I N S T I T U T D E S T A T I S T I Q U E

B I O S T A T I S T I Q U E E T

S C I E N C E S A C T U A R I E L L E S

(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2014/18

Hybrid Copula Estimators

J. SEGERS

Hybrid Copula Estimators

Johan Segers

Université catholique de Louvain
ISBA, Voie du Roman Pays 20, bte L1.04.01
B-1348 Louvain-la-Neuve, Belgium
johan.segers@uclouvain.be

May 8, 2014

Abstract

A variant of the empirical copula is considered which combines an estimator of a multivariate cumulative distribution function with estimators of its margins that are not necessarily equal to the margins of the estimator of the joint distribution. Such a hybrid estimator may be reasonable when there is additional information available for some margins in the form of additional data or stronger modelling assumptions. A functional central limit theorem is established and some examples are developed.

1 Introduction

Let H be a p-variate cumulative distribution function with continuous margins $F_1, \ldots, F_p$ and copula C (Sklar, 1959). We have

$$\begin{aligned} H(\boldsymbol{x}) &= C\bigl(F_1(x_1), \ldots, F_p(x_p)\bigr), && \boldsymbol{x} \in \mathbb{R}^p, \\ C(\boldsymbol{u}) &= H\bigl(F_1^{\leftarrow}(u_1), \ldots, F_p^{\leftarrow}(u_p)\bigr), && \boldsymbol{u} \in [0,1]^p. \end{aligned}$$

Here, $G^{\leftarrow}$ denotes the left-continuous inverse of a univariate cumulative distribution function G, i.e.,

$$G^{\leftarrow}(u) = \inf\{x \in \mathbb{R} : G(x) \geqslant u\}, \qquad u \in [0,1].$$

Throughout, standard conventions regarding infinities are employed: $\inf \varnothing = +\infty$, $G(-\infty) = 0$, and $G(+\infty) = 1$.

Let $\hat{H}_n$ and $\hat{F}_{n,j}$ be estimator sequences of H and F_j ($j = 1, \ldots, p$), respectively. Consider the copula estimator

$$\hat{C}_n(\boldsymbol{u}) = \hat{H}_n\bigl(\hat{F}_{n,j}^{\leftarrow}(u_1), \ldots, \hat{F}_{n,p}^{\leftarrow}(u_p)\bigr), \qquad \boldsymbol{u} \in [0,1]^p. \tag{1.1}$$

Note that $\hat{F}_{n,j}$ is not necessarily equal to the jth marginal distribution function, $\hat{H}_{n,j}$, of $\hat{H}_n$. We call $\hat{C}_n$ a *hybrid copula estimator*.

Given a rate $0 < r_n \to \infty$ (typically $r_n = \sqrt{n}$), the normalized estimation error of the hybrid copula estimator is

$$\mathbb{C}_n(\boldsymbol{u}) = r_n\bigl(\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\bigr), \qquad \boldsymbol{u} \in [0,1]^p. \tag{1.2}$$

The aim is to establish weak convergence of $\mathbb{C}_n$ in the space $\ell^\infty([0,1]^p)$ of bounded, real-valued functions on $[0,1]^p$ equipped with the supremum norm.

If $\hat{H}_n$ and $\hat{F}_{n,j} = \hat{H}_{n,j}$ are the joint and marginal empirical distribution functions of a p-variate sample of size n, then $\hat{C}_n$ is just the Deheuvels–Rüschendorf empirical copula, see Examples 3.1 and 3.2 below. However, there may be good reasons not to estimate F_j by $\hat{H}_{n,j}$ but by a different estimator. It may be that there is information available on the jth margin which cannot directly be used by the joint estimator $\hat{H}_n$.

- A parametric model may be reasonable for some or all of the marginal distributions but not for the joint distribution (Example 3.4). This is the case for instance when the data are vectors of annual maxima. Asymptotic theory then suggests to model the vector of componentwise maxima by a multivariate max-stable distribution (de Haan and Resnick, 1977; Deheuvels, 1978; Galambos, 1978). The marginal distributions are univariate extreme-value distributions, whereas the copula belongs to the infinite-dimensional family of extreme-value copulas.
- Some entries in the $n \times p$ data matrix may be missing (Example 3.5). Then $\hat{H}_n$ may be defined as the empirical distribution function of all data rows which are complete, whereas $\hat{F}_{n,j}$ is the empirical distribution function of all observed entries in the jth column.
- Similarly, in a time series setting, the observation periods of the p univariate series could be different and overlap only partially. Again, one could estimate F_j by the complete series for that variable but estimate H only based on the time period where all series were recorded simultaneously. In the same spirit, there may be additional samples for some of the variables.

The structure of the paper is as follows. The main result, Theorem 2.3, is given in Section 2, stating weak convergence of the hybrid copula estimator process in (1.2) under high-level conditions on the estimators of the joint and marginal distribution functions. Special cases and examples are worked out in Section 3. All proofs and calculations are deferred to Section 4.

Throughout, the following notations are used. For an arbitrary set T, let $\ell^\infty(T)$ be the space of bounded, real-valued functions on T, the space being equipped with the supremum distance $\|f\|_\infty = \sup_{t \in T} |f(t)|$ for $f \in \ell^\infty(T)$. The indicator variable of a set E is denoted by $\mathbb{1}_E$, whereas the identity mapping on a set E is denoted by id_E. Weak convergence in the sense of J. Hoffmann-Jørgensen is denoted by the arrow '$\rightsquigarrow$'; see Part 1 in the monograph by van der Vaart and Wellner (1996).

2 Main result

Besides the continuity of the margins $F_1, \ldots, F_p$, two assumptions will be made. The first assumption imposes among others a bit of smoothness on the target copula C (Segers, 2012) without which there is litte hope of establishing weak convergence of $\mathbb{C}_n$ in (1.2) with respect to the supremum norm on $\ell^\infty([0,1]^p)$. The second assumption is a high-level condition concerning the asymptotic distribution of the estimators $\hat{H}_n$ and $\hat{F}_{n,j}$ and is to be checked on a case-by-case basis. See Remarks 2.4 and 2.5 and see the examples in Section 3.

Condition 2.1. *(a) The p-variate distribution function H has continuous margins $F_1, \ldots, F_p$ and copula C.*

(b) For all $j \in \{1, \ldots, p\}$, the first-order partial derivative $\dot{C}_j(\boldsymbol{u}) = \partial C(\boldsymbol{u})/\partial u_j$ exists and is continuous on the set $\{\boldsymbol{u} \in [0,1]^p : 0 < u_j < 1\}$.

