Electronic Journal of Statistics
Vol. 16 (2022) 5692–5719
ISSN: 1935-7524
https://doi.org/10.1214/22-EJS2075

Conditional empirical copula processes and generalized measures of association*

Alexis Derumigny

Department of Applied Mathematics,
Delft University of Technology,
Mekelweg 4, 2628 CD, Delft, Netherlands
e-mail: a.f.f.derumigny@tudelft.nl

and

Jean-David Fermanian

Crest-Ensae 5,
avenue Henry Le Chatelier,
91764 Palaiseau cedex, France
e-mail: jean-david.fermanian@ensae.fr

Abstract: We study the weak convergence of conditional empirical copula processes indexed by general families of conditioning events that have non zero probabilities. Moreover, we also study the case where the conditioning events are chosen in a data-driven way. The validity of several bootstrap schemes is stated, including the exchangeable bootstrap. We define general multivariate measures of association, possibly given some fixed or random conditioning events. By applying our theoretical results, we prove the asymptotic normality of the estimators of such measures. We illustrate our results with financial data.

MSC2020 subject classifications: Primary 62G05, 62G30; secondary 62H20, 62G09.
Keywords and phrases: Empirical copula process, conditional copula, weak convergence, bootstrap.

Received February 2022.

Contents

*Jean-David Fermanian's work has been supported by the labex Ecodec (reference project ANR-11-LABEX-0047).

1. Introduction

Since their formal introduction by Patton in [43, 44], conditional copulas have become key tools to describe the dependence function between the components of a random vector $\mathbf{X} := (X_1, \dots, X_p) \in \mathbb{R}^p$, given that another random vector of covariates $\mathbf{Z} := (Z_1, \dots, Z_q) \in \mathbb{R}^q$ is observed. This concept, generalized in [21], may be stated as an extension of Sklar's famous theorem: for every borelian subset $A \subset \mathbb{R}^q$ and every vector $\mathbf{x} \in \mathbb{R}^p$, the conditional joint law of $\mathbf{X}$ given $(\mathbf{Z} \in A)$ is written

$$\begin{aligned} F(\mathbf{x}|A) &:= \mathbb{P}\big(\mathbf{X} \leqslant \mathbf{x} | \mathbf{Z} \in A\big) \\ &= C_{\mathbf{X}|\mathbf{Z}}\big(\mathbb{P}(X_1 \leqslant x_1 | \mathbf{Z} \in A), \dots, \mathbb{P}(X_p \leqslant x_p | \mathbf{Z} \in A) \,\big|\, \mathbf{Z} \in A\big), \end{aligned} \tag{1}$$

for some map $C_{\mathbf{X}|\mathbf{Z}}(\cdot|\mathbf{Z} \in A) : [0,1]^p \to [0,1]$ that is a copula (denoted as $C(\cdot|A)$ hereafter to be short). Note that we have denoted inequalities componentwise. This will be our convention hereafter.

Now, Patton's seminal paper ([43]) has been referenced more than 2 000 times in the academic literature. The concept of conditional copulas (also sometimes called "dynamic copulas" or "time-varying copulas") has been applied in many fields: economics ([37, 46]), financial econometrics ([11, 31, 45]), risk management ([40, 42]), agriculture ([27]), actuarial science ([7, 19] and [12] more recently), hydrology ([28, 33]), etc, among many others. The rise of pair-copula constructions, particularly vine models ([1, 5, 6]) has fuelled the interest around conditional copulas. Indeed, generally speaking, any p-dimensional distribution can be described by $p(p-1)/2$ bivariate conditional copulas and p margins. Even if most vine models assume that such conditional copulas do not depend in fact on their own conditioning variables (the so-called "simplifying assumption"; see [13, 26, 29] and the references therein), no consensus has emerged. Therefore, some recent papers propose some model specification for vines and the associated inference procedures by working directly on conditional copulas: see [36, 52, 59, 60], for instance.

Moreover, the statistical theory of conditional copulas is currently an active research topic. In the literature, the conditioning subset A in (1) is most often pointwise, i.e., the authors consider conditioning subsets $\{\omega \in \Omega : \mathbf{Z}(\omega) = \mathbf{z}\}$ for some particular vector $\mathbf{z} \in \mathbb{R}^q$. In such cases, we will denote $C(\cdot|A)$ as $C(\cdot|\mathbf{Z} = \mathbf{z})$. Typically, in a semi-parametric model, it is assumed that $C_{\mathbf{X}|\mathbf{Z}}(\mathbf{x}|\mathbf{Z} = \mathbf{z}) = C_{\theta(\mathbf{z})}(\mathbf{x})$ for some map $\mathbf{z} \mapsto \theta(\mathbf{z}) \in \mathbb{R}^m$ and the main goal is to statistically estimate the latter link function, as in [2, 3, 4, 64]. Under a nonparametric point-of-view, the main quantity of interest is rather the empirical copula process given $(\mathbf{Z} = \mathbf{z})$. For instance, [25, 47, 65] study the weak convergence of such a process.

To the best of our knowledge, almost all the papers in the literature have focused on pointwise conditioning events until now. In a few papers, some box-type conditioning events as $A := \prod_{k=1}^q (a_k, b_k)$ are considered, where $(a_k, b_k) \in \overline{\mathbb{R}}^2$

for every $k \in \{1, \dots, q\}$. For example, [55], p.1127, discusses a Spearman's rho between two random variables X_1 and X_2, knowing that X_1 and/or X_2 is above (or below) some threshold. Nonetheless, the limiting law of such a quantity is not derived. In the same spirit, [17] estimates similar quantities for measuring contagions between two markets, but they do not yield their asymptotic variances. They wrote that "this variance is usually difficult to get in a closed form and can be estimated by means of a bootstrap procedure". See [18] too. Indeed, the limiting law of such statistics cannot be easily deduced from the asymptotic behavior of the usual empirical copula process, and necessitates particular analysis (see below). The aim of our paper is to state general theoretical results that allow to solve such problems.

Actually, such box-type conditioning events provide a natural framework in many situations. For instance, it is often of interest to measure and monitor conditional measures of association between the components of $\mathbf{X}$ given $\mathbf{Z}$ belongs to some particular areas in $\mathbb{R}^q$, through a model-free approach. Therefore, bank stress tests will focus on the events $\{\omega | Z_k(\omega) > q_k^Z, k \in \{1, \dots, q\}\}$ for some quantiles q_k^Z of Z_k. Since the levels of the latter quantiles are often high, it is no longer possible to rely on marginal or joint estimators given pointwise conditioning events (kernel smoothing, e.g.). This justifies the bucketing of $\mathbf{Z}$ values. Moreover, when dealing with high-dimensional vectors of covariates, discretizing the $\mathbf{Z}$-space is often the only feasible way of measuring conditional dependencies. Indeed, it is no longer possible to invoke usual nonparametric estimators, due to the usual curse of dimensionality. Since measures of association are functions of the underlying copula, the key theoretical object will be here the conditional copula $C(\cdot|A)$ of $\mathbf{X}$ given $(\mathbf{Z} \in A)$ for some borelian subsets A, and some of its nonparametric estimators.

Focusing on set-type conditioning events rather than pointwise conditioning events to study dependencies induces an important change of perspective. The interpretation of some empirical results has to be done with care, because the relationship between the two types of conditional copulas is involved and may be counter-intuitive. To illustrate, consider the very simple linear model $X_1 = \beta_1 Z + \epsilon_1$, and $X_2 = \beta_2 Z + \varepsilon_2$, where ϵ_1, ϵ_2 and Z are mutually independent random variables. Then, X_1 and X_2 are conditionally independent given $(Z = z)$. Nonetheless, X_1 and X_2 are in general dependent given $(Z \in [a, b])$, $(a, b) \in \mathbb{R}^2$. See Remark 1 and Section 3.1. in [13] for a deeper discussion.

The goal of this paper is threefold. First, in Section 2, we state the weak convergence of the conditional empirical copula process indexed by borelian subsets under minimal assumptions, extending [57] written for usual copulas; we also state an analogous result when the conditioning subset is random. Second, we prove the validity of the exchangeable bootstrap scheme for the latter process in Section 3. We show that the usual nonparametric Efron's bootstrap ([20]) can still be applied. Third, Section 4 introduces a family of general "conditional" measures of association as mappings of the latter copulas. This family virtually includes and generalizes all measures that have been introduced until now. We apply our theoretical results to prove their asymptotic normality. It is important to note that our results obviously include the particular case of no covariate/conditioning event. Therefore, we contribute to the literature on usual copulas as much as on conditional copulas. In the Supplementary Material [14], Section A provides an empirical application of conditional copulas with data-driven conditioning events to study conditional dependencies between returns on stock prices.

2. Weak convergence of empirical copula processes

2.1. Empirical copula processes indexed by families of subsets

Let us fix the law of $(\mathbf{X},\mathbf{Z}) \in \mathbb{R}^{p+q}$ and consider a family $\mathcal{A}$ of borelian subsets in $\mathbb{R}^q$ satisfying the following regularity condition.

Condition 1. *The class of subsets $\mathcal{A}$ is Donsker and $p_A := \mathbb{P}(\mathbf{Z} \in A)$ is larger than a constant $\underline{p} > 0$ for every $A \in \mathcal{A}$. Moreover, the conditional margins $F_k(\cdot|\mathbf{Z} \in A)$ are continuous, for any $k \in \{1,\dots,p\}$ and any $A \in \mathcal{A}$.*

Note that this property implicitly depends on the probability distribution of the random vector $\mathbf{Z}$. On the other hand, most standard conditioning subsets (intervals, boxes, e.g.) are universally Donsker and the latter point is always satisfied in this case.

To fix notations, let $\big((\mathbf{X}_1,\mathbf{Z}_1),\dots,(\mathbf{X}_n,\mathbf{Z}_n)\big)$ be an i.i.d. sample of realizations of $(\mathbf{X},\mathbf{Z})$ and denote by A a particular element of $\mathcal{A}$. The conditional copula of $\mathbf{X}$ given the event $(\mathbf{Z} \in A)$, that will simply be denoted by $C(\cdot|A)$, can be estimated by

$$\hat{C}_n(\mathbf{u}|A) := \frac{1}{n\hat{p}_A}\sum_{i=1}^n \mathbf{1}\big(F_{n,1}(X_{i,1}|A) \leqslant u_1,\dots,F_{n,p}(X_{i,p}|A) \leqslant u_p, \mathbf{Z}_i \in A\big), \text{ where}$$

$$F_{n,k}(t|A) := \frac{1}{n\hat{p}_A}\sum_{i=1}^n \mathbf{1}(X_{i,k} \leqslant t, \mathbf{Z}_i \in A), \text{ for } k \in \{1,\dots,p\},$$

$$\hat{p}_A := n^{-1}\sum_{i=1}^n \mathbf{1}(\mathbf{Z}_i \in A) =: \frac{n_A}{n} \simeq p_A.$$

Note that n_A is the size of the sub-sample of the observations $\mathbf{X}_i$ such that $\mathbf{Z}_i \in A$. It is a random integer in $\{0,1,\dots,n\}$. When $n_A = 0$, $\hat{p}_A = 0$ and $F_{n,k}(\cdot|A)$ can be set to any distribution, formally.

