain.be) (Johan Segers)

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 (Berghaus et al., 2017, Remark 2.3).

The empirical beta copula (Segers et al., 2017) arises as a particular case of the empirical Bernstein copula (see, e.g., Sancetta and Satchell, 2004; Janssen et al., 2012) if the degrees of the Bernstein polynomials are set to the sample size. In numerical experiments, 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. For this reason it is possible to prove weighted weak convergence for the empirical beta copula process on the whole unit cube. In turn, weak convergence on the whole unit cube rather than on a subset thereof comes in a lot more 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.

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. In Section 3 we demonstrate the ease of application of the main result to the analysis of an empirical beta-copula based variant of the Capéràa–Fougères–Genest estimator (Capéràa et al., 1997) of the Pickands dependence function of a bivariate extreme-value copula. A simulation study shows that, under weak dependence, the new variant performs better than the original estimator. The proofs are deferred to Section 4, whereas a number of detailed arguments are worked out in Section 5.

## 2. Notation and main result

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

$$\mathbb{P}(X_{i,j} \leq x_j) = F_j(x_j), \quad \mathbb{P}(\mathbf{X}_i \leq \mathbf{x}) = F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}.$$

For vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ , the inequality  $\mathbf{x} \leq \mathbf{y}$  means that  $x_j \leq y_j$  for  $j = 1, \dots, d$ . Similar conventions apply for other inequalities and for minima and maxima, denoted by the operators  $\wedge$  and  $\vee$ , respectively. Given the sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$ , 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 too. For instance, if  $(Y_n)_n$  is a univariate, strictly stationary time series, then the  $d$ -variate time series of lagged values  $\mathbf{X}_n = (Y_n, Y_{n-1}, \dots, Y_{n-d+1})$  is strictly stationary too and the instantaneous dependence within the series  $(\mathbf{X}_n)_n$  corresponds to serial dependence within the original series  $(Y_n)_n$  up to lag  $d - 1$ .

For  $i = 1, \dots, n$  and  $j = 1, \dots, d$ , let  $R_{i,j}$  denote the rank of  $X_{i,j}$  among  $X_{1,j}, \dots, X_{n,j}$ . For convenience, we omit the sample size  $n$  in the notation for ranks. The random vectors  $\hat{\mathbf{U}}_i = (\hat{U}_{i,1}, \dots, \hat{U}_{i,d})$ , with  $\hat{U}_{i,j} = n^{-1}R_{i,j}$  and  $i = 1, \dots, n$ , are called pseudo-observations from  $C$ . The empirical copula is defined by

$$\hat{C}_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\hat{\mathbf{U}}_i \leq \mathbf{u}\}}, \quad \mathbf{u} \in [0, 1]^d,$$

where  $\mathbb{1}_A$  denotes the indicator variable of the event  $A$ .

Under mixing conditions on the sequence  $(\mathbf{X}_n)_n$  and smoothness conditions on  $C$ , it was shown in [Bücher and Volgushev \(2013\)](#) that

$$\hat{\mathbb{C}}_n = \sqrt{n}(\hat{C}_n - C) \rightsquigarrow \mathbb{C}_C, \quad n \rightarrow \infty \quad (2.1)$$

in the metric space  $\ell^\infty([0, 1]^d) = \{f : [0, 1]^d \rightarrow \mathbb{R} \mid \sup_{\mathbf{u} \in [0, 1]^d} |f(\mathbf{u})| < \infty\}$  equipped with the supremum distance. The arrow  $\rightsquigarrow$  in (2.1) denotes weak convergence in metric spaces as exposed in [van der Vaart and Wellner \(1996\)](#). The limit process in (2.1) is

$$\mathbb{C}_C(\mathbf{u}) = \alpha_C(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \alpha(1, \dots, 1, u_j, 1, \dots, 1), \quad \mathbf{u} \in [0, 1]^d,$$

where  $\dot{C}_j(\mathbf{u}) = \partial C(\mathbf{u}) / \partial u_j$  and where  $\alpha_C$  is a tight, centered Gaussian process on  $[0, 1]^d$  with covariance function

$$\text{Cov}(\alpha_C(\mathbf{u}), \alpha_C(\mathbf{v})) = \sum_{i \in \mathbb{Z}} \text{Cov}(\mathbb{1}_{\{\mathbf{U}_0 \leq \mathbf{u}\}}, \mathbb{1}_{\{\mathbf{U}_i \leq \mathbf{v}\}}), \quad \mathbf{u}, \mathbf{v} \in [0, 1]^d, \quad (2.2)$$

where  $\mathbf{U}_i = (U_{i,1}, \dots, U_{i,d})$  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  $\mathbf{U}_i$  is  $C$ . The margins  $F_1, \dots, F_d$  being unknown, we cannot observe the  $\mathbf{U}_i$ , and this is why we use the  $\hat{\mathbf{U}}_i$  instead. In the case of serial independence, weak convergence of  $\hat{\mathbb{C}}_n$  has been investigated by many authors, see the references in [Bücher and Volgushev \(2013\)](#); the series in (2.2) simplifies to  $\text{Cov}(\mathbb{1}_{\{\mathbf{U}_0 \leq \mathbf{u}\}}, \mathbb{1}_{\{\mathbf{U}_0 \leq \mathbf{v}\}}) = C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})C(\mathbf{v})$  so that  $\alpha_C$  is a  $C$ -Brownian bridge. In the stationary case, convergence of the series in (2.2) is a consequence of the mixing conditions imposed on  $(\mathbf{X}_n)_n$ .

Weak convergence in (2.1) 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 in [Berghaus et al. \(2017\)](#), 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 with

$$g(\mathbf{u}) = \bigwedge_{j=1}^d \left\{ u_j \wedge \bigvee_{k \neq j} (1 - u_k) \right\}, \quad \mathbf{u} \in [0, 1]^d. \quad (2.3)$$

Note that  $g(\mathbf{u})$  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  $\{\mathbf{v} \in [0, 1]^d : g(\mathbf{v}) \geq c/n\}$ . Relying on such a workaround, Theorem 2.2 in [Berghaus et al. \(2017\)](#) states weak convergence of the weighted empirical copula process  $\hat{\mathbb{C}}_n/g^\omega$  to  $\mathbb{C}_C/g^\omega$ . Note that  $g(\mathbf{v}) = 0$  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  $\mathbb{C}_C(\mathbf{v}) = 0$  almost surely for such  $\mathbf{v}$  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. In [Segers et al. \(2017\)](#), the empirical beta copula is defined as

$$C_n^\beta(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d F_{n, R_{i,j}}(u_j), \quad \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d, \quad (2.4)$$

where  $F_{n,r}$  is the distribution function of the beta distribution  $\mathcal{B}(r, n+1-r)$ , i.e.,

$$F_{n,r}(u) = \sum_{s=r}^n \binom{n}{s} u^s (1-u)^{n-s}, \quad u \in [0, 1], \quad r \in \{1, \dots, n\}.$$

Note that

$$C_n^\beta(\mathbf{u}) = \int_{[0,1]^d} \hat{C}_n(\mathbf{w}) \, d\mu_{n,\mathbf{u}}(\mathbf{w}), \quad (2.5)$$

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

Consider the empirical beta copula process

$$\mathbb{C}_n^\beta = \sqrt{n}(C_n^\beta - C).$$

Under a smoothness condition on  $C$ , it follows from Theorem 3.6(ii) in [Segers et al. \(2017\)](#) that weak convergence in  $\ell^\infty([0, 1]^d)$  of the empirical copula process  $\hat{\mathbb{C}}_n$  to a limit process  $\mathbb{C}$  with continuous trajectories is sufficient to conclude that also

$$\mathbb{C}_n^\beta = \hat{\mathbb{C}}_n + o_{\mathbb{P}}(1) \rightsquigarrow \mathbb{C}, \quad n \rightarrow \infty, \quad (2.6)$$

in the space  $\ell^\infty([0, 1]^d)$ . In quite general circumstances, 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 Segers et al. (2017) revealed the empirical beta copula to be more accurate.