For convenience, collect the marginal distribution and quantile functions into vector-valued functions:

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{x}) &= \big(F_1(x_1), \ldots, F_p(x_p)\big), && \boldsymbol{x} \in \mathbb{R}^p, \\ \boldsymbol{F}^{\leftarrow}(\boldsymbol{u}) &= \big(F_1^{\leftarrow}(u_1), \ldots, F_p^{\leftarrow}(u_p)\big), && \boldsymbol{u} \in [0,1]^p. \end{aligned}$$

Condition 2.2. *There exists $0 < r_n \to \infty$ such that in the space $\ell^\infty(\mathbb{R}^p) \otimes (\ell^\infty(\mathbb{R}) \otimes \cdots \otimes \ell^\infty(\mathbb{R}))$ equipped with the topology of uniform convergence, we have joint weak convergence*

$$\begin{aligned} &\big(r_n(\hat{H}_n - H);\, r_n(\hat{F}_{n,1} - F_1), \ldots, r_n(\hat{F}_{n,p} - F_p)\big) \\ &\qquad \rightsquigarrow (\alpha \circ \boldsymbol{F}; \beta_1 \circ F_1, \ldots, \beta_p \circ F_p), \qquad n \to \infty, \end{aligned} \tag{2.1}$$

the stochastic processes α and β_j taking values in $\ell^\infty([0,1]^p)$ and $\ell^\infty([0,1])$, respectively, and having continuous trajectories almost surely.

Usually, $r_n = \sqrt{n}$, although Condition 2.2 allows for different convergence rates. Joint weak convergence in (2.1) can typically be established when the estimators $\hat{H}_n$ and $\hat{F}_{n,j}$ can be written as functionals of the same underlying empirical process. Because $\dot{C}_j(\boldsymbol{u})$ need not be defined if $u_j \in \{0,1\}$, some care is needed in the formulation of the following theorem.

Theorem 2.3 (Hybrid copula process)**.** *If Conditions 2.1 and 2.2 hold, then, uniformly in $\boldsymbol{u} \in [0,1]^p$,*

$$\begin{aligned} r_n\{\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\} &= r_n\{\hat{H}_n(\boldsymbol{F}^{\leftarrow}(\boldsymbol{u})) - C(\boldsymbol{u})\} \\ &\quad - \sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u})\, r_n\{\hat{F}_{n,j}(F_j^{\leftarrow}(u_j)) - u_j\}\, \mathbb{1}_{(0,1)}(u_j) + o_p(1), \end{aligned} \tag{2.2}$$

as $n \to \infty$. Hence, in $\ell^\infty([0,1]^p)$ equipped with the supremum norm, as $n \to \infty$,

$$\big(r_n\{\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\}\big)_{\boldsymbol{u}\in[0,1]^p} \rightsquigarrow \Big(\alpha(\boldsymbol{u}) - \textstyle\sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u})\, \beta_j(u_j)\Big)_{\boldsymbol{u}\in[0,1]^p}. \tag{2.3}$$

The right-hand side in (2.3) is well-defined because $\beta_j(0) = \beta_j(1) = 0$ almost surely.

Remark 2.4 (No hybridisation)**.** If, as in the standard situation, $\hat{F}_{n,j}$ is equal to the jth margin of $\hat{H}_n$ for each $j \in \{1, \ldots, p\}$, then, rather than assuming (2.1), it suffices to assume

$$r_n(\hat{H}_n - H) \rightsquigarrow \alpha \circ \boldsymbol{F}, \qquad n \to \infty, \tag{2.4}$$

in $\ell^\infty(\mathbb{R}^p)$, where α is a random element in $\ell^\infty([0,1]^p)$ with continuous trajectories almost surely. Indeed, by the continuous mapping theorem (van der Vaart and Wellner, 1996, Theorem 1.3.6), equation (2.4) implies equation (2.1) with

$$\beta_j(u_j) = \alpha(1, \ldots, 1, u_j, 1, \ldots, 1), \qquad u_j \in [0,1],$$

with u_j appearing at the jth coordinate.

Remark 2.5 (Empirical process representation)**.** Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be an independent random sample from H. For $f \in L^2(H)$, put

$$\mathbb{G}_n f = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n f(\boldsymbol{X}_i) - \mathrm{E}\left[f(\boldsymbol{X}_1)\right] \right).$$

Assume there exists functions $f_{\boldsymbol{x}}$ and $f_{x,j}$ in $L^2(H)$ satisfying the following assumptions:

- We have, as $n \to \infty$,

$$\sup_{\boldsymbol{x} \in \mathbb{R}^p} \left| \sqrt{n}\{\hat{H}_n(\boldsymbol{x}) - H(\boldsymbol{x})\} - \mathbb{G}_n f_{\boldsymbol{x}} \right| = o_p(1),$$
$$\sup_{x \in \mathbb{R}} \left| \sqrt{n}\{\hat{F}_{n,j}(x) - F_j(x)\} - \mathbb{G}_n f_{x,j} \right| = o_p(1), \qquad j \in \{1, \ldots, p\}.$$

- We have $f_{\boldsymbol{x}} = f_{\boldsymbol{x}'}$ in $L^2(H)$ as soon as $F_j(x_j) = F_j(x'_j)$ for all $j \in \{1, \ldots, p\}$; similarly $f_{x,j} = f_{x',j}$ in $L^2(H)$ as soon as $F_j(x) = F_j(x')$.

- The collection

$$\mathcal{F} = \{f_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^p\} \cup \{f_{x,j} : x \in \mathbb{R}, 1 \leqslant j \leqslant p\}$$

  is H-Donsker, i.e., $\mathbb{G}_n \rightsquigarrow \mathbb{G}$ as $n \to \infty$ in the space $\ell^\infty(\mathcal{F})$. The limit $\mathbb{G}$ is a tight, centered Gaussian process with covariance function

$$\operatorname{cov}\left[\mathbb{G}f, \mathbb{G}g\right] = \operatorname{cov}\left[f(\boldsymbol{X}_1), g(\boldsymbol{X}_1)\right], \qquad f, g \in \mathcal{F}. \tag{2.5}$$

Then Condition 2.2 is fulfilled with

$$\alpha(\boldsymbol{u}) = \mathbb{G} f_{\boldsymbol{x}(\boldsymbol{u})}, \qquad \boldsymbol{x}(\boldsymbol{u}) = (F_1^{\leftarrow}(u_1), \ldots, F_p^{\leftarrow}(u_p)),$$
$$\beta_j(u) = \mathbb{G} f_{x_j(u)}, \qquad x_j(u) = F_j^{\leftarrow}(u).$$

It follows that, as $n \to \infty$,

$$\sqrt{n}(\hat{C}_n - C) \rightsquigarrow \left(\mathbb{G} f_{\boldsymbol{x}(\boldsymbol{u})} - \textstyle\sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, \mathbb{G} f_{x_j(u_j)} \right)_{\boldsymbol{u} \in [0,1]^p}.$$

For each $\boldsymbol{u}$, the right-hand side is a zero-mean Gaussian random variable whose variance can be computed via (2.5), yielding

$$\operatorname{var}\left[\mathbb{G} f_{\boldsymbol{x}(\boldsymbol{u})} - \textstyle\sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, \mathbb{G} f_{x_j(u_j)} \right]$$
$$= \operatorname{var}\left[f_{\boldsymbol{x}(\boldsymbol{u})}(\boldsymbol{X}_1) - \textstyle\sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, f_{x_j(u_j)}(\boldsymbol{X}_1) \right], \qquad \boldsymbol{u} \in [0,1]^p.$$

For the usual empirical distribution functions, the above assumptions are fulfilled with $f_{\boldsymbol{x}} = \mathbb{1}_{(-\boldsymbol{\infty}, \boldsymbol{x}]}$ and $f_{x,j} = \mathbb{1}_{\{\boldsymbol{y} : y_j \leqslant x\}}$. The conclusion of Theorem 2.3 then leads to the familiar asymptotics for the empirical copula process (Examples 3.1 and 3.2).