Remark 1. *Since $\mathcal{A}$ is a Donsker class, $\sup_{A\in\mathcal{A}} \sqrt{n}|\hat{p}_A - p_A| = O_P(1)$ and then $\inf_{A\in\mathcal{A}} \hat{p}_A \geqslant \underline{p} + o_P(1)$. Therefore, with a probability that tends to 1, $\inf_{A\in\mathcal{A}} n_A > 0$ and we will consider in the following that this event holds.*

The copula process associated with A is denoted as $\hat{\mathbb{C}}_n(\cdot|A)$, i.e., $\hat{\mathbb{C}}_n(\mathbf{u}|A) := \sqrt{n}\big(\hat{C}_n(\mathbf{u}|A) - C(\mathbf{u}|A)\big)$ for any $\mathbf{u} \in [0,1]^p$. Equivalently, one can define the empirical copula as

$$\overline{C}_n(\mathbf{u}|A) := \frac{1}{n\hat{p}_A}\sum_{i=1}^n \mathbf{1}\big(X_{i,1} \leqslant F_{n,1}^{-1}(u_1|A),\dots,X_{i,p} \leqslant F_{n,p}^{-1}(u_p|A), \mathbf{Z}_i \in A\big),$$

invoking usual generalized inverse functions: $F^{-1}(u) := \inf\{t \in \mathbb{R} | F(t) \geqslant u\}$ for every univariate distribution F. Then, the associated copula process becomes $\overline{\mathbb{C}}_n(\cdot|A)$, where

$$\overline{\mathbb{C}}_n(\mathbf{u}|A) := \sqrt{n}\big(\overline{C}_n(\mathbf{u}|A) - C(\mathbf{u}|A)\big),\ \mathbf{u} \in [0,1]^p.$$

Actually, the two latter processes $\hat{\mathbb{C}}_n$ and $\overline{\mathbb{C}}_n$ can be seen as random maps from $[0,1]^p \times \mathcal{A}$ to $\mathbb{R}$, respectively $(\mathbf{u}, A) \mapsto \hat{\mathbb{C}}_n(\mathbf{u}|A)$ and $(\mathbf{u}, A) \mapsto \overline{\mathbb{C}}_n(\mathbf{u}|A)$. In this section, we state the weak convergence of $\hat{\mathbb{C}}_n(\cdot|\cdot)$ and $\overline{\mathbb{C}}_n(\cdot|\cdot)$ in $\ell^\infty([0,1]^p \times \mathcal{A})$.

First, note that the asymptotic behaviors of $\hat{\mathbb{C}}_n$ and $\overline{\mathbb{C}}_n$ are the same. Indeed, adapting the same arguments as in [48], Appendix C, it is easy to check that

$$\sup_{\mathbf{u}\in[0,1]^p} \big|(\hat{C}_n - \overline{C}_n)(\mathbf{u}|A)\big| \leqslant \frac{p}{n\hat{p}_A},$$

everywhere. Since $\inf_{A\in\mathcal{A}} \hat{p}_A \geqslant \underline{p} + o_P(1)$ (Remark 1), we deduce

$$\sup_{A\in\mathcal{A}} \sup_{\mathbf{u}\in[0,1]^p} \big|\sqrt{n}(\hat{C}_n - C)(\mathbf{u}|A) - \sqrt{n}(\overline{C}_n - C)(\mathbf{u}|A)\big| \leqslant \frac{1}{\sqrt{n}(\underline{p} + o_P(1))}, \quad (2)$$

that tends to zero in probability, i.e. $\|\hat{\mathbb{C}}_n - \overline{\mathbb{C}}_n\|_\infty = o_P(1)$. Therefore, the weak limits of $\hat{\mathbb{C}}_n$ and $\overline{\mathbb{C}}_n$ (in particular of $\hat{\mathbb{C}}_n(\cdot|A)$ and $\overline{\mathbb{C}}_n(\cdot|A)$ for a fixed subset $A \in \mathcal{A}$) will be the same.

Second, note that the random variable $U_k^A := F_k(X_k|\mathbf{Z} \in A)$ is uniformly distributed on $[0,1]$, given $(\mathbf{Z} \in A)$, for every $k \in \{1,\dots,p\}$. We denote by $\mathbf{U}^A$ the unobservable random vector $(U_1^A,\dots,U_p^A)$.

Condition 2. *For every $k \in \{1,\dots,p\}$ and $A \in \mathcal{A}$, the partial derivative $\partial_k C(\mathbf{u}|A)$ of $C(\cdot|A)$ w.r.t. u_k exists and is continuous on the set $V_k := \{\mathbf{u} \in [0,1]^p, 0 < u_k < 1\}$. Moreover, the map $\mathbf{u} \mapsto \partial_k C(\mathbf{u}|A)$ is uniformly continuous on $\{\mathbf{u} \in [0,1]^p, u_k \in [\delta, 1-\delta]\}$ and uniformly w.r.t. $A \in \mathcal{A}$, for any $\delta \in (0,1/2)$.*

When there is no conditioning subset, the latter assumption is the standard "minimal" regularity condition, as stated in [57], so that the usual empirical copula process weakly converges in $\ell^\infty([0,1]^p)$. Following [57, Equation (2.2)], we extend the definition of $\partial_k C(\mathbf{u}|A)$ for $\mathbf{u} \in [0,1]^p \backslash V_k$ so that the function $\partial_k C(\cdot|A)$ is defined and continuous on the closed hypercube $[0,1]^p$.

Theorem 2. *If Conditions 1 and 2 hold, then $\hat{\mathbb{C}}_n$ and $\overline{\mathbb{C}}_n$ weakly tend to a centered Gaussian process $\mathbb{C}_\infty$ in $\ell^\infty([0,1]^p \times \mathcal{A})$, where*

$$\begin{aligned}\mathbb{C}_\infty(\mathbf{u}|A) := \frac{\mathbb{B}(\mathbf{u}, A)}{p_A} - \sum_{k=1}^p \frac{\partial_k C(\mathbf{u}|A)}{p_A}\Big\{\mathbb{B}\big((u_k, \mathbf{1}_{-k}), A\big) - u_k \mathbb{B}(\mathbf{1}, A)\Big\} \\ - \frac{C(\mathbf{u}|A)}{p_A}\mathbb{B}(\mathbf{1}, A), \quad (3)\end{aligned}$$

$(u_k, \mathbf{1}_{-k}) := (1,\dots,1,u_k,1,\dots,1)$ with u_k in k-th position, and $\mathbb{B}$ is a Brownian bridge on $[0,1]^p \times \mathcal{A}$, whose covariance function is given by

$$\mathbb{E}\big[\mathbb{B}(\mathbf{u}, A)\mathbb{B}(\mathbf{u}', A')\big] = \mathbb{P}(\mathbf{U}^A \leqslant \mathbf{u}, \mathbf{U}^{A'} \leqslant \mathbf{u}', \mathbf{Z} \in A \cap A')$$

$$- \mathbb{P}(\mathbf{U}^A \leqslant \mathbf{u}, \mathbf{Z} \in A)\mathbb{P}(\mathbf{U}^{A'} \leqslant \mathbf{u}', \mathbf{Z} \in A'), \tag{4}$$

for every $(\mathbf{u}, \mathbf{u}') \in [0,1]^p$ *and* $(A, A') \in \mathcal{A}^2$.

Theorem 2 is proved at the end of this section. By simple calculations, we can explicitly write the covariance function of the limiting conditional copula process $\mathbb{C}_\infty$. Moreover, the latter covariance can be empirically estimated: see the Supplementary Material, Section C.

When there is no conditioning subset, or when $\mathcal{A} = \{\mathbb{R}^q\}$ equivalently, then $p_A = 1$ and $\mathbb{B}(\mathbf{1}, A) = 0$ a.s. because its variance is zero. In this case, $\mathbb{C}_\infty(\mathbf{u}|A)$ is the well-known weak limit of the usual empirical copula process, as stated in [20, 57]. Nonetheless, we stress that Theorem 2 cannot be straightforwardly deduced from the weak convergence of usual empirical copula processes, due to the dependencies between $\mathbf{X}$ and $\mathbf{Z}$.

Remark 3. *Theorem 2 is not a consequence of Theorem 5 in [48] either, where the authors state the weak convergence of the usual empirical copula process in* $\ell^\infty(\mathcal{G})$ *for some set of functions* $\mathcal{G}$ *from* $[0,1]^p$ *to* $\mathbb{R}$. *Indeed, first, such functions are assumed to be right-continuous and of bounded variation (in the sense of Hardy-Krause; see [48]) while we consider general borelian subsets A. Second and more importantly, it is not possible to recover our processes* $\hat{\mathbb{C}}_n(\cdot|A)$ *or* $\overline{\mathbb{C}}_n(\cdot|A)$ *of interest with some quantities* $\int g\, d\mathbb{C}_n$ *for some particular function g and an usual empirical copula process* $\mathbb{C}_n$.

Let us apply the latter results to a finite family $\mathcal{A} := \{A_1, \dots, A_m\}$ of borelian subsets of $\mathbb{R}^q$ such that $p_{A_j} := \mathbb{P}(\mathbf{Z} \in A_j) > 0$ for every $j \in \{1, \dots, m\}$ and a given $m > 0$. The subsets in $\mathcal{A}$ may be disjoint or not. Theorem 2 yields the weak convergence of the process $\vec{\mathbb{C}}_n(\cdot|\mathcal{A})$ defined on $[0,1]^{mp}$ as

$$\vec{\mathbb{C}}_n(\vec{\mathbf{u}}|\mathcal{A}) := \big(\overline{\mathbb{C}}_n(\mathbf{u}_1|A_1), \dots, \overline{\mathbb{C}}_n(\mathbf{u}_m|A_m)\big),$$

for every $\mathbf{u}_j \in [0,1]^p$, $j \in \{1, \dots, m\}$, where $\vec{\mathbf{u}} := (\mathbf{u}_1, \dots, \mathbf{u}_m)$.

Corollary 4. *If, for every* $j \in \{1, \dots, m\}$, $p_{A_j} > 0$ *and Condition 2 holds for* $\mathcal{A} = \{A_1, \dots, A_m\}$, *then* $\vec{\mathbb{C}}_n(\cdot|\mathcal{A})$ *weakly tends to a multivariate centered Gaussian process* $\vec{\mathbb{C}}_\infty(\cdot|\mathcal{A})$ *in* $\ell^\infty([0,1]^{mp}, \mathbb{R}^m)$, *where*

$$\vec{\mathbb{C}}_\infty(\vec{\mathbf{u}}|\mathcal{A}) := \big(\mathbb{C}_\infty(\mathbf{u}_1|A_1), \dots, \mathbb{C}_\infty(\mathbf{u}_m|A_m)\big), \ \ \mathbf{u}_j \in [0,1]^p, \ j \in \{1, \dots, m\},$$

with the same notations as in Theorem 2.