Our aim is to extend the convergence statement in (2.6) for weighted versions  $\mathbb{C}_n^\beta/g^\omega$ . 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 5.1 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 Berghaus et al. (2017).

The analysis of  $\mathbb{C}_n^\beta/g^\omega$  will be based on the one of  $\hat{\mathbb{C}}_n/g^\omega$  via (2.5). We will therefore need the same smoothness condition on  $C$  as imposed in Berghaus et al. (2017, Condition 2.1), combining Conditions 2.1 and 4.1 in Segers (2012). The condition is satisfied by many copula families; see the examples in the latter article.

**Condition 2.1.** (i) For every  $j \in \{1, \dots, d\}$ , the first-order partial derivative  $\dot{C}_j(\mathbf{u}) := \partial C(\mathbf{u})/\partial u_j$  exists and is continuous on  $V_j = \{\mathbf{u} \in [0, 1]^d : u_j \in (0, 1)\}$ .

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

$$|\ddot{C}_{j_1 j_2}(\mathbf{u})| \leq K \min \left\{ \frac{1}{u_{j_1}(1 - u_{j_1})}, \frac{1}{u_{j_2}(1 - u_{j_2})} \right\}, \quad \forall \mathbf{u} \in V_{j_1} \cap V_{j_2}. \quad (2.7)$$

The alpha-mixing coefficients of the sequence  $(\mathbf{X}_n)_n$  are defined as

$$\alpha(k) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \sigma(\mathbf{X}_j, j \leq i), B \in \sigma(\mathbf{X}_{j+k}, j \geq i), i \in \mathbb{Z}\},$$

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

**Theorem 2.2.** *Suppose that  $\mathbf{X}_1, \mathbf{X}_2, \dots$  is a strictly stationary, alpha-mixing sequence with  $\alpha(k) = O(a^k)$ , as  $k \rightarrow \infty$ , for some  $a \in (0, 1)$ . Assume that within each variable, ties do not occur with probability one. If the copula  $C$  satisfies Condition 2.1, then, for any  $\omega \in (0, 1/2)$ , we have, in  $\ell^\infty([0, 1]^d)$ ,*

$$\mathbb{C}_n^\beta/g^\omega \rightsquigarrow \mathbb{C}_C/g^\omega, \quad n \rightarrow \infty.$$

*Remark 2.3.* 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 2.4.* The result also holds under weaker assumptions on the serial dependence. In [Berghaus et al. \(2017\)](#) 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. However, in this case the possible choices of  $\omega$  are more restricted. For more details see Theorem 4.5 in the latter reference.

### 3. 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

$$C(\mathbf{u}) = \exp \left\{ \left( \sum_{j=1}^d \ln u_j \right) A \left( \frac{\ln u_1}{\sum_{j=1}^d \ln u_j}, \dots, \frac{\ln u_{d-1}}{\sum_{j=1}^d \ln u_j} \right) \right\}, \quad \mathbf{u} \in (0, 1]^d \setminus \{\mathbf{1}\},$$

for some function  $A : \Delta_{d-1} \rightarrow [1/d, 1]$  called Pickands dependence function (after [Pickands, 1981](#)) and where  $\Delta_{d-1} = \{\mathbf{t} = (t_1, \dots, t_{d-1}) \in [0, 1]^{d-1} : \sum_{j=1}^{d-1} t_j \leq 1\}$  denotes the unit simplex.

The rank-based Capéràa–Fougères–Genest (CFG) estimator ([Capéràa et al., 1997](#)) of the Pickands dependence function of a bivariate extreme-value copula is given by ([Genest and Segers, 2009](#))

$$\ln\{\hat{A}_n^{\text{CFG}}(t)\} = -\gamma + \int_0^1 \{\hat{C}_n(u^{1-t}, u^t) - \mathbb{1}_{[e^{-1}, 1]}(u)\} \frac{du}{u \ln u}, \quad t \in [0, 1].$$

Here,  $\gamma$  is the Euler–Mascheroni constant and  $\hat{C}_n$  is the empirical copula defined as the bivariate cumulative distribution function of the pairs of pseudo-observations, the ranks being divided by  $n + 1$ . Replacing the empirical copula by the empirical beta copula in (2.4) yields the estimator

$$\ln\{\hat{A}_{n,\beta}^{\text{CFG}}(t)\} = -\gamma + \int_0^1 \{C_n^\beta(u^{1-t}, u^t) - \mathbb{1}_{[e^{-1}, 1]}(u)\} \frac{du}{u \ln u}. \quad (3.1)$$

The technique could also be used for other estimators based upon the empirical copula ([Bücher et al., 2011](#); [Berghaus et al., 2013](#)).

Thanks to Theorem 2.2, the limit of the latter estimator can be derived by a straightforward application of the continuous mapping theorem. The result does not require serial independence and can be extended to higher dimensions.

**Corollary 3.1.** *Let  $C$  be a bivariate extreme-value copula with Pickands dependence function  $A$ . Under the assumptions of Theorem 2.2 we have, as  $n \rightarrow \infty$ ,*

$$\sqrt{n}\{\hat{A}_{n,\beta}^{\text{CFG}}(\cdot) - A(\cdot)\} \rightsquigarrow \mathbb{A}(\cdot) \text{ in } \ell^\infty([0, 1]),$$

where for  $t \in [0, 1]$  we define  $\mathbb{A}(t) = A(t) \int_0^1 \mathbb{C}(u^{1-t}, u^t) \frac{du}{u \ln u}$ .

*Proof.* Let  $0 < \omega < 1/2$ . From the definition of  $\hat{A}_{n,\beta}^{\text{CFG}}$  in (3.1) and a similar formula linking  $A$  and  $C$  we obtain that

$$\begin{aligned}\sqrt{n}[\ln\{\hat{A}_{n,\beta}^{\text{CFG}}(t)\} - \ln\{A(t)\}] &= \int_0^1 \mathbb{C}_n^\beta(u^{1-t}, u^t) \frac{du}{u \ln u} \\ &= \int_0^1 (\mathbb{C}_n^\beta/g^\omega)(u^{1-t}, u^t) g^\omega(u^{1-t}, u^t) \frac{du}{u \ln u}.\end{aligned}$$

The integral  $\int_0^1 g^\omega(u^{1-t}, u^t) \frac{du}{u \ln u}$  is bounded, uniformly in  $t \in [0, 1]$ . Therefore, the linear map that sends a measurable function  $f \in \ell^\infty([0, 1]^2)$  to the bounded function  $[0, 1] \ni t \mapsto \int_0^1 f(u^{1-t}, u^t) g^\omega(u^{1-t}, u^t) \frac{du}{u \ln u}$  is continuous. By Theorem 2.2 and the continuous mapping theorem, we find, as  $n \rightarrow \infty$ ,

$$\sqrt{n}[\ln\{\hat{A}_{n,\beta}^{\text{CFG}}(t)\} - \ln\{A(t)\}] \rightsquigarrow \int_0^1 \mathbb{C}(u^{1-t}, u^t) \frac{du}{u \ln u} \text{ in } \ell^\infty([0, 1]).$$

Finally, the result follows by an application of the functional delta method.  $\square$

For the CFG-estimator we actually use the endpoint-corrected version

$$\ln\{\hat{A}_{n,c}^{\text{CFG}}(t)\} = \ln\{\hat{A}_n^{\text{CFG}}(t)\} - (1-t)\ln\{\hat{A}_n^{\text{CFG}}(0)\} - t\ln\{\hat{A}_n^{\text{CFG}}(1)\}. \quad (3.2)$$

For the estimator based on the empirical beta copula the endpoint correction is immaterial, since  $C_n^\beta$  is a copula itself: we already have  $C_n^\beta(u, 1) = C_n^\beta(1, u) = u$  for all  $u \in [0, 1]$  and thus  $\hat{A}_{n,\beta}^{\text{CFG}}(0) = \hat{A}_{n,\beta}^{\text{CFG}}(1) = 1$ .