3 Special cases and examples

Example 3.1 (Empirical copula I)**.** Let $\boldsymbol{X}_i = (X_{i,1}, \ldots, X_{i,p})$, for $i \in \{1, \ldots, n\}$, be an independent random sample from H. Let $\hat{H}_n$ and $\hat{F}_{n,j}$ be the joint and marginal empirical distribution functions:

$$\hat{H}_n(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\boldsymbol{X}_i \leqslant \boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^p,$$

$$\hat{F}_{n,j}(x_j) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_{i,j} \leqslant x_j), \qquad x_j \in \mathbb{R}.$$

The hybrid copula estimator $\hat{C}_n$ is then equal to the Deheuvels–Rüschendorf empirical copula (Rüschendorf, 1976; Deheuvels, 1979). By classical empirical process theory (see Remark 2.5), Condition 2.2 is satisfied with $r_n = \sqrt{n}$ and α a C-Brownian bridge and $\beta_j(u_j) = \alpha(1, \ldots, 1, u_j, 1, \ldots, 1)$. Theorem 2.3 then just confirms the weak convergence of the empirical copula process (Stute, 1984; Fermanian et al., 2004; Tsukahara, 2005; van der Vaart and Wellner, 2007; Segers, 2012).

Example 3.2 (Empirical copula II)**.** Let the random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ form a stretch of a stationary time series. By Remark 2.4, the argument in Example 3.1 remains valid provided weak convergence (2.4) of the multivariate empirical process holds. The latter is typically true for weakly dependent, strictly stationary time series, in which case α is a centered Gaussian process whose covariance structure also depends on the serial dependence structure of the underlying time series (Rio, 2000; Doukhan et al., 2009; Dehling and Durieu, 2011; Bücher and Volgushev, 2013).

Example 3.3 (Known margins)**.** In the hypothetical situation that the margins are known, one may just set $\hat{F}_{n,j} = F_j$ for every $j \in \{1, \ldots, p\}$. Remark 2.5 applies with $f_{\boldsymbol{x}} = \mathbb{1}_{(-\boldsymbol{\infty}, \boldsymbol{x}]}$ and $f_{x,j} = 0$. The hybrid copula estimator $\hat{C}_n$ is then equal to the empirical distribution function of the vectors of uniform random variables $(F_1(X_{i,1}), \ldots, F_p(X_{i,p}))$, $i = 1, \ldots, n$. The conclusion is the well-known fact that $\sqrt{n}(\hat{C}_n - C)$ converges to a C-Brownian bridge.

In Genest and Segers (2010), this 'ideal' hybrid copula estimator was compared to the usual empirical copula. Surprisingly, it was concluded that for many copulas, the empirical copula actually has the lower asymptotic variance.

Example 3.4 (Margins modelled parametrically)**.** Assume that the jth margin is modelled by a parametric family $(F_j(\,\cdot\,; \theta_j) : \theta_j \in \Theta_j)$, where Θ_j is an open subset of d_j-dimensional Euclidean space. Then one may estimate F_j parametrically rather than by the marginal empirical distribution function.

Specifically, let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a random sample from H. Let $\hat{\theta}_{n,j}$ be an estimator of θ_j. Estimate F_j by plugging in the estimator for θ_j:

$$\hat{F}_{n,j}(x_j) = F_j(x_j; \hat{\theta}_{n,j}), \qquad x_j \in \mathbb{R}.$$

To estimate the joint distribution, take for instance the empirical distribution function

$$\hat{H}_n(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\boldsymbol{X}_i \leqslant \boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^p.$$

Combining $\hat{H}_n$ and $\hat{F}_{n,j}$ yields the hybrid copula estimator

$$\hat{C}_n(\boldsymbol{u}) = \hat{H}_n\big(F_1^{\leftarrow}(u_1; \hat{\theta}_{n,1}), \dots, F_p^{\leftarrow}(u_p; \hat{\theta}_{n,p})\big), \qquad \boldsymbol{u} \in [0,1]^p,$$

containing both parametric and nonparametric components.

To apply Theorem 2.3, we must check Condition 2.2. In particular, we need to establish an asymptotic representation for $\hat{F}_{n,j}(x_j)$. Required are some basic smoothness assumption on the parametrization $\theta_j \mapsto F_j(\,\cdot\,; \theta_j)$ together with a central limit theorem for $\hat{\theta}_j$. Specifically, assume the following:

(i) The map $\Theta_j \to \ell^\infty(\mathbb{R}) : \theta_j \mapsto F_j(\,\cdot\,; \theta_j)$ is differentiable in the sense that

$$\sup_{x_j \in \mathbb{R}} \Big| F_j(x_j; \theta_j + h) - F_j(x_j; \theta_j) - \textstyle\sum_{k=1}^{d_j} h_k \, \dot{F}_{j,k}(x_j; \theta_j) \Big| = o(|h|), \qquad |h| \to 0, \quad (3.1)$$

where $|h|$ is the Euclidean norm of $h \in \mathbb{R}^{d_j}$ and where $\dot{F}_{j,k}(\,\cdot\,; \theta_j) \in \ell^\infty(\mathbb{R})$ is continuous and depends on x_j only through $F_j(x_j; \theta_j)$.

To establish (3.1), check that the partial derivatives of $F_j(x_j; \theta_j)$ with respect to the components of θ_j exist and are continuous and bounded on compact subsets of $[-\infty, +\infty] \times \Theta_j$.

(ii) The estimator $\hat{\theta}_{n,j}$ admits a linear expansion with influence function $\psi_j = (\psi_{j,1}, \dots, \psi_{j,d_j})$, i.e.,

$$\sqrt{n}(\hat{\theta}_{n,j} - \theta_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_j(\boldsymbol{X}_i) + o_p(1), \qquad n \to \infty.$$

Moreover, $\mathrm{E}\,[\psi_{j,k}(\boldsymbol{X}_1)] = 0$ and $\mathrm{E}\left[\psi_{j,k}^2(\boldsymbol{X}_1)\right] < \infty$ for every component $k \in \{1, \dots, d_j\}$.

The influence function ψ_j may and in general will depend on the unknown value of θ_j. Often, $\psi_j(\boldsymbol{x})$ will be a function of $\boldsymbol{x}$ only through x_j, but this is not required.