The latter result is obviously true replacing $\overline{\mathbb{C}}_n$ with $\hat{\mathbb{C}}_n$. It will be useful for building and testing the relevance of some partitions $\mathcal{A}$ of the space of covariates, in the spirit of Pearson's chi-square test. Typically, this means testing the equality between the copulas $C_n(\cdot|A_j)$ and $C_n(\cdot|A_k)$ for several couples $(j,k) \in \{1, \dots, m\}^2$.

We can specify the covariance function of $\vec{\mathbb{C}}_\infty(\vec{\mathbf{u}}|\mathcal{A})$ and $\vec{\mathbb{C}}_\infty(\vec{\mathbf{u}}'|\mathcal{A})$, for any vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{u}}'$ in $[0,1]^{mp}$ by recalling the expression of $\mathbb{C}_\infty$, see Equation (3). Note that we have not imposed that the subsets A_j are disjoint. Nonetheless, in the case of a partition (disjoint subsets A_j), our calculations become significantly simpler because of the nullity of $\mathbb{P}(\mathbf{U}^{A_j} \leqslant \mathbf{u}_j, \mathbf{U}^{A_k} \leqslant \mathbf{u}_k, \mathbf{Z} \in A_j \cap A_k)$ when $j \neq k$, see Equation (4).

The rest of this section is devoted to the proof of Theorem 2.

1. Reduction to $\overline{\mathbb{D}}_n$. For every $k \in \{1, \dots, p\}$, the empirical distribution of the unobservable random variable U_k^A given the event $(\mathbf{Z} \in A)$ is

$$G_{n,k}(u|A) := n_A^{-1} \sum_{i=1}^n \mathbf{1}(U_{i,k}^A \leqslant u, \mathbf{Z}_i \in A),\ U_{i,k}^A := F_k(X_{i,k}|\mathbf{Z} \in A),\ i \in \{1, \dots, n\}.$$

Note that $G_{n,k}(u|A)$ and $F_{n,k}(t|A)$ can be seen as an average of n_A indicator functions, i.e., an average on a sub-sample of observations whose size is random. Obviously, $G_{n,k}(u|A)$ tends to $\mathbb{P}(U_k^A \leqslant u|\mathbf{Z} \in A) = u$ almost surely and its associated empirical process will be $\alpha_{n,k}(u|A) := \sqrt{n_A}\big(G_{n,k}(u|A) - u\big)$, $u \in [0,1]$. Note that the normalizing sample size is random here, contrary to the usual empirical processes. Nonetheless, this will not be a source of worry to state some asymptotic behaviors hereafter and n_A could be replaced by np_A in the definition of $\alpha_{n,k}(\cdot|A)$. Then, set

$$\overline{D}_n(\mathbf{u}, A) := n^{-1} \sum_{i=1}^n \mathbf{1}\big(U_{i,1}^A \leqslant G_{n,1}^{-1}(u_1|A), \dots, U_{i,p}^A \leqslant G_{n,p}^{-1}(u_p|A), \mathbf{Z}_i \in A\big),$$

for any $\mathbf{u} \in [0,1]^p$, that tends to

$$D(\mathbf{u}, A) := \mathbb{P}(\mathbf{U}^A \leqslant \mathbf{u}, \mathbf{Z} \in A) = p_A \mathbb{P}(\mathbf{U}^A \leqslant \mathbf{u}|\mathbf{Z} \in A)$$

almost surely. Note that $\big(X_{i,k} \leqslant F_{n,k}^{-1}(u|A)\big)$ if and only if $\big(U_{i,k}^A \leqslant G_{n,k}^{-1}(u|A)\big)$ for any $k \in \{1, \dots, p\}$, $i \in \{1, \dots, n\}$ and $u \in [0,1]$. This implies

$$\overline{C}_n(\mathbf{u}|A) = \overline{D}_n(\mathbf{u}, A)/\hat{p}_A = \overline{D}_n(\mathbf{u}, A)/\overline{D}_n(\mathbf{1}, A).$$

Therefore, the asymptotic behavior of $\overline{\mathbb{C}}_n$ will be deduced from the weak convergence of the process $\overline{\mathbb{D}}_n$, where $\overline{\mathbb{D}}_n(\mathbf{u}, A) := \sqrt{n}(\overline{D}_n - D)(\mathbf{u}, A)$, since $C(\mathbf{u}|A) = \mathbb{P}(\mathbf{U}^A \leqslant \mathbf{u}|\mathbf{Z} \in A) = D(\mathbf{u}, A)/D(\mathbf{1}, A)$. Simple algebra yields

$$\begin{aligned}
\overline{\mathbb{C}}_n(\mathbf{u}|A) &:= \sqrt{n}\big\{\overline{C}_n(\mathbf{u}|A) - C(\mathbf{u}|A)\big\} = \sqrt{n}\left\{\frac{\overline{D}_n(\mathbf{u}, A)}{\overline{D}_n(\mathbf{1}, A)} - \frac{D(\mathbf{u}, A)}{D(\mathbf{1}, A)}\right\} \\
&= \sqrt{n}\overline{D}_n(\mathbf{u}, A)\left\{\frac{1}{\overline{D}_n(\mathbf{1}, A)} - \frac{1}{D(\mathbf{1}, A)}\right\} + \frac{\sqrt{n}(\overline{D}_n - D)(\mathbf{u}, A)}{D(\mathbf{1}, A)} \\
&= \overline{D}_n(\mathbf{u}, A)\frac{\sqrt{n}(D(\mathbf{1}, A) - \overline{D}_n(\mathbf{1}, A))}{\overline{D}_n(\mathbf{1}, A)D(\mathbf{1}, A)} + \frac{\sqrt{n}(\overline{D}_n - D)(\mathbf{u}, A)}{D(\mathbf{1}, A)} \\
&= \frac{\overline{\mathbb{D}}_n(\mathbf{u}, A)}{p_A} - D(\mathbf{u}, A)\frac{\overline{\mathbb{D}}_n(\mathbf{1}, A)}{p_A^2} + o_P(1). \qquad (5)
\end{aligned}$$

Then, the result will follow if we find the weak limit of $\overline{\mathbb{D}}_n$. To this aim, we now introduce two auxiliary related processes $\mathbb{D}_n$ and $\widetilde{\mathbb{D}}_n$.

2. Weak convergence of $\mathbb{D}_n$. The unfeasible empirical counterpart of $D(\mathbf{u}, A)$ is

$$D_n(\mathbf{u}, A) := n^{-1} \sum_{i=1}^{n} \mathbf{1}\big(\mathbf{U}_i^A \leqslant \mathbf{u}, \mathbf{Z}_i \in A\big).$$

A key process is $\mathbb{D}_n := \sqrt{n}\big(D_n - D\big)$ that is seen as a random map from $[0,1]^p \times \mathcal{A}$ to $\mathbb{R}$. As $\mathcal{A}$ is Donsker, this is still the case for the family of maps $\mathcal{D} := \{f_{\mathbf{u},A}; \mathbf{u} \in [0,1]^p, A \in \mathcal{A}\}$, where

$$f_{\mathbf{u},A} : (\mathbf{x}, \mathbf{z}) \mapsto \mathbf{1}(x_1 \leqslant F_1^-(u_1|A), \ldots, x_p \leqslant F_p^-(u_p|A), \mathbf{z} \in A), \tag{6}$$

due to the permanence of the Donsker property (Example 2.10.8 in [63]). Note that $\mathbb{D}_n(\mathbf{u}, A) = \sqrt{n} \int f_{\mathbf{u},A}(\mathbf{x}, \mathbf{z})\, d(\mathbb{P}_n - P)(\mathbf{x}, \mathbf{z})$ a.s., i.e., it is a usual empirical process indexed by a family of maps. Thus, $\mathbb{D}_n$ weakly tends in $\ell^\infty([0,1]^p \times \mathcal{A})$ to a Gaussian process.

3. Reduction to $\widetilde{\mathbb{D}}_n$. We now define the instrumental empirical process

$$\widetilde{\mathbb{D}}_n(\mathbf{u}, A) := \mathbb{D}_n(\mathbf{u}, A) - p_A^{-1} \sum_{k=1}^{p} \partial_k D\big(\mathbf{u}, A\big)\Big\{\mathbb{D}_n\big((u_k, \mathbf{1}_{-k}), A\big) - u_k \mathbb{D}_n(\mathbf{1}, A)\Big\}, \tag{7}$$

denoting by $\partial_k D(\mathbf{u}, A)$ the partial derivative of the map $\mathbf{u} \mapsto D(\mathbf{u}, A)$ w.r.t. u_k. This new process $\widetilde{\mathbb{D}}_n$ will yield a nice approximation of the process of interest $\overline{\mathbb{D}}_n$, as stated in the theorem below.

Theorem 5. *If Conditions 1-2 hold, then*

$$\sup_{A \in \mathcal{A}, \mathbf{u} \in [0,1]^p} |(\overline{\mathbb{D}}_n - \widetilde{\mathbb{D}}_n)(\mathbf{u}, A)| = o_P(1).$$

See the proof in the Supplementary Material, Section B.1. Note that $\widetilde{\mathbb{D}}_n$ differs from the asymptotic approximation of the usual empirical copula process: compare $\widetilde{\mathbb{D}}_n$ with Equation (3.2) and Proposition 3.1 in [57], for instance. This is due to the additional influence of the random sample size n_A, or, equivalently, the randomness of $\hat{p}_A$. This stresses that our results are not straightforward applications of the existing results in the literature.

Since the process $\mathbb{D}_n$ is weakly convergent in $\ell^\infty([0,1]^p \times \mathcal{A})$, we obtain the weak convergence of $\widetilde{\mathbb{D}}_n$ and then of $\overline{\mathbb{D}}_n$ in the same space.

Corollary 6. *If Conditions 1 and 2 hold, then the process $\overline{\mathbb{D}}_n$ weakly converges in $\ell^\infty([0,1]^p \times \mathcal{A})$ towards the centered Gaussian process $\mathbb{D}_\infty$, where*

$$\mathbb{D}_\infty(\mathbf{u}, A) := \mathbb{B}(\mathbf{u}, A) - p_A^{-1} \sum_{k=1}^{p} \partial_k D\big(\mathbf{u}, A\big)\Big\{\mathbb{B}\big((u_k, \mathbf{1}_{-k}), A\big) - u_k \mathbb{B}(\mathbf{1}, A)\Big\},$$

for every $\mathbf{u} \in [0,1]^p$ and $A \in \mathcal{A}$.