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

$$\int_0^1 \mathbb{E}[\{\hat{A}(t) - A(t)\}^2] dt = \mathbb{E}[\{\hat{A}(T) - A(T)\}^2],$$

where the random variable  $T$  is uniformly distributed on  $(0, 1)$  and 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

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

where  $\hat{A}_n^{(m)}$  denotes the estimator based upon sample number  $m$ , and where the random variables  $T^{(1)}, \dots, T^{(m)}$  are uniformly distributed on  $(0, 1)$  and are independent of each other and of the copula samples. The approximation error is  $O_{\mathbb{P}}(1/\sqrt{M})$ , aggregating both the sampling error and the integration error. A similar trick was used in Segers et al. (2017) and is more efficient than first estimating the pointwise mean squared error through a Monte Carlo procedure and then integrating this out via numerical integration.



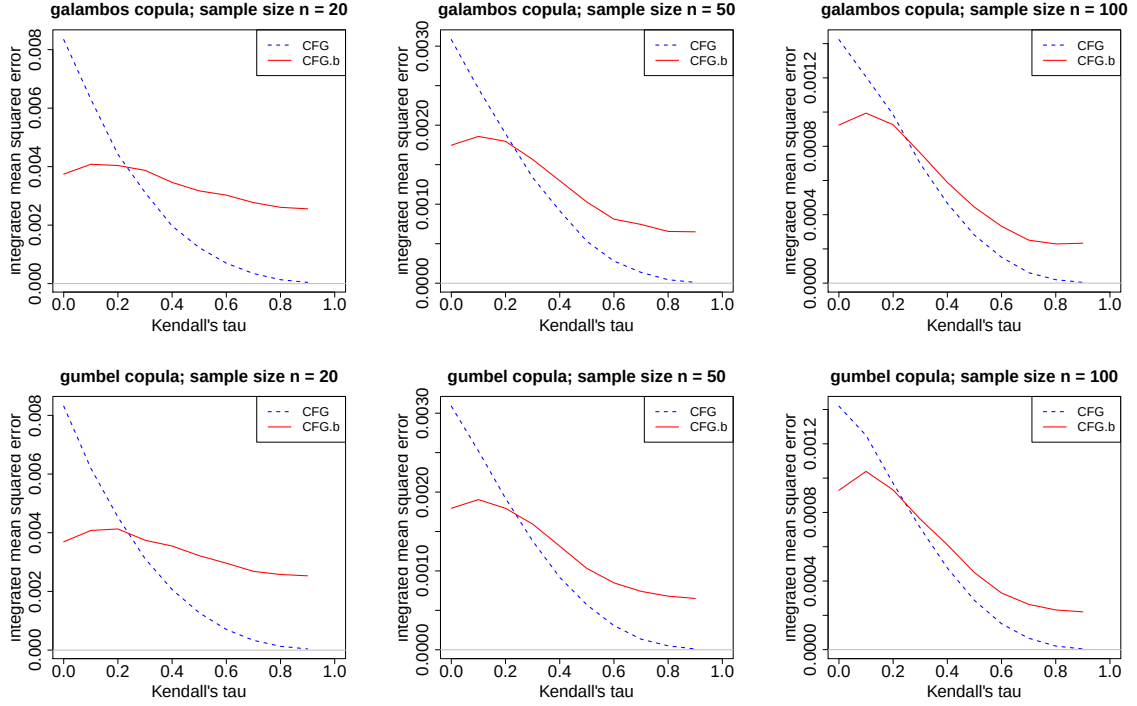


Figure 1: Integrated mean squared error (vertical axis) of the endpoint-corrected CFG-estimator (3.2) and the empirical beta variant (3.1) for the Pickands dependence function of the Galambos (top) and Gumbel (bottom) copulas as a function of Kendall's  $\tau \in \{0, 0.1, \dots, 0.9\}$  (horizontal axis), each point being based on 10 000 random samples of size  $n \in \{20, 50, 100\}$  (left to right).

The results are shown in Figure 1 for the Galambos and Gumbel copula families for  $M = 10\,000$  samples of size  $n \in \{20, 50, 100\}$ . The experiments were also done for the Hüsler–Reiss and t-EV copula families, yielding similar results, not shown to save space (see, e.g., [Gudendorf and Segers, 2010](#), for the definitions of these copulas). The copula families are parametrized via the value of Kendall's  $\tau$ . The calculations were done in the statistical software environment R ([R Core Team, 2017](#)) using the package `copula` ([Kojadinovic and Yan, 2010](#)). For weak dependence (small  $\tau$ ), the beta variant is the more efficient one, whereas for strong dependence (larger  $\tau$ ), it is the usual CFG estimator which is more accurate.

In order to gain a better understanding, we have also traced some trajectories of estimated Pickands dependence functions for the Gumbel copula. See Figure 2 for the results of  $\tau \in \{0.3, 0.9\}$  and  $n \in \{20, 50, 100\}$ . For large  $\tau$ , the true copula and its Pickands dependence function are strongly curved, and the empirical beta copula suffers from a bias due to oversmoothing. For weaker dependence, it is the CFG estimator which seems to be a bit biased.

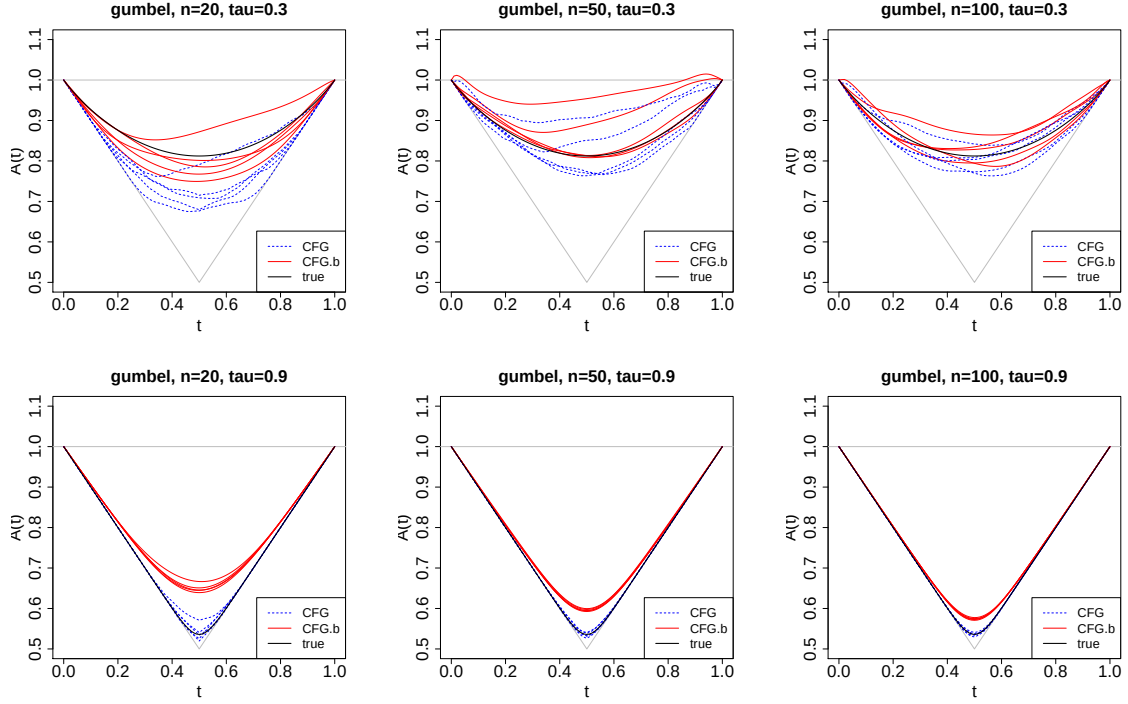


Figure 2: Plots of estimated Pickands dependence functions by the CFG-estimator (3.2) and the empirical beta variant (3.1) for the Gumbel copula with Kendall's  $\tau = 0.3$  (top) and  $\tau = 0.9$  (bottom) and for  $n \in \{20, 50, 100\}$  (left to right).