By the functional delta method (van der Vaart and Wellner, 1996, Theorem 3.9.4), Assumptions (i) and (ii) imply that, as $n \to \infty$,

$$\begin{aligned}
\sqrt{n}\{F_j(\,\cdot\,; \hat{\theta}_{n,j}) - F_j(\,\cdot\,; \theta_j)\} &= \sum_{k=1}^{d_j} \sqrt{n}(\hat{\theta}_{n,j,k} - \theta_{j,k}) \, \dot{F}_{j,k}(\,\cdot\,; \theta_j) + o_p(1) \\
&= \sum_{k=1}^{d_j} \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_{j,k}(\boldsymbol{X}_i) \, \dot{F}_{j,k}(\,\cdot\,; \theta_j) + o_p(1),
\end{aligned}$$

the $o_p(1)$ terms referring to remainder terms that converge weakly to zero in the space $\ell^\infty(\mathbb{R})$. Remark 2.5 applies with $f_{\boldsymbol{x}} = \mathbb{1}_{(-\boldsymbol{\infty}, \boldsymbol{x}]}$ and

$$f_{x,j}(\,\cdot\,) = \sum_{k=1}^{d_j} \psi_{j,k}(\,\cdot\,) \, \dot{F}_{j,k}(x; \theta_j).$$

We obtain (2.4) with

$$\beta_j(u_j) = \sum_{k=1}^{d_j} \dot{F}_{j,k}(F_j^{\leftarrow}(u_j;\theta_j);\theta_j)\, \mathbb{G}\psi_{j,k}.$$

In view of the conclusion at the end of Example 3.3, it is not certain that the hybrid copula estimator performs better than the empirical copula: bringing in the parametric models for the margins in this way is not necessarily helpful. As the above analysis shows, both the parametric models for the margins and the parameter estimators play a role.

Example 3.5 (Missing data)**.** To show the use of the hybrid copula estimator if some data are missing, consider the following bivariate set-up. Given is an $n \times 2$ data matrix, in each row of which one or both entries may be missing. Formally, the observations consist of a sample of independent, identically distributed quadruples

$$(I_i, J_i, I_i X_i, J_i Y_i), \qquad i \in \{1, \ldots, n\}.$$

The indicator variable I_i (J_i) is equal to 1 or 0 according to whether X_i (Y_i) is observed or not. The pairs (I_i, J_i) and (X_i, Y_i) are supposed to be independent, i.e., the data are missing completely at random. The indicators I_i and J_i may be dependent, and the probabilities of observing a data-row partially or completely are $\mathrm{P}[I_i = 1] = p_X > 0$, $\mathrm{P}[J_i = 1] = p_Y > 0$, and $\mathrm{P}[I_i = J_i = 1] = p_{XY} > 0$. The estimation target is the copula, C, of the bivariate distribution, H, of the pairs (X_i, Y_i). The margins, F and G, of H are assumed to be continuous and Condition 2.1 is assumed to hold.

The marginal and joint distribution functions may be estimated using the data-rows for which the relevant information is available. For $(x, y) \in \mathbb{R}^2$, put

$$\hat{F}_n(x) = \frac{\sum_{i=1}^n \mathbb{1}(X_i \leqslant x, I_i = 1)}{\sum_{i=1}^n \mathbb{1}(I_i = 1)},$$
$$\hat{G}_n(y) = \frac{\sum_{i=1}^n \mathbb{1}(Y_i \leqslant y, J_i = 1)}{\sum_{i=1}^n \mathbb{1}(J_i = 1)},$$
$$\hat{H}_n(x, y) = \frac{\sum_{i=1}^n \mathbb{1}(X_i \leqslant x, Y_i \leqslant y, I_i = J_i = 1)}{\sum_{i=1}^n \mathbb{1}(I_i = J_i = 1)}.$$

Condition 2.2 can be verified by embedding the previous estimators in a certain empirical process. The resulting formulas resemble those for the classical empirical copula process, but now the asymptotic variances and covariances are to be multiplied by (the reciprocals of) the observation probabilities p_X, p_Y and p_{XY}. Details are given at the end of Section 4.

4 Proofs

The proof of Theorem 2.3 is based on a differentiability property of the map that sends a distribution function to its inverse function. In contrast to Lemma 3.9.20 in van der Vaart and Wellner (1996), Lemma 4.1 below does not require the distribution function F to have a density; F need not even be strictly increasing between the two endpoints of its support.

Lemma 4.1. *Let $F_n, F : \mathbb{R} \to [0,1]$ be cumulative distribution functions. Assume that F is continuous and assume that there exists a sequence $0 < r_n \to \infty$ and a continuous function $\beta : [0,1] \to \mathbb{R}$ such that*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |r_n\{F_n(x) - F(x)\} - \beta(F(x))| = 0. \tag{4.1}$$

Then $\beta(0) = \beta(1) = 0$ and

$$\lim_{n\to\infty} \sup_{u\in[0,1]} |r_n\{F(F_n^{\leftarrow}(u)) - u\} + \beta(u)| = 0. \tag{4.2}$$

In particular,

$$\lim_{n\to\infty} \sup_{u\in[0,1]} |r_n\{F(F_n^{\leftarrow}(u)) - u\} + r_n\{F_n(F^{\leftarrow}(u)) - u\}| = 0. \tag{4.3}$$

An abstract way of stating (4.2) is that the map sending a cumulative distribution function G on $\mathbb{R}$ to the distribution function $F \circ G^{\leftarrow}$ on $[0,1]$ is Hadamard differentiable at F tangentially to all functions of the form $\beta \circ F$ for some continuous function $\beta : [0,1] \to \mathbb{R}$, the derivative being given by the map $\beta \circ F \mapsto -\beta$.

Proof. First, note that $\beta(0) = \beta(1) = 0$. Indeed, since $F_n(x) - F(x) \to 0$ as $x \to -\infty$ for each fixed n, we can find a sequence $x_n \to -\infty$ sufficiently fast such that $r_n\{F_n(x_n) - F(x_n)\} \to 0$ as $n \to \infty$ and thus

$$\beta(0) = \lim_{n\to\infty} \beta(F(x_n)) = \lim_{n\to\infty} r_n\{F_n(x_n) - F(x_n)\} = 0$$

by uniform convergence. Similarly $\beta(1) = 0$.

It follows that in (4.2), we can restrict the range in the supremum to $u \in (0,1]$, since $F(F_n^{\leftarrow}(0)) = F(-\infty) = 0$. [However, $F(F_n^{\leftarrow}(1))$ could be smaller than 1.] Write

$$\gamma_n = r_n(F_n - F), \qquad \gamma = \beta \circ F,$$

and note that $F_n = F + r_n^{-1}\gamma_n$. On the one hand, for every $u \in (0,1]$,

$$u \leqslant F_n(F_n^{\leftarrow}(u)) = F(F_n^{\leftarrow}(u)) + r_n^{-1}\gamma_n(F_n^{\leftarrow}(u)),$$

and thus

$$r_n\{F(F_n^{\leftarrow}(u)) - u\} \geqslant -\gamma_n(F_n^{\leftarrow}(u)).$$

On the other hand, for every $u \in (0,1]$ and every $\delta > 0$, we have

$$\begin{aligned} u &> F_n(F_n^{\leftarrow}(u) - \delta) \\ &= F(F_n^{\leftarrow}(u) - \delta) + r_n^{-1}\gamma_n(F_n^{\leftarrow}(u) - \delta) \\ &= F(F_n^{\leftarrow}(u)) + F(F_n^{\leftarrow}(u) - \delta) - F(F_n^{\leftarrow}(u)) + r_n^{-1}\gamma_n(F_n^{\leftarrow}(u) - \delta), \end{aligned}$$

and thus

$$r_n\{F(F_n^{\leftarrow}(u)) - u\} < -\gamma_n(F_n^{\leftarrow}(u) - \delta) + r_n\{F(F_n^{\leftarrow}(u)) - F(F_n^{\leftarrow}(u) - \delta)\}.$$