4. End of the proof of Theorem 2. We deduce from (5) and Corollary 6 that $\overline{\mathbb{C}}_n$ is weakly convergent in $\ell^\infty([0,1]^p \times \mathcal{A})$, with limit

$$\mathbb{C}_\infty(\mathbf{u}|A) = \frac{\mathbb{D}_\infty(\mathbf{u}, A)}{p_A} - D(\mathbf{u}, A)\frac{\mathbb{D}_\infty(\mathbf{1}, A)}{p_A^2},$$

finishing the proof of Theorem 2. □

Remark 7. *From (5), Theorem 5 and (7), note that*

$$\overline{\mathbb{C}}_n(\mathbf{u}|A) = \frac{\widetilde{\mathbb{D}}_n(\mathbf{u},A)}{p_A} - D(\mathbf{u},A)\frac{\widetilde{\mathbb{D}}_n(\mathbf{1},A)}{p_A^2} + o_P(1).$$

Since $D(\mathbf{u},A) = C(u|A)p_A$ and $D\big((u_k,\mathbf{1}_{-k}),A\big) = u_k p_A$, we deduce another insightful asymptotic representation of $\overline{\mathbb{C}}_n$:

$$\overline{\mathbb{C}}_n(\mathbf{u}|A) = \frac{1}{p_A\sqrt{n}}\sum_{i=1}^n \big\{\mathbf{1}(\mathbf{U}_i \leqslant \mathbf{u}) - C(\mathbf{u}|A)\big\}\mathbf{1}(\mathbf{Z}_i \in A)$$
$$- \quad \frac{1}{p_A\sqrt{n}}\sum_{k=1}^p \partial_k C(\mathbf{u}|A)\sum_{i=1}^n \big\{\mathbf{1}(U_{ik} \leqslant u_k) - u_k\big\}\mathbf{1}(\mathbf{Z}_i \in A) + o_P(1). \quad (8)$$

The previous expansion (8), suggested by a reviewer, clearly shows the close link between the usual empirical copula process and its conditional version we consider here.

2.2. Empirical copula processes conditionally on random subsets

For each $n \geqslant 1$, let A_n be a random borelian subset that depends on a sample of observations $\mathcal{S}_n := \{\mathbf{X}_1, \mathbf{Z}_1, \ldots, \mathbf{X}_n, \mathbf{Z}_n\}$. Assume that $A_n \in \mathcal{A}$ almost surely. Let $A_\infty \in \mathcal{A}$ be a fixed subset that will be considered as "the limit of A_n". Consider a couple $(\mathbf{X}, \mathbf{Z})$ that is independent of the sample $\mathcal{S}_n$. We focus on the law of $\mathbf{X}$ given the event $(\mathbf{Z} \in A_n)$, especially its underlying copula. Therefore, define

$$F(\mathbf{x}|A_n) := \mathbb{P}\big(\mathbf{X} \leqslant \mathbf{x}\big|A_n\,, \mathbf{Z} \in A_n\big) = \frac{\mathbb{P}\big(\mathbf{X} \leqslant \mathbf{x}, \mathbf{Z} \in A_n\big|A_n\big)}{\mathbb{P}\big(\mathbf{Z} \in A_n\big|A_n\big)},$$

for any $\mathbf{x} \in \mathbb{R}^p$. Due to the fact that A_n is a set-valued random variable, note that the random variable defined above is different from the real number

$$F(\mathbf{x}|\mathbf{Z} \in A_n) := \mathbb{P}\big(\mathbf{X} \leqslant \mathbf{x}\big|\mathbf{Z} \in A_n\big) = \frac{\mathbb{P}\big(\mathbf{X} \leqslant \mathbf{x}, \mathbf{Z} \in A_n\big)}{\mathbb{P}\big(\mathbf{Z} \in A_n\big)},$$

where the latter probabilities are relative to the joint law of $(\mathbf{X}, \mathbf{Z}, \mathcal{S}_n)$. Similarly, we define the (random) conditional copula $C(\cdot|A_n)$ as the copula of the random distribution $F(\cdot|A_n)$. The process of interest will be $\sqrt{n}\big\{\overline{C}_n(\cdot|A_n) - C(\cdot|A_\infty)\big\}$ and we want to state some sufficient conditions to obtain its weak convergence in $\ell^\infty\big([0,1]^p\big)$.

First note that

$$\sqrt{n}\big\{\overline{C}_n(\cdot|A_n) - C(\cdot|A_\infty)\big\} = \overline{\mathbb{C}}_n(\cdot|A_\infty) + \sqrt{n}\big\{C(\cdot|A_n) - C(\cdot|A_\infty)\big\}$$
$$+ \quad \big\{\overline{\mathbb{C}}_n(\cdot|A_n) - \overline{\mathbb{C}}_n(\cdot|A_\infty)\big\} =: R_{1,n} + R_{2,n} + R_{3,n}. \quad (9)$$

The results of Section 2.1 typically insure the weak convergence of $R_{1,n}$ in $\ell^\infty\big([0,1]^p\big)$. The behavior of the second term $R_{2,n}$ depends on the definition of A_n and the regularity of $C(\cdot|A_\infty)$. Globally, in many cases, $R_{1,n}+R_{2,n}$ will be weakly convergent to a Gaussian process. To deal with $R_{3,n}$, recall (6) and assume that "A_∞ is the limit of A_n" in the following sense.

Condition 3.

$$\sup_{\mathbf{u}\in[0,1]^p} P\big(f_{\mathbf{u},A_n}-f_{\mathbf{u},A_\infty}\big)^2 \xrightarrow[n\to\infty]{P} 0.$$

Let us state some sufficient conditions so that Condition 3 applies.

Proposition 8. *Condition 3 is satisfied when*

(i) every map $t\mapsto F_k(t|A)$ is strictly increasing for every $A\in\mathcal{A}$, and
(ii) $P(\mathbf{Z}\in A_n\triangle A_\infty|A_n)$ tends to zero with n almost surely.

Denote by $F_{n,\mathbf{Z}}$ the empirical c.d.f. of $(\mathbf{Z}_1,\ldots,\mathbf{Z}_n)$ and by $\mathcal{D}(\mathbb{R}^q)$ the set of all cumulative distribution functions on $\mathbb{R}^q$.

Condition 4. *There exists a map $H:\mathcal{D}(\mathbb{R}^q)\to\ell^\infty([0,1]^p)$ s.t. $C(\cdot|A_n)=H\big(F_{n,\mathbf{Z}}\big)$, and $C(\cdot|A_\infty)=H\big(F_{\mathbf{Z}}\big)$. Moreover, H is Hadamard differentiable at $F_{\mathbf{Z}}$.*

Theorem 9. *Assume Conditions 1-4 hold. Then, $\sqrt{n}\big\{\overline{C}_n(\cdot|A_n)-C(\cdot|A_\infty)\big\}$ weakly tends to a centered Gaussian process in $\ell^\infty\big([0,1]^p\big)$.*

The proofs of Theorem 9 and Proposition 8 can be found in the Supplementary Material, Section B.2.

We now give sufficient conditions so that the map H in Condition 4 is Hadamard-differentiable.

Proposition 10. *Assume Conditions 1-2 hold. Moreover, assume that there exists a map $\widetilde{H}:\mathcal{D}(\mathbb{R}^q)\to\ell^\infty([0,1]^p)$ such that $F(\cdot|A_\infty)=\widetilde{H}(F_{\mathbf{Z}})$, $F(\cdot|A_n)=\widetilde{H}(F_{n,\mathbf{Z}})$, and such that $\widetilde{H}$ is Hadamard-differentiable at $F_{\mathbf{Z}}$. Then Condition 4 applies.*

Proof. Let $\phi:\mathcal{D}(\mathbb{R}^p)\to\ell([0,1]^p)$ be the function that maps a cdf to its copula. Then ϕ is Hadamard-differentiable by Theorem 2.4 in [9]. Conclude by the chain rule since $H=\phi\circ\widetilde{H}$. □

Example 11. *It is natural to define $A_n := \times_{k=1}^p\big[a_k(F_{n,\mathbf{Z}}),b_k(F_{n,\mathbf{Z}})\big]$, for some regular maps a_k and b_k, $a_k<b_k$, $k\in\{1,\ldots,p\}$. Obviously, set $A_\infty := \times_{k=1}^p\big[a_k(F_{\mathbf{Z}}),b_k(F_{\mathbf{Z}})\big]$. Assume the latter maps a_k and b_k are Hadamard differentiable at $F_{\mathbf{Z}}$. Typically, this is the case when $a_k(F_{n,\mathbf{Z}})=F^-_{n,\mathbf{Z}}(q_k)$ for some constants $q_k\in(0,1)$ and every k, and similarly for the maps b_k. In other words, the boxes A_n can be defined through the empirical quantiles of the $\mathbf{Z}$'s components. By the chain rule, we deduce there exists an Hadamard-differentiable $\widetilde{H}$ such that*

$$F(\cdot|A_n)=\mathbb{P}(\mathbf{X}\leqslant\cdot,\mathbf{Z}\in A_n|A_n)/\mathbb{P}(\mathbf{Z}\in A_n|A_n)=\widetilde{H}(F_{n,\mathbf{Z}}),$$

and similarly $F(\cdot|A_\infty) = \widetilde{H}(F_{\mathbf{Z}})$*. Thus, we can apply Proposition 10. If, in addition, Conditions 1-3 are fulfilled, then Theorem 9 applies too.*

3. Bootstrap approximations

The limiting laws of the previous empirical processes $\hat{\mathbb{C}}_n, \overline{\mathbb{C}}_n$ are complex. Therefore, it is difficult to evaluate the weak limits of some functionals of the latter processes. This point is particularly crucial for the estimation of the asymptotic variances or the p-values of some test statistics that may be built from $\hat{\mathbb{C}}_n$ or $\overline{\mathbb{C}}_n$. The usual answer to this problem is to rely on bootstrap. In this section, we study the validity of some bootstrap schemes for our particular empirical copula processes. We will prove the validity of the general exchangeable bootstrap for such processes, a result that has never been formally stated in the literature even in the case of usual copulas, to the best of our knowledge. Moreover, we extend the nonparametric bootstrap and the multiplier bootstrap techniques to the case of conditioning events that have a non-zero probability (the case of pointwise events is dealt in [41]).

3.1. The exchangeable bootstrap

For the sake of generality, we rely on the exchangeable bootstrap (also called "wild bootstrap" by some authors), as introduced in [63]. For every n, let $\mathbf{W}_n := (W_{n,1}, \dots, W_{n,n})$ be an exchangeable nonnegative random vector and $\overline{W}_n := (W_{n,1}, \dots, W_{n,n})/n$ its average. For any borelian subset A, $p_A > 0$, the weighted empirical bootstrap process of $\mathbb{D}_n(\cdot, A)$ that is related to our initial i.i.d. sample $(\mathbf{X}_i, \mathbf{Z}_i)_{i=1,\dots,n}$ is defined as

$$\mathbb{D}^*_n(\mathbf{u}, A) := \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{n,i} \mathbf{1}\Big\{ X_{i,1} \leqslant F^{-1}_{n,1}(u_1|A), \dots, X_{i,p} \leqslant F^{-1}_{n,p}(u_p|A), \mathbf{Z}_i \in A \Big\} \\ - \sqrt{n}\, \overline{W}_n \overline{D}_n(\mathbf{u}, A) \qquad .$$

We require some standard conditions on the weights (Theorem (3.6.13) in [63]).