#### 4. Proof of Theorem 2.2

Recall the empirical copula process  $\hat{C}_n = \sqrt{n}(\hat{C}_n - C)$  and the empirical beta copula process  $\hat{C}_n^\beta = \sqrt{n}(C_n^\beta - C)$ . The link between the empirical copula  $\hat{C}_n$  and the empirical beta copula  $C_n^\beta$  is given in (2.5). In the derivation of the limit of the weighted empirical beta copula process the following decomposition plays a central role:

$$\begin{aligned} \frac{C_n^\beta(\mathbf{u})}{g(\mathbf{u})^\omega} &= \frac{\hat{C}_n(\mathbf{u})}{g(\mathbf{u})^\omega} \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) \\ &\quad + \int_{[0,1]^d} \left\{ \frac{\hat{C}_n(\mathbf{w})}{g(\mathbf{w})^\omega} - \frac{\hat{C}_n(\mathbf{u})}{g(\mathbf{u})^\omega} \right\} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) \\ &\quad + \int_{[0,1]^d} \sqrt{n} \frac{C(\mathbf{w}) - C(\mathbf{u})}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}). \end{aligned} \quad (4.1)$$

It is reasonable to assume that the last two terms on the right-hand side vanish as  $n \rightarrow \infty$ . Indeed, the measure  $\mu_{n,\mathbf{u}}$  concentrates around its mean  $\mathbf{u}$ , if the sample size grows, and both integrands are small if  $\mathbf{w}$  is close to  $\mathbf{u}$ . By the same reason, the integral

in the first term should be close to one. 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.

Recall that  $0 < \omega < 1/2$ . Fix a scalar  $\gamma$  such that  $1/\{2(1-\omega)\} < \gamma < 1$ . Consider the abbreviations  $\{g \geq n^{-\gamma}\} = \{\mathbf{v} \in [0, 1]^d \mid g(\mathbf{v}) \geq n^{-\gamma}\}$  and similarly  $\{g < n^{-\gamma}\}$ . By Lemma 5.1, we have

$$\begin{aligned}\mathbb{C}_n^\beta/g^\omega &= \mathbb{C}_n^\beta/g^\omega \mathbb{1}_{\{g \geq n^{-\gamma}\}} + \mathbb{C}_n^\beta/g^\omega \mathbb{1}_{\{g < n^{-\gamma}\}} \\ &= \mathbb{C}_n^\beta/g^\omega \mathbb{1}_{\{g \geq n^{-\gamma}\}} + o(1), \quad n \rightarrow \infty, \quad \text{a.s.}\end{aligned}$$

The three terms on the right-hand side of (4.1) are treated in Lemmas 5.2, 5.3 and 5.4. We find

$$\mathbb{C}_n^\beta/g^\omega = \hat{\mathbb{C}}_n/g^\omega \mathbb{1}_{\{g \geq n^{-\gamma}\}}(1 + o(1)) + o_{\mathbb{P}}(1), \quad n \rightarrow \infty. \quad (4.2)$$

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

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\mathbf{U}_i \leq \mathbf{u}\}}, \quad \alpha_n(\mathbf{u}) = \sqrt{n}\{C_n(\mathbf{u}) - C(\mathbf{u})\},$$

respectively, for  $\mathbf{u} \in [0, 1]^d$ . Consider the process

$$\bar{\mathbb{C}}_n(\mathbf{u}) = \alpha_n(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \alpha_n(1, \dots, 1, u_j, 1, \dots, 1), \quad \mathbf{u} \in [0, 1]^d,$$

with  $u_j$  appearing at the  $j$ -th coordinate. Note the slight but convenient abuse of notation in the definition of  $\bar{\mathbb{C}}_n$ : if  $\mathbf{u}$  is such that  $u_j \in \{0, 1\}$ , then  $\alpha_n(1, \dots, 1, u_j, 1, \dots, 1) = 0$  almost surely, so that the fact that for such  $\mathbf{u}$ , the partial derivative  $\dot{C}_j(\mathbf{u})$  has been left undefined in Condition 2.1 plays no role.

By (4.2) above and by Theorem 2.2 in Berghaus et al. (2017) (see also Remark 4.1 below),

$$\begin{aligned}\mathbb{C}_n^\beta/g^\omega &= \{\bar{\mathbb{C}}_n/g^\omega + o_{\mathbb{P}}(1)\} \mathbb{1}_{\{g \geq n^{-\gamma}\}}(1 + o(1)) + o_{\mathbb{P}}(1) \\ &= \bar{\mathbb{C}}_n/g^\omega \mathbb{1}_{\{g \geq n^{-\gamma}\}} + o_{\mathbb{P}}(1), \quad n \rightarrow \infty.\end{aligned}$$

In view of Lemma 4.9 in Berghaus et al. (2017), the indicator function can be omitted, and, applying Theorem 2.2 in the same reference again, we obtain

$$\mathbb{C}_n^\beta/g^\omega = \bar{\mathbb{C}}_n/g^\omega + o_{\mathbb{P}}(1) \rightsquigarrow \mathbb{C}_C/g^\omega, \quad n \rightarrow \infty,$$

as required. This finishes the proof of Theorem 2.2.

*Remark 4.1.* Some of the results in Berghaus et al. (2017) have to be adapted to the present situation.

- In the latter reference, the pseudo-observations are defined as  $\hat{U}_{i,j} = (n+1)^{-1}R_{i,j}$  rather than  $n^{-1}R_{i,j}$ . However, this does not affect the asymptotics, since the difference of the two empirical copulas is at most  $d/n$ , almost surely. For  $\mathbf{u} \in \{g \geq n^{-\gamma}\}$ , this modification makes a difference of the order  $O_{\mathbb{P}}(n^{\gamma\omega+1/2-1}) = o_{\mathbb{P}}(1)$ , as  $n \rightarrow \infty$ .
- In Theorem 2.2 in [Berghaus et al. \(2017\)](#), 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  $\{g \geq c/n\}$ . See Section 5.5 below for details.

## 5. Auxiliary results

Throughout and unless otherwise stated, we assume the conditions of Theorem 2.2.

### 5.1. Negligibility of the boundary regions

**Lemma 5.1.** *For  $\gamma > 1/\{2(1 - \omega)\}$ , we have*

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

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

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

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

an upper bound which vanishes as  $n \rightarrow \infty$  by the choice of  $\gamma$ .

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

$$\begin{aligned} |\mathbb{C}_n^\beta(\mathbf{u})/g(\mathbf{u})^\omega| &\leq \sqrt{n}(1 - u_1)^{-\omega} \{|C_n^\beta(\mathbf{u}) - u_2| + |C(\mathbf{u}) - u_2|\} \\ &\leq 2\sqrt{n}(1 - u_1)^{-\omega} \sum_{j \neq 2} (1 - u_j) \\ &\leq 2\sqrt{n} \sum_{j \neq 2} (1 - u_j)^{1-\omega} \\ &\leq 2(d-1)n^{1/2+\gamma\omega-\gamma} = o(1), \quad n \rightarrow \infty. \end{aligned}$$

The upper bounds do not depend on  $\mathbf{u} \in \{g \leq n^{-\gamma}\}$ , whence the uniformity in  $\mathbf{u}$ .  $\square$

### 5.2. The three terms in the decomposition (4.1)

The following lemma is to be compared with Proposition 3.5 in Segers et al. (2017). There, a pointwise approximation rate of  $O(n^{-1})$  was established. Here, we state a rate which is slightly slower,  $O(n^{-1} \ln n)$ , but uniformly in  $\mathbf{u}$ .