Since the latter inequality is true for every $\delta > 0$, we can take the limit as $\delta \to 0$. As F is continuous, we obtain

$$r_n\{F(F_n^{\leftarrow}(u)) - u\} \leqslant -\gamma_n(F_n^{\leftarrow}(u)-)$$

where $\gamma_n(x-)$ is the left-hand limit of γ_n at x, a limit which must exist since γ_n is the rescaled difference of two cumulative distribution functions. In combination, we find

$$-\gamma_n(F_n^{\leftarrow}(u)) \leqslant r_n\{F(F_n^{\leftarrow}(u)) - u\} \leqslant -\gamma_n(F_n^{\leftarrow}(u)-), \qquad u \in (0,1]. \tag{4.4}$$

The difference between the left-hand and right-hand sides converges uniformly to zero: indeed, since the sequence γ_n converges uniformly to the continuous function γ, we have

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |\gamma_n(x-) - \gamma_n(x)| = 0.$$

To show (4.2), it then suffices to show that

$$\lim_{n\to\infty} \sup_{u\in(0,1]} |-\gamma_n(F_n^{\leftarrow}(u)) + \beta(u)| = 0.$$

By the triangle inequality and since $\gamma = \beta \circ F$,

$$\begin{aligned}
&|-\gamma_n(F_n^{\leftarrow}(u)) + \beta(u)| \\
&\qquad \leqslant |-\gamma_n(F_n^{\leftarrow}(u)) + \gamma(F_n^{\leftarrow}(u))| + |-\beta(F(F_n^{\leftarrow}(u))) + \beta(u)| \,.
\end{aligned}$$

The first term on the right-hand side converges to zero uniformly in $u \in (0,1]$ by uniform convergence of γ_n to γ on $\mathbb{R}$. By uniform continuity of β on $[0,1]$, the second term on the right-hand side will converge to zero uniformly in $u \in [0,1]$ if we can show that

$$\lim_{n\to\infty} \sup_{u\in[0,1]} |F(F_n^{\leftarrow}(u)) - u| = 0.$$

But the latter equation is a consequence of (4.4), uniform convergence of γ_n to the bounded function γ, and the fact that $r_n \to \infty$ as $n \to \infty$.

Finally, (4.3) follows from by choosing $x = F^{\leftarrow}(u)$ in (4.1), yielding

$$\lim_{n\to\infty} \sup_{u\in[0,1]} |r_n\{F_n(F^{\leftarrow}(u)) - u\} - \beta(u)| = 0$$

[note that $F(F^{\leftarrow}(u)) = u$ by continuity of F] and then using (4.2) and the triangle inequality. □

Lemma 4.2. *Let $F : \mathbb{R} \to [0,1]$ be a continuous cumulative distribution function. Let $0 < r_n \to \infty$ and let $\hat{F}_n$ be a sequence of random cumulative distribution functions such that, in $\ell^\infty(\mathbb{R})$,*

$$r_n(\hat{F}_n - F) \rightsquigarrow \beta \circ F, \qquad n \to \infty, \tag{4.5}$$

where β is a random element of $\ell^\infty([0,1])$ with continuous trajectories. Then $\beta(0) = \beta(1) = 0$ almost surely and

$$\sup_{u\in[0,1]} \Big| r_n\{F(\hat{F}_n^{\leftarrow}(u)) - u\} + r_n\{\hat{F}_n(F^{\leftarrow}(u)) - u\} \Big| = o_p(1). \tag{4.6}$$

As a consequence, in $\ell^\infty([0,1])$,

$$\big(r_n\{F(\hat{F}_n^{\leftarrow}(u)) - u\}\big)_{u\in[0,1]} \rightsquigarrow -\beta, \qquad n \to \infty. \tag{4.7}$$

Proof. First, we show that $\beta(0) = \beta(1) = 0$ almost surely. Define the map $g : \ell^\infty(\mathbb{R}) \to \mathbb{R}$ by $g(\gamma) = \inf_{M>0} \sup_{x:|x|\geqslant M} |\gamma(x)| = \limsup_{|x|\to\infty} |\gamma(x)|$. The map g is continuous with respect to the supremum distance. As $\hat{F}_n$ and F are cumulative distribution functions, $g(r_n(\hat{F}_n - F)) = 0$ almost surely. By weak convergence (4.5) and the continuous mapping theorem (van der Vaart and Wellner, 1996, Theorem 1.3.6), it follows that $g(\beta \circ F) = \max\{|\beta(0)|, |\beta(1)|\} = 0$ almost surely too.

Equation (4.7) follows from combining (4.5) and (4.6); use the triangle inequality and the fact that $u = F(F^{\leftarrow}(u))$.

We will show equation (4.6) by an application of the extended continuous mapping theorem (van der Vaart and Wellner, 1996, Theorem 1.11.1).

Let $\mathbb{D}_n$ be the collection of all functions $\gamma \in \ell^\infty(\mathbb{R})$ such that $F + r_n^{-1}\gamma$ is a cumulative distribution function. In particular, $\gamma(\pm\infty) = \lim_{x\to\pm\infty} \gamma(x) = 0$. Define the map $g_n : \mathbb{D}_n \to \ell^\infty([0,1])$ by

$$(g_n(\gamma))(u) = r_n\{F((F + r_n^{-1}\gamma)^{\leftarrow}(u)) - u\} + \gamma(F^{\leftarrow}(u)).$$

Let $\gamma_n \in \mathbb{D}_n$ be such that $\gamma_n \to \delta \circ F$ in $\ell^\infty(\mathbb{R})$, where $\delta : [0,1] \to \mathbb{R}$ is continuous. Put $F_n = F + r_n^{-1}\gamma_n$. Then $\gamma_n = r_n(F_n - F)$ and the conditions of Lemma 4.1 are fulfilled. It follows that, in $\ell^\infty(\mathbb{R})$,

$$g_n(\gamma_n) = r_n(F \circ F_n^{\leftarrow} - \mathrm{id}_{[0,1]}) + r_n(F_n \circ F^{\leftarrow} - \mathrm{id}_{[0,1]}) \to 0, \qquad n \to \infty,$$

where 'id' refers to the identity mapping. By construction, the maps $\hat{\gamma}_n = r_n(\hat{F}_n - F)$ take values in $\mathbb{D}_n$. Given the assumption (4.5) and the previous limit relation, we can then apply the extended continuous mapping theorem. We find that, in $\ell^\infty([0,1])$,

$$g_n(r_n(\hat{F}_n - F)) \rightsquigarrow 0, \qquad n \to \infty.$$

But this is precisely (4.6). □

Lemma 4.3. *Let C be a p-variate copula satisfying Condition 2.1(b). Let $0 < r_n \to \infty$ and, for each $n \in \mathbb{N}$ and $j \in \{1, \ldots, p\}$, let $\beta_{n,j} \in \ell^\infty([0,1])$ be such that $0 \leqslant u + r_n\beta_{n,j}(u) \leqslant 1$ for all $u \in [0,1]$. If, for each $j \in \{1, \ldots, p\}$, we have $\beta_{n,j} \to \beta_j$ in $\ell^\infty([0,1])$ and if β_j is continuous and $\beta_j(0) = \beta_j(1) = 0$, then, uniformly in $\boldsymbol{u} \in [0,1]^p$,*

$$\begin{aligned} r_n\{C(u_1 + r_n^{-1}\beta_{n,1}(u_1), \ldots, u_p + r_n^{-1}\beta_{n,p}(u_p)) - C(\boldsymbol{u})\} \\ = \sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u})\, \beta_{n,j}(u_j)\, \mathbb{1}_{(0,1)}(u_j) + o(1), \qquad n \to \infty. \end{aligned} \tag{4.8}$$

Observe that $\dot{C}_j(\boldsymbol{u})$ is not defined if $u_j \in \{0,1\}$. This is the reason for including the indicator $\mathbb{1}_{(0,1)}(u_j)$ on the right-hand side of (4.8).