Condition 5.

$$\sup_n \int_0^\infty \sqrt{\mathbb{P}\big(|W_{n,1} - \overline{W}_n| > t\big)}\, dt < \infty,$$

$$n^{-1/2}\mathbb{E}\big[\max_{1\leqslant i\leqslant n} |W_{n,i} - \overline{W}_n|\big] \to 0, \ \textit{and}\ n^{-1} \sum_{i=1}^n (W_{n,i} - \overline{W}_n)^2 \xrightarrow{P} 1.$$

Note that $\mathbb{D}^*_n(\mathbf{u}, A)$ can be calculated, contrary to $\mathbb{D}_n(\cdot, A)$. Since its asymptotic law will be "close to" the limiting law of $\mathbb{D}_n(\cdot, A)$ when n tends to the infinity, resampling the vector $\mathbf{W}_n$ many times allows the calculation of many realizations of $\mathbb{D}^*_n(\mathbf{u}, A)$, given the initial sample. This will yield a numerical way of approximating the limiting law of $\mathbb{D}_n(\mathbf{u}, A)$ or some functionals of the latter process. We recover the usual and fruitful idea of most resampling techniques. Here, our goal is to formally state the validity of this approach.

Consider the same set $\mathcal{A}$ of borelian subsets as in Section 2. The same reasoning will apply to the copula processes $\hat{\mathbb{C}}_n$ and $\overline{\mathbb{C}}_n$ (seen as processes indexed by $(\mathbf{u}, A) \in [0,1]^p \times \mathcal{A}$), due to the relationships (7) and (5): to prove the validity of an exchangeable bootstrap scheme for the latter copula processes, we first approximate the unfeasible process $\mathbb{D}_n$ by the weighted empirical bootstrapped process $\mathbb{D}_n^*$; second, we invoke Theorem 5 to obtain a similar results for $\overline{\mathbb{D}}_n$; third, we use the relationship between $\overline{\mathbb{D}}_n$ and $\overline{\mathbb{C}}_n$ and deduce a bootstrap approximation for our "conditioned" copula processes.

To be specific, for any integer M, consider M independent realizations of the vector of weights $\mathbf{W}_n$ (that are independent of the initial sample), and the associated processes $\mathbb{D}_{n,k}^*$, $k \in \{1, \dots, M\}$. We first prove the validity of our bootstrap scheme for $\mathbb{D}_n$. Denote by $\mathcal{D}_{n,M}^*$ the process defined on $\big([0,1]^p \times \mathcal{A}\big)^{M+1}$ as

$$\begin{aligned} & \mathcal{D}_{n,M}^*(\mathbf{u}_0, A_0, \mathbf{u}_1, A_1, \dots, \mathbf{u}_M, A_M) \\ := \; & \big(\mathbb{D}_n(\mathbf{u}_0, A_0), \mathbb{D}_{n,1}^*(\mathbf{u}_1, A_1), \dots, \mathbb{D}_{n,M}^*(\mathbf{u}_M, A_M)\big), \end{aligned}$$

for every vectors $\mathbf{u}_0, \dots, \mathbf{u}_M$ in $[0,1]^p$ and every subsets $A_k \in \mathcal{A}$, $k \in \{0, 1, \dots, M\}$. Moreover, denote by $\vec{\mathbb{B}}_\infty$ a process on $\big([0,1]^p \times \mathcal{A}\big)^{M+1}$ that concatenates $M+1$ independent versions of the Brownian bridge $\mathbb{B}$ introduced in Corollary 6.

Theorem 12. *Under Conditions 1, 2, and 5, for any $M \geqslant 2$ and when $n \to \infty$, the process $\mathcal{D}_{n,M}^*$ weakly tends to $\vec{\mathbb{B}}_\infty$ in $\ell^\infty\Big(\big([0,1]^p \times \mathcal{A}\big)^{M+1}, \mathbb{R}^{M+1}\Big)$.*

See the proof in the Supplementary Material, Section B.3. The latter result validates the use of the considered bootstrap scheme. In particular, it implies the weak convergence of $\mathcal{D}_{n,M}^*(\cdot, A_0, \cdot, A_1, \dots, \cdot, A_M)$, seen as a random map from $[0,1]^{p(M+1)}$ to $\mathbb{R}^{M+1}$.

Therefore, we can easily build a bootstrap counterpart of $\widetilde{\mathbb{D}}_n$, and then of $\overline{\mathbb{D}}_n$. Recalling Equation (7), we evaluate the partial derivatives of $D(\cdot, A)$ as in [34]: for any $A \in \mathcal{A}$ and every $\mathbf{u} \in [0,1]^p$, we approximate $\partial_k D(\mathbf{u}, A)$ by

$$\widehat{\partial_k D}(\mathbf{u}, A) := \frac{1}{u_{k,n}^+ - u_{k,n}^-} \Big\{ \overline{D}_n(\mathbf{u}_{-k}, u_{k,n}^+, A) - \overline{D}_n(\mathbf{u}_{-k}, u_{k,n}^-, A) \Big\}, \tag{10}$$

where $u_{k,n}^+ := \min(u_k + n^{-1/2}, 1)$, $u_{k,n}^- := \max(u_k - n^{-1/2}, 0)$ and with obvious notations. Now, the bootstrapped version of $\widetilde{\mathbb{D}}_n(\mathbf{u}, A)$ is defined as

$$\widetilde{\mathbb{D}}_n^*(\mathbf{u}, A) := \mathbb{D}_n^*(\mathbf{u}, A) - \hat{p}_A^{-1} \sum_{k=1}^p \widehat{\partial_k D}\big(\mathbf{u}, A\big) \Big\{ \mathbb{D}_n^*\big((u_k, \mathbf{1}_{-k}), A\big) - u_k \mathbb{D}_n^*(\mathbf{1}, A) \Big\}. \tag{11}$$

Importantly, note that the latter process is a valid bootstrapped approximation of $\overline{\mathbb{D}}_n(\mathbf{u}, A)$ too, because $\widetilde{\mathbb{D}}_n$ and $\overline{\mathbb{D}}_n$ have the same limiting law (Theorem 5).

Denote by $\overline{\mathcal{D}}^*_{n,M}$ the process defined on $\big([0,1]^p \times \mathcal{A}\big)^{M+1}$ by

$$\begin{aligned} & \overline{\mathcal{D}}^*_{n,M}(\mathbf{u}_0, A_0, \mathbf{u}_1, A_1, \dots, \mathbf{u}_M, A_M) \\ := \; & \big(\overline{\mathbb{D}}_n(\mathbf{u}_0, A_0), \widetilde{\mathbb{D}}^*_{n,1}(\mathbf{u}_1, A_1), \dots, \widetilde{\mathbb{D}}^*_{n,M}(\mathbf{u}_M, A_M)\big). \end{aligned}$$

Moreover, denote by $\mathcal{D}_\infty$ a process on $\big([0,1]^p \times \mathcal{A}\big)^{M+1}$ that concatenates $M+1$ independent versions of $\mathbb{D}_\infty$, as defined in Corollary 6. Then, we are able to state the validity of the exchangeable bootstrap for $\overline{\mathbb{D}}_n$.

Theorem 13. *If Conditions 1, 2, and 5 hold, then the process $\overline{\mathcal{D}}^*_{n,M}$ weakly tends to $\mathcal{D}_\infty$ in $\ell^\infty\Big(\big([0,1]^p \times \mathcal{A}\big)^{M+1}, \mathbb{R}^{M+1}\Big)$.*

Proof. With the same arguments as in the proof of Proposition 2 in [34], it can be proved that $\sup_{\mathbf{u}\in[0,1]^p} |\widehat{\partial_k D}(\mathbf{u}, A)| \leqslant 5$ for every $k \in \{1, \dots, p\}$ and every $A \in \mathcal{A}$. Moreover, by Lemma 2 in [34] and recalling the uniform continuity of $\mathbf{u} \mapsto \partial_k D(\mathbf{u}, A)$ (Condition 2), we have, for every a, b s.t. $0 < a < b < 1$,

$$\sup_{A\in\mathcal{A}} \sup_{\mathbf{u}_{-k}\in[0,1]^{p-1}} \sup_{u_k\in[a,b]} |\partial_k D(\mathbf{u}, A) - \widehat{\partial_k D}(\mathbf{u}, A)| \xrightarrow{P} 0.$$

By applying the same arguments as in Proposition 3.2 in [57], we obtain the result. □

Recalling Equation (5), we deduce an exchangeable bootstrapped version of $\overline{\mathbb{C}}_n$, defined as

$$\widetilde{\mathbb{C}}^*_n(\mathbf{u}|A) := \frac{\widetilde{\mathbb{D}}^*_n(\mathbf{u}, A)}{\hat{p}_A} - \overline{D}_n(\mathbf{u}, A)\frac{\widetilde{\mathbb{D}}^*_n(\mathbf{1}, A)}{\hat{p}^2_A}. \tag{12}$$

Still considering M independent random realizations of $\mathbf{W}_n$, we finally introduce the joint process $\mathcal{C}^*_{n,M}$ whose trajectories are

$$\begin{aligned} & (\mathbf{u}_0, A_0, \mathbf{u}_1, A_1, \dots, \mathbf{u}_M, A_M) \mapsto \mathcal{C}^*_{n,M}(\mathbf{u}_0, A_0, \dots, \mathbf{u}_M, A_M) \\ := \; & \big(\overline{\mathbb{C}}_n(\mathbf{u}_0|A_0), \widetilde{\mathbb{C}}^*_{n,1}(\mathbf{u}_1|A_1), \dots, \widetilde{\mathbb{C}}^*_{n,M}(\mathbf{u}_M|A_M)\big), \end{aligned}$$

for every $\mathbf{u}_0, \dots, \mathbf{u}_M$ in $[0,1]^p$ and $A_k \in \mathcal{A}$, $k \in \{0, \dots, M\}$.

Corollary 14. *If Conditions 1, 2, and 5 hold, then, for every $M \geqslant 2$ and when $n \to \infty$, the process $\mathcal{C}^*_{n,M,}$ weakly tends in $\ell^\infty\Big(\big([0,1]^p \times \mathcal{A}\big)^{M+1}, \mathbb{R}^{M+1}\Big)$ to a process that concatenates $M+1$ independent versions of $\mathbb{C}_\infty$, as defined in Theorem 2.*

Therefore, we can approximate the limiting law of $\overline{\mathbb{C}}_n(\mathbf{u}|A)$ by the law of $\widetilde{\mathbb{C}}^*_n(\mathbf{u}|A)$, that is obtained by simulating many times independent realizations of the vector of weights $\mathbf{W}_n$, given the initial sample $(\mathbf{X}_i, \mathbf{Z}_i)_{i=1,\dots,n}$. The same result applies with a finite family $A_1, \dots, A_m$ of subsets in $\mathcal{A}$

Remark 15. *Let $(\xi_i)_{i\geqslant 1}$ be a sequence of i.i.d. random variables, with mean zero and variance one. Formally, we can set $W_{n,k} = \xi_k$ for every n and every $k \in \{1,\dots,n\}$, even if the ξ_i are not always nonnegative. The same formulas as before yield some feasible bootstrapped processes that are similar to those obtained with the multiplier bootstrap of [49], or in [57], Prop. 3.2. With the same techniques of proofs as above, it can be proved that this bootstrap scheme is valid, invoking Theorem 10.1 and Corollary 10.3 in [35] instead of Theorem 3.6.13 in [63]. Therefore, we can state that Corollary 14 applies, replacing $\mathbf{W}_n$ with i.i.d. normalized weights. In other words, the multiplier bootstrap methodology applies with empirical copula processes "indexed by" borelian subsets.*

3.2. The nonparametric bootstrap

When $\mathbf{W}_n$ is drawn along a multinomial law with parameter n and probabilities $(1/n,\dots,1/n)$, we recover the original idea of Efron's nonparametric bootstrap, here applied to the estimation of the limiting law of $\mathbb{D}_n$. Nonetheless, our final bootstrap counterparts $\tilde{\mathbb{C}}_n^*$ for $\hat{\mathbb{C}}_n$ or $\overline{\mathbb{C}}_n$ are not the same as the commonly met nonparametric bootstrap processes. In particular, our techniques require the estimation of some partial derivatives (see (10) and (11)), as in the popular multiplier bootstrap proposed in [49]. As pointed out in [8], this can be seen as a drawback, even if the numerical performances of such bootstrap schemes seem to be good.