**Lemma 5.2.** *If  $C$  satisfies Condition 2.1, then*

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

*Proof.* Put  $\varepsilon_n = n^{-1} \ln n$ . First, we show that we can ignore those  $\mathbf{u}$  for which  $u_j \leq \varepsilon_n$  for some  $j \in \{1, \dots, d\}$ . Indeed, for such  $\mathbf{u}$ , the absolute value in the statement is bounded by

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

Let  $\mathbf{u} \in [\varepsilon_n, 1]^d$ . We show how to reduce the analysis to the case where  $\mathbf{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$ . Let  $J = J(\mathbf{u})$  denote the set of indices  $j = 1, \dots, d$  such that  $u_j > 1 - \varepsilon_n$  and suppose that  $J$  is not empty. Consider the vector  $\mathbf{e} \in \{0, 1\}^d$  which has components  $e_j = 1$  for  $j \in J$  and  $e_j = 0$  otherwise. For  $\mathbf{v} \in [0, 1]^d$ , the vector  $\mathbf{v} \vee \mathbf{e}$  has components  $(\mathbf{v} \vee \mathbf{e})_j$  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

$$\begin{aligned} & \left| \int_{[0,1]^d} \{C(\mathbf{w}) - C(\mathbf{u})\} d\mu_{n,\mathbf{u}}(\mathbf{w}) \right| \\ & \leq \left| \int_{[0,1]^d} \{C(\mathbf{w} \vee \mathbf{e}) - C(\mathbf{u} \vee \mathbf{e})\} d\mu_{n,\mathbf{u}}(\mathbf{w}) \right| \\ & \quad + \int_{[0,1]^d} |C(\mathbf{w}) - C(\mathbf{w} \vee \mathbf{e})| d\mu_{n,\mathbf{u}}(\mathbf{w}) + |C(\mathbf{u} \vee \mathbf{e}) - C(\mathbf{u})|. \end{aligned}$$

- 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  $\{1, \dots, d\} \setminus J$ . The copula of those variables is a multivariate margin of the original copula and Condition 2.1 applies to it as well. By construction, all remaining  $u_j$  are in the interval  $[\varepsilon_n, 1 - \varepsilon_n]$ , as required.
- We have  $|C(\mathbf{w}) - C(\mathbf{w} \vee \mathbf{e})| \leq \sum_{j \in J} |w_j - 1| \leq \sum_{j \in J} (|w_j - u_j| + \varepsilon_n)$ . By the Cauchy-Schwarz inequality,  $\int_{[0,1]^d} |w_j - u_j| d\mu_{n,\mathbf{u}}(\mathbf{w}) \leq \{n^{-1} u_j (1 - u_j)\}^{1/2} \leq n^{-1/2} \varepsilon_n^{1/2} \leq \varepsilon_n$ . Hence  $\int_{[0,1]^d} |C(\mathbf{w}) - C(\mathbf{w} \vee \mathbf{e})| d\mu_{n,\mathbf{u}}(\mathbf{w}) \leq 2d\varepsilon_n$ .
- Finally,  $|C(\mathbf{u} \vee \mathbf{e}) - C(\mathbf{u})| \leq \sum_{j \in J} (1 - u_j) \leq d\varepsilon_n$ .

It remains to consider the case  $\mathbf{u} \in [\varepsilon_n, 1 - \varepsilon_n]^d$ . As in the proof of Proposition 3.4 in Segers et al. (2017), we have

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

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

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  $0 \leq \dot{C}_j \leq 1$ ), and the mass on the boundary is  $\mu_{n,\mathbf{u}}([0, 1]^d \setminus (0, 1)^d) = \mathbb{P}[\exists j : S_j \in \{0, n\}] \leq 2d(1 - \varepsilon_n)^n \leq 2d \exp(-n\varepsilon_n) = 2dn^{-1} = o(\varepsilon_n)$  as  $n \rightarrow \infty$ .

In view of the second part of Condition 2.1, we have

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

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

We apply the bound in (2.7) to  $\ddot{C}_{jk}(\mathbf{u} + st(\mathbf{w} - \mathbf{u}))$ . We have  $\min(a^{-1}, b^{-1}) \leq (ab)^{-1/2}$ , 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,

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

We obtain

$$\begin{aligned} & \left| \int_{(0,1)^d} (w_j - u_j)(w_k - u_k) \ddot{C}_{jk}(\mathbf{u} + st(\mathbf{w} - \mathbf{u})) d\mu_{n,\mathbf{u}}(\mathbf{w}) \right| \\ & \leq K \int_{(0,1)^d} \left[ \frac{|(w_j - u_j)(w_k - u_k)|}{\{u_j(1 - u_j)u_k(1 - u_k)\}^{1/2}} + \frac{|(w_j - u_j)(w_k - u_k)|}{\{w_j(1 - w_j)w_k(1 - w_k)\}^{1/2}} \right] d\mu_{n,\mathbf{u}}(\mathbf{w}). \end{aligned}$$

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

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

Second, again by Cauchy–Schwarz inequality,

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

Each of the two integrals ( $i = j$  and  $i = k$ ), and therefore their geometric mean, will be bounded by the same quantity. Note that  $\frac{1}{w_i(1-w_i)} = \frac{1}{w_i} + \frac{1}{1-w_i}$  and that the integral involving  $\frac{1}{1-w_i}$  is equal to the one involving  $\frac{1}{w_i}$  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,

$$\begin{aligned} \int_{(0,1)^d} \frac{(w_i - u_i)^2}{w_i} d\mu_{n,\mathbf{u}}(\mathbf{w}) &\leq \int_{[0,1]^d} \mathbb{1}_{(0,1]}(w_i) \frac{(w_i - u_i)^2}{w_i} d\mu_{n,\mathbf{u}}(\mathbf{w}) \\ &= \int_{[0,1]^d} \mathbb{1}_{(0,1]}(w_i) (w_i - 2u_i + \frac{u_i^2}{w_i}) d\mu_{n,\mathbf{u}}(\mathbf{w}) \\ &= u_i - 2u_i \mathbb{P}[S_i/n > 0] + u_i^2 \mathbb{E}[\frac{1}{S_i/n} \mathbb{1}_{\{S_i/n > 0\}}] \\ &\leq -u_i + 2\mathbb{P}[S_i = 0] + nu_i^2 \mathbb{E}[\frac{1}{S_i} \mathbb{1}_{\{S_i \geq 1\}}]. \end{aligned}$$

Recall that  $u_i \in [\varepsilon_n, 1 - \varepsilon_n]$  and thus  $\mathbb{P}[S_i = 0] \leq (1 - \varepsilon_n)^n \leq \exp(-n\varepsilon_n) = n^{-1} = o(\varepsilon_n)$ . Further, the expectation of the reciprocal of a binomial random variable is treated in Lemma 5.7. Note that  $n\varepsilon_n = \ln n \rightarrow \infty$ . We find

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

The proof is complete.  $\square$

**Lemma 5.3.** *For any  $1/\{2(1 - \omega)\} < \gamma < 1$ , we have*

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

*Proof.* Since  $g(\frac{S_1}{n}, \dots, \frac{S_d}{n})$  is a random variable taking values in  $[0, 1]$ , we can write

$$\begin{aligned} \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) &= \frac{1}{g(\mathbf{u})^\omega} \mathbb{E} \left[ g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right)^\omega \right] \\ &= \frac{1}{g(\mathbf{u})^\omega} \int_0^1 \mathbb{P} \left\{ g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) > t^{1/\omega} \right\} dt. \end{aligned} \quad (5.1)$$

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

On the one hand, we find

$$\begin{aligned} \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) &\leq \frac{a_{n,+}}{g(\mathbf{u})^\omega} + \frac{1-a_{n,+}}{g(\mathbf{u})^\omega} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) > a_{n,+}^{1/\omega}\right\} \\ &\leq (1+\varepsilon_n)^\omega + g(\mathbf{u})^{-\omega} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) > g(\mathbf{u})(1+\varepsilon_n)\right\} \\ &\leq 1+\varepsilon_n + g(\mathbf{u})^{-\omega} 2d \exp\{-ng(\mathbf{u})h(1+\varepsilon_n)\} \end{aligned}$$

where we used (5.6) in the last step. Since  $h(1+\varepsilon_n) \geq \frac{1}{3}\varepsilon_n^2$  for  $0 \leq \varepsilon_n \leq 1$  and since  $g(\mathbf{u}) \geq n^{-\gamma}$ , the upper bound is bounded by

$$1 + \varepsilon_n + 2dn^{\gamma\omega} \exp\{-\frac{1}{3}(\ln n)^2\} = 1 + \varepsilon_n + o(\varepsilon_n), \quad n \rightarrow \infty.$$