Proof. For convenience, write

$$\boldsymbol{\beta}_n(\boldsymbol{u}) = \big(\beta_{n,1}(u_1), \ldots, \beta_{n,p}(u_p)\big), \qquad \boldsymbol{u} \in [0,1]^p.$$

Fix $\boldsymbol{u} \in [0,1]^p$ and $n \in \mathbb{N}$. Define $f : [0,1] \to \mathbb{R}$ by

$$f(x) = C(\boldsymbol{u} + x\, r_n^{-1}\, \boldsymbol{\beta}_n(\boldsymbol{u})).$$

The function f is continuous on $[0,1]$ and continuously differentiable on $(0,1)$. Indeed, if $\beta_{n,j}(u_j) \neq 0$, then u_j and $u_j + r_n^{-1}\beta_{n,j}(u_j)$ are two different points in $[0,1]$, and thus

$$\beta_{n,j}(u_j) \neq 0 \quad \Longrightarrow \quad \forall x \in (0,1) : 0 < u_j + x\, r_n^{-1}\, \beta_{n,j}(u_j) < 1. \tag{4.9}$$

The derivative of f is

$$f'(x) = \sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u} + x\, r_n^{-1}\, \boldsymbol{\beta}_n(\boldsymbol{u}))\, r_n^{-1}\, \beta_{n,j}(u_j), \qquad x \in (0,1). \tag{4.10}$$

Because of (4.9), the right-hand side of (4.10) is well-defined. By the mean value theorem, there exists $x_n(\boldsymbol{u}) \in (0,1)$ such that

$$\begin{aligned} r_n\{C(\boldsymbol{u} + r_n^{-1}\boldsymbol{\beta}_n(\boldsymbol{u})) - C(\boldsymbol{u})\} &= r_n\{f(1) - f(0)\} = r_n\, f'(x_n(\boldsymbol{u})) \\ &= \sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u} + x_n(\boldsymbol{u})\, r_n^{-1}\, \boldsymbol{\beta}_n(\boldsymbol{u}))\, \beta_{n,j}(u_j). \end{aligned}$$

By the triangle inequality,

$$\left| r_n\{C(\boldsymbol{u} + r_n^{-1}\boldsymbol{\beta}_n(\boldsymbol{u})) - C(\boldsymbol{u})\} - \sum_{j=1}^{p} \dot{C}_j(\boldsymbol{u})\, \beta_{n,j}(u_j)\, \mathbb{1}_{(0,1)}(u_j) \right| \leqslant \sum_{j=1}^{p} \Delta_{n,j}(\boldsymbol{u})$$

where, for $\boldsymbol{u} \in [0,1]^p$,

$$\Delta_{n,j}(\boldsymbol{u}) = \left| \dot{C}_j(\boldsymbol{u} + x_n(\boldsymbol{u})\, r_n^{-1}\, \boldsymbol{\beta}_n(\boldsymbol{u})) - \dot{C}_j(\boldsymbol{u})\, \mathbb{1}_{(0,1)}(u_j) \right| \, |\beta_{n,j}(u_j)| \, .$$

Fix $j \in \{1, \ldots, p\}$. We need to show that $\lim_{n\to\infty} \|\Delta_{n,j}\|_\infty = 0$. Since $\lim_{n\to\infty} \beta_{n,j} = \beta_j$ in $\ell^\infty([0,1])$, we have $\sup_{n\in\mathbb{N}} \|\beta_{n,j}\|_\infty = M < \infty$. Fix $\varepsilon > 0$. As $\beta_j(0) = \beta_j(1) = 0$ and β_j is continuous, there exists $n(\varepsilon) \in \mathbb{N}$ and $\delta(\varepsilon) \in (0, 1/2)$ such that

$$\sup\{|\beta_{n,j}(u_j)| : n \geqslant n(\varepsilon),\ u_j \in [0, \delta(\varepsilon)] \cup [1 - \delta(\varepsilon), 1]\} \leqslant \varepsilon.$$

By increasing $n(\varepsilon)$ if necessary, we can also ensure that $M/r_n \leqslant \delta(\varepsilon)/2$ for all $n \geqslant n(\varepsilon)$. Split the supremum of $\Delta_{n,j}(\boldsymbol{u})$ over $\boldsymbol{u} \in [0,1]^p$ into two parts, according to whether $u_j \in [\delta(\varepsilon), 1 - \delta(\varepsilon)]$ or not. Write $V_j(\delta) = \{\boldsymbol{u} \in [0,1]^p : \delta \leqslant u_j \leqslant 1 - \delta\}$.

- On the one hand, writing $|\boldsymbol{w}|_\infty = \max\{|w_1|, \ldots, |w_p|\}$ for $\boldsymbol{w} \in \mathbb{R}^p$,

  $$\sup_{\boldsymbol{u} \in V_j(\delta(\varepsilon))} |\Delta_{n,j}(\boldsymbol{u})| \leqslant M \sup_{\substack{\boldsymbol{u},\boldsymbol{v} \in V_j(\delta(\varepsilon)/2) \\ |\boldsymbol{u}-\boldsymbol{v}|_\infty \leqslant M/r_n}} \left| \dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{v}) \right| .$$

  By uniform continuity of $\dot{C}_j$ on $V_j(\delta)$ for any $\delta > 0$, the right-hand side converges to zero as $n \to \infty$.

- On the other hand, for $n \geqslant n(\varepsilon)$, since $0 \leqslant \dot{C}_j \leqslant 1$,

  $$\sup_{\boldsymbol{u} \in [0,1]^p \setminus V_j(\delta(\varepsilon))} |\Delta_{n,j}(\boldsymbol{u})| \leqslant \varepsilon.$$

It follows that $\limsup_{n\to\infty} \|\Delta_{n,j}\|_\infty \leqslant \varepsilon$. As $\varepsilon > 0$ was arbitrary, we conclude that $\lim_{n\to\infty} \|\Delta_{n,j}\|_\infty = 0$, as required. □

Proof of Theorem 2.3. By Lemma 4.2, we have, in $\ell^\infty([0,1])$,

$$r_n(F_j \circ \hat{F}_{n,j}^{\leftarrow} - \mathrm{id}_{[0,1]}) \\ = -r_n(\hat{F}_{n,j} \circ F_j^{\leftarrow} - \mathrm{id}_{[0,1]}) + o_p(1) \rightsquigarrow -\beta_j, \qquad n \to \infty. \quad (4.11)$$

Moreover, $\beta_j(0) = \beta_j(1) = 0$ almost surely.