Alternatively and more directly, one can still rely on the standard nonparametric bootstrap scheme: simply resample with replacement the initial set of observations and recalculate the statistics of interest with the bootstrapped sample exactly in the same manner as with the initial sample. In practical terms, all analytics and IT codes can be reused as many times as necessary without any additional work. No derivatives have to be numerically evaluated. Indeed, note that the empirical copula may be seen as a regular functional of F_n, the usual empirical distribution of $(\mathbf{X}_i,\mathbf{Z}_i)_{i=1,\dots,n}$, i.e., $\overline{C}_n(\mathbf{u},A) = \psi_0(F_n)(\mathbf{u},A)$ for every $(\mathbf{u},A) \in [0,1]^p \times \mathcal{A}$. Now, apply Efron's initial idea by resampling with replacement n realizations of $(\mathbf{X},\mathbf{Z})$ among the initial sample, and set $\overline{C}_n^* = \psi_0(F_n^*)$, F_n^* being the empirical cdf associated with the bootstrapped sample $(\mathbf{X}_i^*,\mathbf{Z}_i^*)_{i=1,\dots,n}$. Actually, this standard bootstrap scheme is still valid in our framework and we now prove it.

To be specific, let A be a borelian subset, $p_A > 0$ and $\mathbf{x} \in \mathbb{R}^p$. We impose $A \in \mathcal{A}$, for some class $\mathcal{A}$ that follows Assumption 1. As in [9] and many other copula-related papers ([20, 57], etc.), it is more convenient to work with the (unobservable) random vectors $\mathbf{U} \in [0,1]^p$ rather than $\mathbf{X} \in \mathbb{R}^p$. For any $(\mathbf{u},A) \in [0,1]^p \times \mathcal{A}$, set

$$H_n(\mathbf{u},A) := \frac{1}{n\hat{p}_A}\sum_{i=1}^n \mathbf{1}(\mathbf{U}_i \leqslant \mathbf{u}, \mathbf{Z}_i \in A),$$

that is the empirical counterpart of $H(\mathbf{u},A) := \mathbb{P}(\mathbf{U} \leqslant \mathbf{u}|\mathbf{Z} \in A)$. Let $F_{0,n}$ be the empirical cdf of $(\mathbf{U}_i,\mathbf{Z}_i)_{i=1,\dots,n}$ and F_0 denotes the cdf of $(\mathbf{U},\mathbf{Z})$. Obviously,

$\sqrt{n}(F_{0,n}-F_0)$ weakly tends to a tight Gaussian $\mathcal{F}_0$ in $\ell^\infty\big([0,1]^p\times\mathbb{R}^q\big)$. Note that $H_n(\mathbf{u},A)=\chi(F_{0,n})(\mathbf{u},A)$ for some functional χ from the space of distribution functions on $[0,1]^p\times\mathbb{R}^q$, with values in the space of distribution functions on $[0,1]^p$ indexed by $\mathcal{A}$. The latter functional is defined by

$$\chi(F)(\mathbf{u}_0,A)=\int \mathbf{1}(\mathbf{u}\leqslant\mathbf{u}_0,\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z})\,/\int \mathbf{1}(\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z}),$$

when $\mathbf{u}_0\in[0,1]^p$ and $A\in\mathcal{A}$. It is easy to check that the latter function χ is Hadamard differentiable at every cdf F on $[0,1]^p\times\mathbb{R}^q$ s.t. $\int\mathbf{1}(\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z})>0$, tangentially to $C_0([0,1]^p\times\mathbb{R}^q)$, the space of continuous maps on $[0,1]^p\times\mathbb{R}^q$. Its derivative at F is given by

$$\begin{aligned}\chi'(F)(h)(\mathbf{u}_0,A)&=\frac{\int\mathbf{1}(\mathbf{u}\leqslant\mathbf{u}_0,\mathbf{z}\in A)\,h(d\mathbf{u},d\mathbf{z})}{\int\mathbf{1}(\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z})}\\ -\quad&\int\mathbf{1}(\mathbf{u}\leqslant\mathbf{u}_0,\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z})\cdot\frac{\int\mathbf{1}(\mathbf{z}\in A)\,h(d\mathbf{u},d\mathbf{z})}{\Big\{\int\mathbf{1}(\mathbf{z}\in A)\,F(d\mathbf{u},d\mathbf{z})\Big\}^2}.\end{aligned}$$

When h is not of bounded variation, the latter integrals are defined by an integration by parts formula (see Theorem 15 in [48]). Moreover, $\overline{C}_n(\mathbf{u}|A)=\phi\big(H_n(\cdot,A)\big)(\mathbf{u})=:\tilde\phi(H_n)(\mathbf{u},A)$, introducing a map ϕ from the space of distribution functions on $[0,1]^p$ to $\ell^\infty([0,1]^p)$ by

$$\phi(F)(\mathbf{u})=F\big(F_1^-(u_1),\dots,F_p^-(u_p)\big).$$

Let $C_0([0,1]^p)$ be the space of continuous maps on $[0,1]^p$. Moreover $D_0([0,1]^p)$ denotes the set of maps $f\in C_0([0,1]^p)$ s.t. $f(1,\dots,1)=0$, and $f(\mathbf{x})=0$ if some component of $\mathbf{x}$ is zero. Theorem 2.4 in [9] proved that ϕ is Hadamard-differentiable tangentially to $D_0([0,1]^p)$. By a careful reading of their proof, it can be checked that the latter property is uniform w.r.t. $A\in\mathcal{A}$, due to the assumed uniform continuity of the maps $(\mathbf{u},A)\mapsto\partial_k C(\mathbf{u}|A)$, $k\in\{1,\dots,p\}$ (see Condition 2). This means $\tilde\phi$ is Hadamard-differentiable tangentially to $D_{\mathcal{A}}([0,1]^p)$, the space of functions h on $[0,1]^p\times\mathcal{A}$ s.t. $h(\cdot,A)\in D_0([0,1]^p)$ for every $A\in\mathcal{A}$. Since $\chi'(F_0)(h)$ belongs to $D_{\mathcal{A}}([0,1]^p)$ for any $h\in C_0([0,1]^p\times\mathbb{R}^q)$, we can invoke the chain rule (Lemma 3.9.3 in [63]). This means that $\psi:=\tilde\phi\circ\chi$ is Hadamard differentiable at F_0 tangentially to $C_0([0,1]^p\times\mathbb{R}^q)$ and its derivative is $\psi'(F_0)=\tilde\phi'\big(\chi(F_0)\big)\circ\chi'(F_0)$. This is the main condition to apply the Delta-Method for bootstrap (Theorem 3.9.11 in [63], e.g.), because $\sqrt{n}\big\{\overline{C}_n(\mathbf{u}|A)-C(\mathbf{u}|A)\big\}=\sqrt{n}\big\{\psi(F_{0,n})-\psi(F_0)\big\}(\mathbf{u},A)$.

The nonparametric bootstrapped empirical copula associated with $\overline{C}_n(\cdot|\cdot)$ is then defined as

$$\begin{aligned}&\overline{C}^*_n(\mathbf{u}|A)\\ :=\quad&\frac{1}{n\hat p^*_A}\sum_{i=1}^n\mathbf{1}\big\{X^*_{i,1}\leqslant(F^*_{n,1})^{-1}(u_1|A),\dots,X^*_{i,p}\leqslant(F^*_{n,p})^{-1}(u_p|A),\mathbf{Z}^*_i\in A\big\},\end{aligned}$$

and the associated bootstrapped copula process is given by

$$\overline{\mathbb{C}}_n^*(\mathbf{u}|A) := \sqrt{n}\big(\overline{C}_n^*(\mathbf{u}|A) - \overline{C}_n(\mathbf{u}|A)\big),\ \mathbf{u} \in [0,1]^p, A \in \mathcal{A}.$$

Obviously, F_n^* and the $F_{n,k}^*$ respectively denote the associated empirical cdf and the empirical marginal cdfs' associated with the nonparametric bootstrap sample $(\mathbf{X}_i^*, \mathbf{Z}_i^*)_{i=1,\dots,n}$. By mimicking the arguments of [20], Theorem 5, it is easy to state the validity of the nonparametric bootstrap scheme for $\overline{C}_n(\cdot|\cdot)$. Details are left to the reader.

To simply announce the result, introduce the random map

$$\begin{aligned} & \mathcal{C}_{n,M}(\mathbf{u}_0, A_0, \mathbf{u}_1, A_1, \dots, \mathbf{u}_M, A_M) \\ := \quad & \big(\overline{\mathbb{C}}_n(\mathbf{u}_0|A_0), \overline{\mathbb{C}}_{n,1}^*(\mathbf{u}_1|A_1), \dots, \overline{\mathbb{C}}_{n,M}^*(\mathbf{u}_M|A_M)\big), \end{aligned}$$

for every vectors $\mathbf{u}_0, \dots, \mathbf{u}_M$ in $[0,1]^p$ and every subsets $A_0, A_1, \dots, A_M$ in $\mathcal{A}$.