On the other hand, restricting the integral in (5.1) to  $[0, a_{n,-}]$ , we have

$$\begin{aligned} \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) &\geq \frac{a_{n,-}}{g(\mathbf{u})^\omega} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) > a_{n,-}^{1/\omega}\right\} \\ &= (1-\varepsilon_n)^\omega \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) > g(\mathbf{u})(1-\varepsilon_n)\right\} \\ &\geq (1-\varepsilon_n)^\omega [1 - 4d \exp\{-ng(\mathbf{u})h(1+\varepsilon_n)\}], \end{aligned}$$

where we used (5.7) in the last step. Since  $0 \leq \varepsilon_n \rightarrow 0$  and  $g(\mathbf{u}) \geq n^{-\gamma}$ , the lower bound is bounded from below by

$$\begin{aligned} (1-\varepsilon_n)^\omega [1 - 4d \exp\{-ng(\mathbf{u})h(1+\varepsilon_n)\}] &\geq (1-\varepsilon_n)[1 - 4d \exp\{-\frac{1}{3}(\ln n)^2\}] \\ &\geq 1 - \varepsilon_n - o(\varepsilon_n), \quad n \rightarrow \infty. \end{aligned}$$

□

**Lemma 5.4.** *As  $n \rightarrow \infty$ , we have, for any  $\gamma \in (1/(2(1-\omega)), 1)$*

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

*Proof.* Let  $\delta_n = 1/\ln(n)$ . Write  $\|\hat{\mathbb{C}}_n\|_\infty = \sup\{|\hat{\mathbb{C}}_n(\mathbf{v})| : \mathbf{v} \in [0,1]^d\}$ . We have  $\|\hat{\mathbb{C}}_n\|_\infty = O_p(1)$  as  $n \rightarrow \infty$  by weak convergence of  $\hat{\mathbb{C}}_n$  in  $\ell^\infty([0,1]^d)$ .

We split the integral over  $\mathbf{w} \in [0,1]^d$  into two pieces: the integral over the domain

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

and the integral over its complement; here  $|\mathbf{x}|_\infty = \max\{|x_j| : j = 1, \dots, d\}$ .

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

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



Moreover, for all  $\mathbf{u} \in \{g \geq n^{-\gamma}\}$ , using Chebyshev's inequality and the concentration inequality (5.7), we have

$$\begin{aligned}\mu_{n,\mathbf{u}}(A_{n,\mathbf{u}}) &\leq \sum_{j=1}^d \mathbb{P}\left\{\left|\frac{S_j}{n} - u_j\right| > \delta_n\right\} + \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) < g(\mathbf{u})(1 - \delta_n)\right\} \\ &\leq dn^{-1}\delta_n^{-2} + 4d \exp\{-n^{1-\gamma}h(1 + \delta_n)\}.\end{aligned}$$

Since  $0 < \omega < 1/2$ ,  $0 < \gamma < 1$ ,  $\delta_n = 1/\ln(n)$  and  $h(1 + \delta_n) \geq \frac{1}{3}\delta_n^2$ , it follows that

$$\begin{aligned}\sup_{\mathbf{u} \in \{g \geq n^{-\gamma}\}} \int_{A_{n,\mathbf{u}}} R_n(\mathbf{u}, \mathbf{w}) d\mu_{n,\mathbf{u}}(\mathbf{w}) \\ \leq n^{2\gamma\omega} [n^{-1}\delta_n^{-2} + \exp\{-n^{1-\gamma}h(1 + \delta_n)\}] O_p(1) = o_p(1), \quad n \rightarrow \infty.\end{aligned}$$

It remains to consider the integral over  $\mathbf{w} \in [0, 1]^d \setminus A_{n,\mathbf{u}}$ , i.e.,  $|\mathbf{w} - \mathbf{u}|_\infty \leq \delta_n$  and  $g(\mathbf{w}) \geq n^{-\gamma}(1 - \delta_n) > n^{-1}$ , at least for sufficiently large  $n$ . By Lemma 4.1 in [Berghaus et al. \(2017\)](#), we have

$$\sup_{\substack{\mathbf{u}, \mathbf{w} \in \{g \geq n^{-1}\} \\ |\mathbf{u} - \mathbf{w}|_\infty \leq \delta_n}} \left| \frac{\hat{\mathbb{C}}_n(\mathbf{w})}{g(\mathbf{w})^\omega} - \frac{\hat{\mathbb{C}}_n(\mathbf{u})}{g(\mathbf{u})} \right| = o_p(1), \quad n \rightarrow \infty. \quad (5.2)$$

In view of Lemma 5.3, we obtain that

$$\sup_{\mathbf{u} \in \{g \geq n^{-\gamma}\}} \int_{[0,1]^d \setminus A_{n,\mathbf{u}}} R_n(\mathbf{u}, \mathbf{w}) d\mu_{n,\mathbf{u}}(\mathbf{w}) \leq o_p(1) \int_{[0,1]^d} \frac{g(\mathbf{w})^\omega}{g(\mathbf{u})^\omega} d\mu_{n,\mathbf{u}}(\mathbf{w}) = o_p(1),$$

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

Note that in Lemma 4.1 in [Berghaus et al. \(2017\)](#), the supremum in (5.2) 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  $\{g \geq n^{-1}\}$ . Furthermore, in the latter reference, the pseudo-observations are defined as  $\hat{U}_{i,j} = \frac{1}{n+1}R_{i,j}$ . 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  $\{g \geq n^{-1}\}$  which is of the order  $O_p(n^{\omega+1/2-1}) = o_p(1)$ , as  $n \rightarrow \infty$ .  $\square$

### 5.3. On the expectation of the reciprocal of a binomial random variable

**Lemma 5.5.** *Let  $0 < u \leq 1$  and let  $n \geq 2$  be integer. If  $S \sim \text{Bin}(n, u)$  and  $T \sim \text{Bin}(n-1, u)$ , then*

$$\mathbb{E} \left[ \frac{1}{S} \mathbb{1}_{\{S \geq 1\}} \right] = nu \mathbb{E} \left[ \frac{1}{(1+T)^2} \right] = nu \int_0^1 (1-u+us)^{n-1} (-\ln s) ds. \quad (5.3)$$

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

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

We obtain that

$$\mathbb{E}\left[\frac{1}{S} \mathbb{1}_{\{S \geq 1\}}\right] = \sum_{k=1}^n \frac{1}{k} \mathbb{P}(S = k) = \sum_{k=1}^n \frac{1}{k} \frac{nu}{k} \mathbb{P}(T + 1 = k) = (nu) \mathbb{E}\left[\frac{1}{(1 + T)^2}\right].$$

Now we apply a trick due to [Chao and Strawderman \(1972\)](#): we have

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

Taking expectations and using Fubini's theorem, we obtain

$$\mathbb{E}\left[\frac{1}{(1 + T)^2}\right] = \int_0^1 \mathbb{E}[s^T] (-\ln s) \, ds = \int_0^1 (1 - u + us)^{n-1} (-\ln s) \, ds,$$

as required.  $\square$

**Lemma 5.6.** *Let  $0 < u_n \leq 1$  and let  $S_n \sim \text{Bin}(n, u_n)$ . If  $nu_n \rightarrow \infty$ , then*

$$(nu_n^2) \mathbb{E}\left[\frac{1}{S_n} \mathbb{1}_{\{S_n \geq 1\}}\right] = u_n + O(n^{-1}), \quad n \rightarrow \infty.$$

*Proof.* We start from (5.3):

$$(nu_n^2) \mathbb{E}\left[\frac{1}{S_n} \mathbb{1}_{\{S_n \geq 1\}}\right] = n^2 u_n^3 \int_0^1 (1 - u_n + u_n s)^{n-1} (-\ln s) \, ds$$

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

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

$$\begin{aligned} n^2 u_n^3 \int_0^{1/2} (1 - u_n + u_n s)^{n-1} (-\ln s) \, ds &\leq n^2 u_n^3 (1 - u_n/2)^{n-1} \int_0^{1/2} (-\ln s) \, ds \\ &\leq K n^2 u_n^3 (1 - u_n/2)^n \\ &\leq K n^2 u_n^3 \exp(-nu_n/2). \end{aligned}$$

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

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

$$\begin{aligned} n^2 u_n^3 \int_{1/2}^1 (1 - u_n + u_n s)^{n-1} (-\ln s) \, ds \\ = u_n \int_0^{(nu_n/2)} (1 - v/n)^{n-1} [-(nu_n) \ln\{1 - v/(nu_n)\}] \, dv. \end{aligned} \quad (5.4)$$

We need to show that this integral is  $u_n + O(n^{-1})$  as  $n \rightarrow \infty$ .