For notational convenience, consider the random vector

$$\hat{\boldsymbol{F}}_n^{\leftarrow}(\boldsymbol{u}) = \big(\hat{F}_{n,j}^{\leftarrow}(u_1), \ldots, \hat{F}_{n,p}^{\leftarrow}(u_p)\big), \qquad \boldsymbol{u} \in [0,1]^p.$$

The following decomposition is fundamental to the analysis of the hybrid copula estimator $\hat{C}_n = \hat{H}_n \circ \hat{\boldsymbol{F}}_n^{\leftarrow}$:

$$r_n(\hat{C}_n - C) = r_n(\hat{H}_n \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - H \circ \hat{\boldsymbol{F}}_n^{\leftarrow}) + r_n(H \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - C). \quad (4.12)$$

We will treat both terms on the right-hand side of (4.12) in turn.

As $H = C \circ \boldsymbol{F}$, the first term on the right-hand side in (4.12) is

$$r_n(\hat{H}_n \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - H \circ \hat{\boldsymbol{F}}_n^{\leftarrow}) = r_n(\hat{H}_n \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - \hat{H}_n \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow}) \\ + r_n(\hat{H}_n \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - C \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow}). \quad (4.13)$$

- The first term on the right-hand of (4.13) is $o_p(1)$ in $\ell^\infty([0,1]^p)$ provided we can show that

  $$r_n(\hat{H}_n - \hat{H}_n \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F}) = o_p(1), \qquad n \to \infty.$$

  But the latter holds in view of the identity $H = H \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F}$ (the margins of H are $F_1, \ldots, F_p$ and these are continuous), Condition 2.2, and the identity $\boldsymbol{F} \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F} = \boldsymbol{F}$.

- Regarding the second term on the right-hand side of (4.13), note that, by (4.11), for every $j \in \{1, \ldots, p\}$,

  $$F_j \circ \hat{F}_{n,j}^{\leftarrow} \rightsquigarrow \mathrm{id}_{[0,1]}, \qquad n \to \infty,$$

  in $\ell^\infty([0,1])$. Moreover, by Condition 2.2 and the identities $C = H \circ \boldsymbol{F}^{\leftarrow}$ and $\boldsymbol{F} \circ \boldsymbol{F}^{\leftarrow} = \mathrm{id}_{[0,1]^p}$, we have, in $\ell^\infty([0,1]^p)$,

  $$r_n(\hat{H}_n \circ \boldsymbol{F}^{\leftarrow} - C) \rightsquigarrow \alpha \circ \boldsymbol{F} \circ \boldsymbol{F}^{\leftarrow} = \alpha, \qquad n \to \infty.$$

  By asymptotic uniform equicontinuity (van der Vaart and Wellner, 1996, Theorem 1.5.7 and Addendum 1.5.8), as $n \to \infty$,

  $$r_n(\hat{H}_n \circ \boldsymbol{F}^{\leftarrow} \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - C \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow}) = r_n(\hat{H}_n \circ \boldsymbol{F}^{\leftarrow} - C) + o_p(1)$$

We find that, in $\ell^\infty([0,1]^p)$,

$$r_n(\hat{H}_n \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - H \circ \hat{\boldsymbol{F}}_n^{\leftarrow}) = r_n(\hat{H}_n \circ \boldsymbol{F}^{\leftarrow} - C) + o_p(1) \rightsquigarrow \alpha, \qquad n \to \infty. \quad (4.14)$$

The second term on the right-hand side in (4.12) is $r_n(C \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow} - C)$. For $n \in \mathbb{N}$, let $\mathbb{D}_n$ be the collection of p-tuples $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p) \in \ell^\infty(\mathbb{R}) \otimes \cdots \otimes \ell^\infty(\mathbb{R})$ such that map $x \mapsto F_j(x) + r_n^{-1}\gamma_j(x)$ is a cumulative distribution function for each $j \in \{1, \ldots, p\}$. Define the map $g_n : \mathbb{D}_n \to \ell^\infty([0,1]^p)$ by

$$(g_n(\boldsymbol{\gamma}))(\boldsymbol{u}) = r_n\{C \circ \boldsymbol{F} \circ (\boldsymbol{F} + r_n^{-1}\boldsymbol{\gamma})^{\leftarrow}(\boldsymbol{u}) - C(\boldsymbol{u})\}$$
$$- \sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, r_n\{F_j \circ (F_j + r_n^{-1}\gamma_j)^{\leftarrow}(u_j) - u_j\}\, \mathbb{1}_{(0,1)}(u_j), \qquad \boldsymbol{u} \in [0,1]^p.$$

Let $\boldsymbol{\gamma}_n \in \mathbb{D}_n$ be such that $\lim_{n\to\infty} \boldsymbol{\gamma}_n = \boldsymbol{\gamma}$ where $\gamma_j = \beta_j \circ F$ and $\beta_j \in \ell^\infty([0,1])$ is continuous and satisfies $\beta_j(0) = \beta_j(1) = 1$ for every $j \in \{1, \ldots, p\}$. By Lemma 4.1 with $F_{n,j} = F_j + r_n^{-1}\gamma_{n,j}$, we then have

$$r_n\{F_j \circ (F_j + r_n^{-1}\gamma_j)^{\leftarrow} - \mathrm{id}_{[0,1]}\}\, \mathbb{1}_{(0,1)} \to -\beta_j, \qquad n \to \infty.$$

By Lemma 4.3, it then follows that

$$g_n(\boldsymbol{\gamma}_n) \to 0, \qquad n \to \infty.$$

By the extended continuous mapping theorem (van der Vaart and Wellner, 1996, Theorem 1.11.1), it follows that

$$g_n(r_n(\hat{\boldsymbol{F}}_n - \boldsymbol{F})) \rightsquigarrow 0, \qquad n \to \infty.$$

But this says exactly that, uniformly in $\boldsymbol{u} \in [0,1]^p$, as $n \to \infty$,

$$r_n\{C \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow}(\boldsymbol{u}) - C(\boldsymbol{u})\}$$
$$= \sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, r_n\{F_j \circ \hat{F}_{n,j}^{\leftarrow}(u_j) - u_j\}\, \mathbb{1}_{(0,1)}(u_j) + o_p(1).$$

Insert (4.11) to deduce that, uniformly in $\boldsymbol{u} \in [0,1]^p$ and as $n \to \infty$,

$$r_n\{C \circ \boldsymbol{F} \circ \hat{\boldsymbol{F}}_n^{\leftarrow}(\boldsymbol{u}) - C(\boldsymbol{u})\}$$
$$= -\sum_{j=1}^p \dot{C}_j(\boldsymbol{u})\, r_n\{\hat{F}_{n,j} \circ F_j^{\leftarrow}(u_j) - u_j\}\, \mathbb{1}_{(0,1)}(u_j) + o_p(1). \quad (4.15)$$