Theorem 16. *If Condition 1 and 2 are satisfied, then the process $\mathcal{C}_{n,M}$ weakly converges in $\ell^\infty\Big(\big([0,1]^p \times \mathcal{A}\big)^{M+1}, \mathbb{R}^{M+1}\Big)$ to a process that concatenates $M+1$ independent versions of $\mathbb{C}_\infty$.*

Let us detail the latter result when dealing with a finite family of subsets $\mathcal{A} := \{A_1, \dots, A_m\}$. Then, for every $\vec{\mathbf{u}}_j := (\mathbf{u}_{j,1}, \dots, \mathbf{u}_{j,m})$, $\mathbf{u}_{j,k} \in [0,1]^p$ for every $j \in \{0, 1, \dots, M\}$, $k \in \{1, \dots, m\}$, set

$$\vec{\mathbb{E}}_{n,j}^*(\vec{\mathbf{u}}_j, \mathcal{A}) := \big(\overline{\mathbb{C}}_n^*(\mathbf{u}_{j,1}, A_1), \dots, \overline{\mathbb{C}}_n^*(\mathbf{u}_{j,m}, A_m)\big), \text{ and}$$

$$\vec{\mathcal{E}}_{n,M,\mathcal{A}}(\vec{\mathbf{u}}_0, \dots, \vec{\mathbf{u}}_M) := \big(\vec{\mathbb{C}}_n(\vec{\mathbf{u}}_0|\mathcal{A}), \vec{\mathbb{E}}_{n,1}^*(\vec{\mathbf{u}}_1|\mathcal{A}), \dots, \vec{\mathbb{E}}_{n,M}^*(\vec{\mathbf{u}}_M|\mathcal{A})\big).$$

Theorem 17. *If Condition 1 and Condition 2 is satisfied, then, for every $M \geqslant 2$ and when n tends to the infinity, the process $\vec{\mathcal{E}}_{n,M,\mathcal{A}}$ weakly converges in $\ell^\infty([0,1]^{p(M+1)m}, \mathbb{R}^{m(M+1)})$ to a process that concatenates $M+1$ independent versions of $\vec{\mathbb{C}}_\infty(\cdot|\mathcal{A})$ (as defined in Corollary 4).*

As in Section 2.2, consider the case of a sequence of subsets $(A_n)_{n\geqslant 1}$ in $\mathcal{A}$ such that $A_n = H(F_{n,\mathbf{Z}})$ for some Hadamard differentiable map H. The limit of this sequence is $A_\infty := H(F_{\mathbf{Z}})$. This is particularly the case when the A_n are defined by some empirical quantiles of $\mathbf{Z}$'s components (random "boxes"). Then, the process $\sqrt{n}\big\{\overline{C}_n(\cdot|A_n) - C(\cdot|A_\infty)\big\}$ can most often be nonparametrically bootstrapped exactly as above when we can apply Theorem 9, i.e., when we can write

$$\sqrt{n}\big\{\overline{C}_n(\cdot|A_n) - C(\cdot|A_\infty)\big\} = \overline{\mathbb{C}}_n(\cdot|A_\infty) + \sqrt{n}\big\{C(\cdot|A_n) - C(\cdot|A_\infty)\big\} + o_P(1),$$

as in Equation (9). Therefore, for every bootstrap sample and its associated empirical cdf F_n^*, set $A_n^* := H(F_{n,\mathbf{Z}}^*)$ and the associated bootstrapped process is then $\sqrt{n}\big\{\overline{C}_n^*(\cdot|A_n^*) - \overline{C}_n(\cdot|A_n)\big\}$.

4. Generalized multivariate measures of association processes

Measures of association (in particular "measures of concordance" and "measures of dependence"; see [38], Def. 5.1.7. and 5.3.1.) are real numbers that summarize the amount of dependencies across the components of a random vector. Most of the time, they are defined for bivariate vectors, as originally formalized in [51]. The most usual ones are Kendall's tau, Spearman's rho, Gini's measures of association and Blomqvist's beta. Denoting by C the copula of a bivariate random vector (X_1, X_2), most of the measures of association that have been proposed in the literature can be rewritten as weighted sums of quantities as $\rho_1(\psi, \alpha) := \int \psi(u,v)\, C^\alpha(u,v) C(du,dv)$ for some measurable map $\psi : [0,1]^2 \to \mathbb{R}$, $\alpha \geqslant 0$, or as $\rho_2(\psi, \alpha, \mu) := \int \psi(u,v)\, C^\alpha(u,v) \mu(du,dv)$ for some measure μ on $[0,1]^2$. For example, in the case of Kendall's tau (resp. Spearman's rho), the first case (resp. second case) applies by setting $\psi = 1$ and $\alpha = 1$ (resp. $\alpha = 1$, $\mu(du,dv) = du\,dv$). Gini's index is $\rho_1(\psi_G, 0)$, with $\psi_G(u,v) := 2\big(|u+v-1| - |u-v|\big)$. Blomqvist's beta is obtained with $\rho_2(1, 1, \delta_{(1/2,1/2)})$, where $\delta_{(1/2,1/2)}$ denotes the Dirac measure at $(1/2, 1/2)$. See [38], Chapter 5, or [39] for some justifications of the latter results and additional results.

A few multivariate extensions of the latter measures of association have been introduced in the literature for many years. The axiomatic justification of such measures for p-dimensional random vectors has been developed in [61], and many proposals followed, sometimes in passing. The most extensive analysis has been led in a series of papers by F. Schmid, R. Schmidt and some co-authors: see [53, 54, 55, 56].

Actually, we can significantly extend the previous ideas by considering general formulas for multivariate measures of association, possibly indexed by subsets (of covariates), as in the previous sections. To be specific, we still consider a random vector $(\mathbf{X}, \mathbf{Z}) \in \mathbb{R}^p \times \mathbb{R}^q$ and we will be interested in measures of association between the components of $\mathbf{X}$, when $\mathbf{Z}$ belongs to some borelian subset A in $\mathbb{R}^q$. We will focus on the wide range of measures of association that are defined as functionals of the underlying conditional copulas only (not margins). For any (possibly empty) subsets K and K' that are included in $I := \{1, \dots, p\}$, let us define

$$\rho_{K,K'}(A) := \int \psi(\mathbf{u})\, C_K(\mathbf{u}_K | \mathbf{Z} \in A) C_{K'}(d\mathbf{u}_{K'} | \mathbf{Z} \in A)\, d\mathbf{u}_{I \setminus K'}, \qquad (13)$$

for some measurable function ψ. Obviously, $C_K(\cdot | \mathbf{Z} \in A)$ denotes the conditional copula of $\mathbf{X}_K := (X_j, j \in K)$ given $(\mathbf{Z} \in A)$. In particular, $C_I(\mathbf{u}|\mathbf{Z} \in A) = C_{\{1,\dots,p\}}(\mathbf{u}|\mathbf{Z} \in A) = C_{\mathbf{X}|\mathbf{Z}}(\mathbf{u}|\mathbf{Z} \in A)$, for every $\mathbf{u} \in [0,1]^p$. When $K' = \varnothing$ (resp. $K' = I$) there is no integration w.r.t. $C_{K'}(d\mathbf{u}_{K'}|\mathbf{Z} \in A)$ (resp. $d\mathbf{u}_{I \setminus K'}$).

The latter definition virtually includes and/or extends all unconditional and conditional measures of association that have been introduced until now. Indeed, such measures are linear combinations (or even ratios, possibly) of our quantities $\rho_{K,K'}(A)$, for conveniently chosen (K, K') and ψ. Note that, by setting $A = \mathbb{R}^q$, we recover unconditional measures of association. Moreover, setting $A = (\mathbf{Z} = \mathbf{z})$ allows to recover pointwise conditional measures of association.

A few examples of such $\rho_{K,K'}(A)$ that have already been met in the literature:

- $\psi(\mathbf{u}) = 1$, $K = K' = I$ and $A = \mathbb{R}^q$ provides a multivariate version of the Kendall's taus' of $\mathbf{X}$, that are affine functions of $\int C_{\mathbf{X}}(\mathbf{u})\, C_{\mathbf{X}}(d\mathbf{u})$. See [22, 23, 30], among others;
- $\psi(\mathbf{u}) = 1$, $K = I$, $K' = \varnothing$ and $A = \mathbb{R}^q$ yields ρ_1, the multivariate Spearmans's rho of $\mathbf{X}$, as in [53]; see [67] too.
- $\psi(\mathbf{u}) = 1$, $K = \varnothing$, $K' = I$ and $A = \mathbb{R}^q$ yields the multivariate Spearmans's rho of $\mathbf{X}$ introduced in [50], also called ρ_2 in [53];
- $\psi(\mathbf{u}) = 1$, $K = K' = I$ and choosing A as a (small) neighborhood of $\mathbf{z}$ is similar to a p-dimensional extension of the pointwise conditional Kendall's tau studied in [65] or [15, 16];
- $\psi(\mathbf{u}) = \prod_{j\in I} 1(u_j \leqslant 1/2)$, $K = \varnothing$ and $K' = I$ corresponds to a conditional version of Blomqvist coefficient ([38]);
- $\psi(\mathbf{u}) = 1(\mathbf{u} \leqslant \mathbf{u}_0) + 1(\mathbf{u} \geqslant \mathbf{v}_0)$, $K = \varnothing$ and $K' = I$ yields a conditional version of the tail-dependence coefficient considered in [54];
- if ψ is a density on $[0,1]^p$, $K = I$ and $K' = \varnothing$, we get some conditional product measures of concordance, as defined in [61];
- when $\psi(\mathbf{u})$ is a weighted sum of reflection indicators of the type

  $$\mathbf{u} \in [0,1]^p \mapsto \big(\epsilon_1 u_1 + (1-\epsilon_1)(1-u_1), \ldots, (\epsilon_p u_p + (1-\epsilon_p)(1-u_p)\big),$$

  where $\epsilon_k \in \{0,1\}$ for every $k \in \{1,\ldots,p\}$, we obtain some generalizations of measures of association (Kendall's tau, Blomqvist coefficient, etc), as introduced in [30]. For conveniently chosen weights, such linear combinations of $\rho_{K,K'}(\mathbb{R}^q)$ for different subsets K and K' yield measures of association that are increasing w.r.t. a so-called "concordance ordering" property. See [61], Examples 7 and 8, too. Etc.

Note that our methodology includes as particular cases some multivariate measures of association that are calculated as averages of "usual" measures of association when they are calculated for many pairs (X_k, X_l), $k,l \in \{1,\ldots,p\}^2$. This old and simple idea (see [32]) has been promoted by some authors. See such types of multivariate measures in [56] and the references therein.

Generally speaking, it is possible to estimate the latter quantities $\rho_{K,K'}(A)$ by replacing the conditional copulas with their estimates in Equation (13). This yields the estimator

$$\hat{\rho}_{K,K'}(A) := \int \psi(\mathbf{u})\, \hat{C}_{n,K}(\mathbf{u}_K|\mathbf{Z}\in A)\hat{C}_{n,K'}(d\mathbf{u}_{K'}|\mathbf{Z}\in A)\, d\mathbf{u}_{I\setminus K'}, \tag{14}$$

where we define

$$\hat{C}_{n,K}(\mathbf{u}_K|A) := \frac{1}{n\hat{p}_A}\sum_{i=1}^n \mathbf{1}\big(F_{n,j}(X_{i,j}|A) \leqslant u_j, \forall j \in K; \mathbf{Z}_i \in A\big),$$

and similarly for the induced measure $\hat{C}_{n,K'}(d\mathbf{u}_{K'}|\mathbf{Z}\in A)$.

Now, we want to derive the limiting law of $\mathbb{G}_n(A,K,K',\psi) := \sqrt{n}\big(\hat{\rho}_{K,K'}(A) - \rho_{K,K'}(A)\big)$. Here, $\mathbb{G}_n$ will be seen as a process indexed by $(A,K,K',\psi) \in \mathcal{A} \times$

$\mathcal{P}_p^2 \times \Psi$, where $\mathcal{P}_p$ denotes the set of subsets of $I = \{1, \dots, p\}$ and Ψ denotes a set of "sufficiently regular maps" ψ from $[0,1]^p$ to $\mathbb{R}$.