For facility of writing, put  $k_n = nu_n$ . Recall that  $k_n \rightarrow \infty$  as  $n \rightarrow \infty$  by assumption. The inequalities  $x \leq -\ln(1-x) \leq x/(1-x)$  for  $0 \leq x < 1$  imply that

$$0 \leq -k_n \ln(1 - v/k_n) - v \leq \frac{v^2}{k_n - v} \leq \frac{2v^2}{k_n}, \quad v \in [0, k_n/2].$$

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

$$\begin{aligned} u_n \int_0^{k_n/2} (1 - v/n)^{n-1} | -k_n \ln(1 - v/k_n) - v | \, dv &\leq \frac{4u_n}{k_n} \int_0^{k_n/2} \exp(-v) v^2 \, dv \\ &= O(n^{-1}), \quad n \rightarrow \infty. \end{aligned}$$

As a consequence, replacing  $-k_n \ln(1 - v/k_n)$  by  $v$  in (5.4) produces an error of the required order  $O(n^{-1})$ .

It remains to consider the integral

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

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  $\{1 - k_n/(2n)\}^n \leq \exp(-k_n/2)$ , the previous expression is

$$u_n + O(u_n n^{-1}) + O(u_n \exp(-k_n/2) k_n), \quad n \rightarrow \infty,$$

The error term is  $O(n^{-1})$ , as required. □

**Lemma 5.7.** *If  $0 < u_n \leq 1$  is such that  $nu_n \rightarrow \infty$  as  $n \rightarrow \infty$ , then*

$$\sup_{u_n \leq u \leq 1} \left| nu^2 \mathbb{E} \left[ \frac{1}{S} \mathbb{1}_{\{S \geq 1\}} \right] - u \right| = O(n^{-1}), \quad n \rightarrow \infty,$$

where the expectation is taken for  $S \sim \text{Bin}(n, u)$ .

*Proof.* The function sending  $u \in [u_n, 1]$  to  $|nu^2 \mathbb{E}[S^{-1} \mathbb{1}_{\{S \geq 1\}}] - u|$ , with  $S \sim \text{Bin}(n, u)$ , is continuous and therefore attains its supremum at some  $v_n \in [u_n, 1]$ . Since  $nv_n \geq nu_n \rightarrow \infty$  as  $n \rightarrow \infty$ , we can apply Lemma 5.6 to find that the supremum is  $O(n^{-1})$  as  $n \rightarrow \infty$ . □

#### 5.4. Inequalities for binomial random variables

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

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

for  $\lambda > 0$ , where  $\psi(x) = 2h(1+x)/x^2$  and  $h(x) = x(\ln x - 1) + 1$ ; see for instance [van der Vaart and Wellner \(1996, Proposition A.6.2\)](#). Setting  $\lambda = \sqrt{nu}\delta$ , we find

$$\Pr\left(\left|\frac{S}{n} - u\right| \geq u\delta\right) \leq 2 \exp\{-nu h(1 + \delta)\}, \quad \delta > 0. \quad (5.5)$$

Note that  $h(1 + \delta) = \int_0^\delta \ln(1 + t) dt \geq \int_0^\delta (t - \frac{1}{2}t^2) dt = \frac{1}{2}\delta^2(1 - \frac{1}{3}\delta)$  for  $\delta \geq 0$  and thus  $h(1 + \delta) \geq \frac{1}{3}\delta^2$  for  $0 \leq \delta \leq 1$ . We extend (5.5) to a vector of independent binomial random variables and in terms of the weight function  $g$  in (2.3).

**Lemma 5.8.** *If  $S_1, \dots, S_d$  are independent random variables with  $S_j \sim \text{Bin}(n, u_j)$  and  $0 < u_j < 1$  for all  $j \in \{1, \dots, d\}$ , then, for  $\delta > 0$ ,*

$$\mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) \geq g(\mathbf{u})(1 + \delta)\right\} \leq 2d \exp\{-ng(\mathbf{u})h(1 + \delta)\}, \quad (5.6)$$

$$\mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) \leq g(\mathbf{u})(1 - \delta)\right\} \leq 4d \exp\{-ng(\mathbf{u})h(1 + \delta)\}, \quad (5.7)$$

with  $h$  as above; in particular,  $h(1 + \delta) \geq \frac{1}{3}\delta^2$  for  $0 < \delta \leq 1$ .

*Proof.* Let us start with (5.7). The definition of the weight function  $g$  in (2.3) yields

$$\begin{aligned} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) \leq g(\mathbf{u})(1 - \delta)\right\} \\ \leq \sum_{j=1}^d \left[ \mathbb{P}\left\{\frac{S_j}{n} \leq g(\mathbf{u})(1 - \delta)\right\} + \mathbb{P}\left\{\max_{k \neq j} \left(1 - \frac{S_k}{n}\right) \leq g(\mathbf{u})(1 - \delta)\right\} \right]. \end{aligned}$$

Let us first consider the first term on the right-hand side, i.e.,  $\mathbb{P}\left\{\frac{S_j}{n} \leq g(\mathbf{u})(1 - \delta)\right\}$ . By definition of the weight function we have  $g(\mathbf{u}) \leq u_j$ . By Bennett's inequality (5.5),

$$\begin{aligned} \mathbb{P}\left\{\frac{S_j}{n} \leq g(\mathbf{u})(1 - \delta)\right\} &\leq \mathbb{P}\left\{\frac{S_j}{n} \leq u_j(1 - \delta)\right\} \\ &\leq \mathbb{P}\left\{\left|\frac{S_j}{n} - u_j\right| \geq u_j\delta\right\} \\ &\leq 2 \exp\{-nu_j h(1 + \delta)\} \\ &\leq 2 \exp\{-ng(\mathbf{u})h(1 + \delta)\}. \end{aligned}$$

Second, consider the term  $\mathbb{P}\left\{\max_{k \neq j} \left(1 - \frac{S_k}{n}\right) \leq g(\mathbf{u})(1 - \delta)\right\}$ . Suppose  $j = 1$ ; the other cases can be treated exactly along the same lines. We have  $g(\mathbf{u}) \leq \max_{k \neq 1} (1 - u_k)$ .

Assume without loss of generality that  $\max_{k \neq 1}(1 - u_k) = 1 - u_2$ . Then we obtain  $g(\mathbf{u}) \leq 1 - u_2$  and, by Bennett's inequality (5.5) applied to  $n - S_2 \sim \text{Bin}(n, 1 - u_2)$ ,

$$\begin{aligned} \mathbb{P}\left\{\max_{k \neq 1}\left(1 - \frac{S_k}{n}\right) \leq g(\mathbf{u})(1 - \delta)\right\} &\leq \mathbb{P}\left\{\max_{k \neq 1}\left(1 - \frac{S_k}{n}\right) \leq (1 - u_2)(1 - \delta)\right\} \\ &\leq \mathbb{P}\left\{1 - \frac{S_2}{n} \leq (1 - u_2)(1 - \delta)\right\} \\ &\leq \mathbb{P}\left\{\left|1 - \frac{S_2}{n} - (1 - u_2)\right| \geq (1 - u_2)\delta\right\} \\ &\leq 2 \exp\{-n(1 - u_2)h(1 + \delta)\} \\ &\leq 2 \exp\{-ng(\mathbf{u})h(1 + \delta)\}. \end{aligned}$$