Collect the representations in (4.14) and (4.15) of the two terms on the right-hand side of (4.12) and apply Condition 2.2 to arrive at the stated conclusion. □

Details for Example 3.5. Consider the following functions from $\{0,1\}^2 \times \mathbb{R}^2$ into $\mathbb{R}$: for $(x, y) \in \mathbb{R}^2$,

$$f_1(I,J,X,Y) = \mathbb{1}(I=1), \qquad g_{1,x}(I,J,X,Y) = \mathbb{1}(X \leqslant x, I=1),$$
$$f_2(I,J,X,Y) = \mathbb{1}(J=1), \qquad g_{2,y}(I,J,X,Y) = \mathbb{1}(Y \leqslant y, J=1),$$
$$f_3 = f_1 f_2, \qquad g_{3,x,y} = g_{1,x} g_{2,y}.$$

Let P denote the common distribution of the quadruples (I_i, J_i, X_i, Y_i). The collection of functions

$$\mathcal{F} = \{f_1, f_2, f_3\} \cup \{g_{1,x} : x \in \mathbb{R}\} \cup \{g_{2,y} : y \in \mathbb{R}\} \cup \{g_{3,x,y} : (x,y) \in \mathbb{R}^2\}$$

is a finite union of VC-classes and thus P-Donsker (van der Vaart and Wellner, 1996, Chapter 2.6). The empirical process $\mathbb{G}_n$ defined by

$$\mathbb{G}_n(f) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} f(I_i, J_i, X_i, Y_i) - \mathrm{E}\left[f(I_1, J_1, X_1, Y_1)\right] \right), \qquad f \in \mathcal{F},$$

converges in $\ell^\infty(\mathcal{F})$ to a P-Brownian bridge $\mathbb{G}$. For $(x, y) \in \mathbb{R}^2$,

$$\hat{F}_n(x) = \frac{p_X\, F(x) + n^{-1/2}\mathbb{G}_n g_{1,x}}{p_X + n^{-1/2}\mathbb{G}_n f_1},$$
$$\hat{G}_n(y) = \frac{p_Y\, G(y) + n^{-1/2}\mathbb{G}_n g_{2,y}}{p_y + n^{-1/2}\mathbb{G}_n f_2},$$
$$\hat{H}_n(x, y) = \frac{p_{XY}\, H(x, y) + n^{-1/2}\mathbb{G}_n g_{3,x,y}}{p_{XY} + n^{-1/2}\mathbb{G} f_3}.$$

It follows that, as $n \to \infty$ and uniformly in $(x, y) \in \mathbb{R}^2$,

$$\sqrt{n}\{\hat{F}_n(x) - F(x)\} = p_X^{-1}\mathbb{G}_n(g_{1,x} - F(x)\, f_1) + O_p(n^{-1/2}),$$
$$\sqrt{n}\{\hat{G}_n(x) - G(x)\} = p_Y^{-1}\mathbb{G}_n(g_{2,x} - G(x)\, f_2) + O_p(n^{-1/2}),$$
$$\sqrt{n}\{\hat{H}_n(x, y) - H(x, y)\} = p_{XY}^{-1}\mathbb{G}_n(g_{3,x,y} - H(x, y) f_3) + O_p(n^{-1/2}).$$

As a consequence, Condition 2.2 is fulfilled with, for $(u, v) \in [0, 1]^2$,

$$\beta_1(u) = p_X^{-1}\mathbb{G}(g_{1,F^{\leftarrow}(u)} - u\, f_1),$$
$$\beta_2(v) = p_Y^{-1}\mathbb{G}(g_{2,G^{\leftarrow}(v)} - v\, f_2),$$
$$\alpha(u, v) = p_{XY}^{-1}\mathbb{G}(g_{3,F^{\leftarrow}(u),G^{\leftarrow}(v)} - C(u, v)\, f_3).$$

From these formulas, the variances and covariances can be easily computed: for $(u, u_1, u_2, v, v_1, v_2) \in [0, 1]^6$,

$$\mathrm{cov}\left[\beta_1(u_1), \beta_1(u_2)\right] = p_X^{-1}\, \{u_1 \wedge u_2 - u_1 u_2\},$$
$$\mathrm{cov}\left[\beta_2(v_1), \beta_2(v_2)\right] = p_Y^{-1}\, \{v_1 \wedge v_2 - v_1 v_2\},$$
$$\mathrm{cov}\left[\beta_1(u), \beta_2(v)\right] = \frac{p_{XY}}{p_X p_Y}\, \{C(u, v) - uv\},$$

and

$$\mathrm{cov}\left[\alpha(u_1, v_1), \alpha(u_2, v_2)\right] = p_{XY}^{-1}\, \{C(u_1 \wedge u_2, v_1 \wedge v_2) - C(u_1, v_1)\, C(u_2, v_2)\},$$
$$\mathrm{cov}\left[\alpha(u_1, v), \beta_1(u_2)\right] = p_X^{-1}\, \{C(u_1 \wedge u_2, v) - C(u_1, v)\, u_2\},$$
$$\mathrm{cov}\left[\alpha(u, v_1), \beta_2(v_2)\right] = p_Y^{-1}\, \{C(u, v_1 \wedge v_2) - C(u, v_1)\, v_2\}.$$

□

Acknowledgments

Fruitful discussions with Christian Genest (McGill University) are gratefully acknowledged. The research project was funded 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 (Belgian Science Policy).

References

Bücher, A. and S. Volgushev (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis 119*, 61–70.

de Haan, L. and S. I. Resnick (1977). Limit theory for multivariate sample extremes. *Z. Warhsch. Verw. Gebiete 40*, 317–337.

Deheuvels, P. (1978). Caractérisation complète des lois extrêmes multivariées et de la convergence aux types extrêmes. *Publ. Inst. Statist. Univ. Paris 23.*

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. *Bulletin de la Classe des Sciences, Académie Royale de Belgique 65*, 274–292.

Dehling, H. and O. Durieu (2011). Empirical processes of multidimensional systems with multiple mixing properties. *Stochastic Processes and their Applications 121*(5), 1076–1096.

Doukhan, P., J.-D. Fermanian, and G. Lang (2009). An empirical central limit theorem with applications to copulas under weak dependence. *Statistical Inference for Stochastic Processes 12*(1), 65–87.

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

Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics.* New York: Wiley.

Genest, C. and J. Segers (2010). On the covariance of the asymptotic empirical copula process. *Journal of Multivariate Analysis 101*, 1837–1845.

Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, Volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications].* Berlin: Springer-Verlag.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *The Annals of Statistics 4*(5), 912–923.

Segers, J. (2012). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli 18*(3), 764–782.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, 229–231.

Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. *The Annals of Probability 12*, 361–379.

Tsukahara, H. (2005). Semiparametric estimation in copula models. *The Canadian Journal of Statistics 33*, 357–375.

van der Vaart, A. and J. A. Wellner (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*, Volume 55 of *IMS Lecture Notes–Monograph Series*, pp. 234–252.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes.* New York: Springer-Verlag.