Let us now show (5.6). First suppose  $g(\mathbf{u}) = u_1$ . Since  $g(\frac{S_1}{n}, \dots, \frac{S_d}{n}) \leq \frac{S_1}{n}$  we have, by Bennett's inequality (5.5),

$$\begin{aligned} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) \geq g(\mathbf{u})(1 + \delta)\right\} &\leq \mathbb{P}\left\{\frac{S_1}{n} \geq u_1(1 + \delta)\right\} \\ &\leq \mathbb{P}\left\{\left|\frac{S_1}{n} - u_1\right| \geq u_1\delta\right\} \\ &\leq 2 \exp\{-nu_1h(1 + \delta)\} = 2 \exp\{-ng(\mathbf{u})h(1 + \delta)\}. \end{aligned}$$

Finally, suppose that  $g(\mathbf{u}) = 1 - u_1 \geq 1 - u_k$ , for  $k = 3, \dots, d$ . Note that  $g(\frac{S_1}{n}, \dots, \frac{S_d}{n}) \leq \max_{k \neq 2}(1 - \frac{S_k}{n})$ , which yields

$$\begin{aligned} \mathbb{P}\left\{g\left(\frac{S_1}{n}, \dots, \frac{S_d}{n}\right) \geq g(\mathbf{u})(1 + \delta)\right\} &\leq \mathbb{P}\left\{\max_{k \neq 2}\left(1 - \frac{S_k}{n}\right) \geq (1 - u_1)(1 + \delta)\right\} \\ &\leq \sum_{k \neq 2} \mathbb{P}\left\{1 - \frac{S_k}{n} \geq (1 - u_1)(1 + \delta)\right\}. \end{aligned}$$

By Bennett's inequality (5.5) applied to  $n - S_k \sim \text{Bin}(n, 1 - u_k)$  for every  $k \neq 2$ , we have, since  $(1 - u_1)/(1 - u_k) \geq 1$ ,

$$\begin{aligned} \mathbb{P}\left\{1 - \frac{S_k}{n} \geq (1 - u_1)(1 + \delta)\right\} &\leq \mathbb{P}\left\{\left|1 - \frac{S_k}{n} - (1 - u_k)\right| \geq (1 - u_1)(1 + \delta) - (1 - u_k)\right\} \\ &\leq 2 \exp\left\{-n(1 - u_k)h\left(\frac{1 - u_1}{1 - u_k}(1 + \delta)\right)\right\}. \end{aligned}$$

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

$$\begin{aligned} \mathbb{P}\left\{1 - \frac{S_k}{n} \geq (1 - u_1)(1 + \delta)\right\} &\leq 2 \exp\left\{-n(1 - u_k)\frac{1 - u_1}{1 - u_k}h(1 + \delta)\right\} \\ &= 2 \exp\{-n(1 - u_1)h(1 + \delta)\} = 2 \exp\{-ng(\mathbf{u})h(1 + \delta)\}. \end{aligned}$$

□

---

<sup>1</sup>Or, since  $h(x) = \int_1^x \ln(t) dt$ , we have  $h(a(1 + \delta)) = \int_1^{a(1 + \delta)} \ln(t) dt = a \int_{1/a}^{1 + \delta} \ln(as) ds \geq a \int_1^{1 + \delta} \ln(s) ds = a h(1 + \delta)$  for  $a \geq 1$  and  $\delta \geq 0$ .

### 5.5. Extensions of results in [Berghaus et al. \(2017\)](#)

For any sequence  $\delta_n > 0$  that converges to zero as  $n \rightarrow \infty$ , Lemma 4.10 in [Berghaus et al. \(2017\)](#) can be extended to

$$\sup_{\substack{\mathbf{u}, \mathbf{u}' \in \{g \geq c/n\} \\ |\mathbf{u} - \mathbf{u}'| \leq \delta_n}} \left| \frac{\mathbb{C}_n(\mathbf{u})}{g(\mathbf{u})^\omega} - \frac{\mathbb{C}_n(\mathbf{u}')}{g(\mathbf{u}')^\omega} \right| = o_{\mathbb{P}}(1), \quad n \rightarrow \infty. \quad (5.8)$$

Furthermore, Theorem 4.5 in the same reference can be extended to

$$\sup_{\mathbf{u} \in \{g \geq c/n\}} \left| \frac{\hat{\mathbb{C}}_n(\mathbf{u})}{g(\mathbf{u})^\omega} - \frac{\bar{\mathbb{C}}_n(\mathbf{u})}{g(\mathbf{u})^\omega} \right| = o_{\mathbb{P}}(1), \quad n \rightarrow \infty. \quad (5.9)$$

*Proof.* Let us start with (5.8). The result is similar to the result in Lemma 4.10, in particular Equation (4.1), in [Berghaus et al. \(2017\)](#). A look at the proof of the result shows that the restriction  $\mathbf{u}, \mathbf{u}' \in [c/n, 1 - c/n]^d$  instead of  $\mathbf{u}, \mathbf{u}' \in \{g \geq c/n\}$  is not needed. The proof of Equation (4.1) in Lemma 4.10 in [Berghaus et al. \(2017\)](#) is based on Lemma 4.7, 4.8 and Equations (4.8) and (4.8) 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  $\mathbf{u}, \mathbf{u}' \in \{g \geq c/n\}$  instead of  $\mathbf{u}, \mathbf{u}' \in [c/n, 1 - c/n]^d$ , which gives us exactly (5.8).

For the proof of (5.9) note that for any  $\mathbf{u} \in \{g \geq c/n\}$  we can find  $\mathbf{u}' \in \{g \geq n^{-1/2}\}$  such that  $|\mathbf{u} - \mathbf{u}'| \leq dn^{-1/2}$ . To find such a  $\mathbf{u}'$  is all that it is needed to extend the proof of Theorem 4.5 in [Berghaus et al. \(2017\)](#) to obtain (5.9).  $\square$

### Acknowledgments

Betina Berghaus gratefully acknowledges support by the Collaborative Research Center “Statistical modeling of nonlinear dynamic processes” (SFB 823) of the German Research Foundation (DFG).

Johan Segers gratefully acknowledges funding by contract “Projet d’Actions de Recherche Concertées” No. 12/17-045 of the “Communauté française de Belgique” and by IAP research network Grant P7/06 of the Belgian government.

### References

- Berghaus, B., A. Bücher, and S. Volgushev (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli* 23(1), 743–772.
- Berghaus, B., A. Bücher, and H. Dette (2013). Minimum distance estimators of the Pickands dependence function and related tests of multivariate extreme-value dependence. *Journal de la Société Française de Statistique* 154(1), 116–137.
- Bücher, A., H. Dette, and S. Volgushev (2011). New estimators of the Pickands dependence function and a test for extreme-value dependence. *The Annals of Statistics* 39(4), 1963–2006.

- Bücher, A. and S. Volgushev (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis* 119, 61–70.
- Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84(3), 567–577.
- Chao, M. T. and W. E. Strawderman (1972). Negative moments of positive random variables. *Journal of the American Statistical Association* 67(338), 429–431.
- 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(6), 274–292.
- Fermanian, J.-D., D. Radulović, and M. Wegkamp (2004). Weak convergence of empirical copula processes. *Bernoulli* 10(5), 847–860.
- Gaenssler, P. and W. Stute (1987). *Seminar on empirical processes*, Volume 9 of *DMV Seminar*. Basel: Birkhäuser Verlag.
- Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist.* 37(5B), 2990–3022.
- Gudendorf, G. and J. Segers (2010). Extreme-value copulas. In *Copula theory and its applications*, Volume 198 of *Lect. Notes Stat. Proc.*, pp. 127–145. Springer, Heidelberg.
- Janssen, P., J. Swanepoel, and N. Veraverbeke (2012). Large sample behavior of the Bernstein copula estimator. *Journal of Statistical Planning and Inference* 142, 1189–1197.
- Kojadinovic, I. and J. Yan (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software* 34(9), 1–20.
- Pickands, III, J. (1981). 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. With a discussion.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Annals of Statistics* 4, 912–923.
- Sancetta, A. and S. Satchell (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory* 20, 535–562.
- Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* 18(3), 764–782.

- Segers, J., M. Sibuya, and H. Tsukahara (2017). The empirical beta copula. *Journal of Multivariate Analysis* 155, 35–51.
- van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes - Springer Series in Statistics*. New York: Springer.