Condition 6. *Any $\psi \in \Psi$ is right-continuous, i.e. coordinatewise right-continuous in each coordinate and at every point, and of bounded variation in the sense of Hardy-Krause (see [48]). Moreover, $\sup_{\psi\in\Psi} \|\psi\|_\infty < +\infty$. For every $\epsilon > 0$ and $\psi \in \Psi$, there exists a partition of $[0,1]^p$ with $q = q(\varepsilon, \psi)$ disjoint hyper-rectangles $R_{j,\psi} = (\mathbf{a}_{j,\psi}, \mathbf{b}_{j,\psi}]$ and some coefficients $c_{j,\psi} \in \mathbb{R}$ such that the stepwise functions $s_\varepsilon(\mathbf{u}, \psi) := \sum_{j=1}^q c_{j,\psi} \mathbf{1}(\mathbf{u} \in R_{j,\psi})$ satisfy*

$$\sup_{\psi\in\Psi} \sup_{\mathbf{u}\in[0,1]^p} \big| s_\varepsilon(\mathbf{u}, \psi) - \psi(\mathbf{u}) \big| < \varepsilon, \text{ with } \sup_{\psi\in\Psi} \sum_{j=1}^q |c_{j,\psi}| < \infty.$$

The latter condition is satisfied for any finite family of right-continuous bounded ψ functions of bounded variation, in particular.

We will deduce the weak convergence of $\mathbb{G}_n$ in $\ell^\infty(\mathcal{A} \times \mathcal{P}_p^2 \times \Psi)$ from the weak convergence of the process $\hat{\mathbb{C}}_n$. Indeed, note that $\mathbb{G}_n = \sqrt{n}\big\{\Psi(\hat{C}_n) - \Psi(C)\big\}$, where Ψ is a map from $\mathcal{C}_{p,\mathcal{A}}$, the space of the cdfs' on $[0,1]^p$ indexed by a parameter $A \in \mathcal{A}$, to $\ell^\infty(\mathcal{A} \times \mathcal{P}_p^2 \times \Psi)$. It is defined as

$$\Psi(C) : (A, K, K', \psi) \mapsto \int \psi(\mathbf{u})\, C_K(\mathbf{u}_K|A) C_{K'}(d\mathbf{u}_{K'}|A)\, d\mathbf{u}_{I\setminus K'} \tag{15}$$

for every $C \in \mathcal{C}_{p,\mathcal{A}}$.

To apply the Delta-Method, we need to prove that Ψ is Hadamard-differentiable. To this aim, the trajectories of our limiting process have to be sufficiently regular uniformly w.r.t. $A \in \mathcal{A}$.

Definition 18. *A map $h \in \ell^\infty([0,1]^p \times \mathcal{A})$ is said to be $\mathcal{A}$-regular if it satisfies the following conditions:*

(i) for every $A \in \mathcal{A}$, the map $\mathbf{u} \mapsto h(\mathbf{u}, A)$ is continuous on $[0,1]^p$;
(ii) for every $\epsilon > 0$ and $A \in \mathcal{A}$, there exists a partition of $[0,1]^p$ with $m = m(\varepsilon, A)$ disjoint hyper-rectangles $R_{k,A} = (\mathbf{a}_{k,A}, \mathbf{b}_{k,A}]$ and some coefficients $d_{k,A} \in \mathbb{R}$ such that the stepwise functions $w_\varepsilon(\mathbf{u}, A) := \sum_{k=1}^m d_{k,A} \mathbf{1}(\mathbf{u} \in R_{k,A})$ satisfy

$$\sup_{A\in\mathcal{A}} \sup_{\mathbf{u}\in[0,1]^p} \big| w_\varepsilon(\mathbf{u}, A) - h(\mathbf{u}, A) \big| < \varepsilon, \text{ and } \sup_{A\in\mathcal{A}} \sum_{k=1}^m |d_{k,A}| < \infty.$$

For every $K, K' \in \mathcal{P}_p$ and every $\mathbf{u}_{K'} \in [0,1]^{|K'|}$, denote

$$\chi_{K,K'}(\mathbf{u}_{K'}; \psi, A) := \int \psi(\mathbf{u})\, C_K(\mathbf{u}_K|A)\, d\mathbf{u}_{I\setminus K'}.$$

Lemma 19. *Let $C(\cdot|\cdot) \in \mathcal{C}_{p,\mathcal{A}}$ be a set of p-dimensional conditional copulas and let Ψ be a family that satisfies Condition 6. Assume that, for every $K, K' \in \mathcal{P}_p$, the map $\chi_{K,K'}(\cdot; \psi, A)$ is of bounded variation on $[0,1]^{|K'|}$, uniformly over*

$\psi \in \Psi$ and $A \in \mathcal{A}$. Then, the map $\Psi : \mathcal{C}_{p,\mathcal{A}} \longrightarrow \ell^\infty(\mathcal{A} \times \mathcal{P}_p^2 \times \Psi)$ is Hadamard-differentiable at C, tangentially to the set $\mathcal{H}$ of $\mathcal{A}$-regular maps. Its derivative is given by

$$\big(\Psi'(C)(h)\big)(A, K, K', \psi) = \int \psi(\mathbf{u})\, h_K(\mathbf{u}_K, A) C_{K'}(d\mathbf{u}_{K'}|A)\, d\mathbf{u}_{I\setminus K'} + \int \psi(\mathbf{u})\, C_K(\mathbf{u}_K|A) h_{K'}(d\mathbf{u}_{K'}, A)\, d\mathbf{u}_{I\setminus K'},$$

for any map $h \in \mathcal{H}$.

When h is not of bounded variation, we define the second integral of $\Psi'_{K,K'}(C)(h)$ by an integration by parts, as detailed in [48]. See the proof of Lemma 19 in the Supplementary Material, Section B.5. The natural sets $\mathcal{H}$ will be given by the trajectories of the limiting law of $\hat{\mathbb{C}}_n$, i.e. the trajectories of $\mathbb{C}_\infty$ indexed by $A \in \mathcal{A}$ (see Theorem 2). For any fixed borelian subset A, there exists a version of $\mathbb{C}_\infty$ with continuous trajectories. Thus, when $\mathcal{A}$ is a finite set and for the latter version of $\mathbb{C}_\infty$, the map $(\mathbf{u}, A) \mapsto \mathbb{C}_\infty(\mathbf{u}|A)(\omega)$ from $[0,1]^p \times \mathcal{A}$ to $\mathbb{R}$ is $\mathcal{A}$-regular for every realization $\omega \in \Omega$.

As a consequence, by applying the Delta Method (Theorem 3.9.4 in [63]) to the copula process $\sqrt{n}\big(\hat{C}_n(\cdot|\cdot) - C(\cdot|\cdot)\big)$, we obtain the weak convergence of $\mathbb{G}_n$.

Theorem 20. *Assume that, for almost every realization $\omega \in \Omega$, Lemma 19 can be applied with*

$$\mathcal{H} = \mathcal{H}_\omega := \{(\mathbf{u}, A) \mapsto \mathbb{C}_\infty(\mathbf{u}|A)(\omega),\ \mathbf{u} \in [0,1]^p, A \in \mathcal{A}\},$$

for some version of $\mathbb{C}_\infty$. Then, under the assumptions of Theorem 2, the process $\mathbb{G}_n(A, K, K', \psi) := \sqrt{n}\big(\hat{\rho}_{K,K'}(A) - \rho_{K,K'}(A)\big)$ weakly tends to a centered Gaussian process $\mathbb{G}_\infty$ in $\ell^\infty(\mathcal{A} \times \mathcal{P}_p^2 \times \Psi)$, whose covariance function is given by

$$\begin{aligned}
&\mathbb{E}\big[\mathbb{G}_\infty(A_1, K_1, K_1', \psi_1)\mathbb{G}_\infty(A_2, K_2, K_2', \psi_2)\big] := \int \psi_1(\mathbf{u})\psi_2(\mathbf{v}) \\
&\quad \times \mathbb{E}\Big[\big\{\mathbb{C}_{\infty,K_1}(\mathbf{u}_{K_1}|A_1) C_{K_1'}(d\mathbf{u}_{K_1'}|A_1) + C_{K_1}(\mathbf{u}_{K_1}|A_1)\mathbb{C}_{\infty,K_1'}(d\mathbf{u}_{K_1'}|A_1)\big\} \\
&\quad \times \big\{\mathbb{C}_{\infty,K_2}(\mathbf{v}_{K_2}|A_2) C_{K_2'}(d\mathbf{v}_{K_2'}|A_2) \\
&\quad + C_{K_2}(\mathbf{v}_{K_2}|A_2)\mathbb{C}_{\infty,K_2'}(d\mathbf{v}_{K_2'}|A_2)\big\}\Big]\, d\mathbf{u}_{I\setminus K_1'}\, d\mathbf{v}_{I\setminus K_2'}.
\end{aligned}$$

Let us specify the previous general result in a usual situation.

Corollary 21. *Consider a fixed A such that $p_A > 0$, some fixed subsets K, K' and a bounded function ψ that is right-continuous and of bounded variation. Assume that the conditional margins $F_k(\cdot|\mathbf{Z} \in A)$ are continuous, for any $k \in \{1, \ldots, p\}$ and that the map $\chi_{K,K'}(\cdot;\psi, A)$ is of bounded variation on $[0,1]^{|K'|}$. Then, under Assumption 2, we have*

$$\sqrt{n}\big(\hat{\rho}_{K,K'}(A) - \rho_{K,K'}(A)\big) \xrightarrow{w} \mathcal{N}\big(0, \sigma^2_{K,K'}(A)\big), \quad \textit{where}$$

$$\sigma^2_{K,K'}(A) := \mathrm{V}ar\Big(\int \psi(\mathbf{u})\, \mathbb{C}_{\infty,K}(\mathbf{u}_K|A) C_{K'}(d\mathbf{u}_{K'}|A)\, d\mathbf{u}_{I\setminus K'}$$
$$+ \quad \int \psi(\mathbf{u})\, C_K(\mathbf{u}_K|A) \mathbb{C}_{\infty,K'}(d\mathbf{u}_{K'}|A)\, d\mathbf{u}_{I\setminus K'} \Big).$$

As an example, let us consider the multivariate Spearman's rho obtained by setting $\psi(\mathbf{u}) = 1$, $K = I$, $K' = \varnothing$, $p = q$, $\mathbf{X} = \mathbf{Z}$ and $A = \prod_{j=1}^p]-\infty, a_j]$, for some threshold $(a_1, \dots, a_p)$ in $\mathbb{R}^p$. In other words, we focus on

$$\rho_S(\mathbf{a}) := \int C_{\mathbf{X}}(\mathbf{u}|X_j \leqslant a_j, \forall j \in \{1, \dots, p\}) \prod_{j=1}^p du_j.$$

This measure is related to the average dependencies among the components of $\mathbf{X}$, knowing that all these components are observed in their own tails. Indeed, we are interested in the joint tail event $X_j \leqslant a_j$ for every $j \in \{1, \dots, p\}$. A similar measure has been introduced in [53] but its properties have not been studied, since the techniques developed in this article were not available. Therefore, they preferred to concentrate on other Spearman's rho-type measures of association. Now, we fill this gap by applying Theorem 20. With our notations, a natural estimator of $\rho_S(\mathbf{a})$ is

$$\hat{\rho}_S(\mathbf{a}) := \int \hat{C}_n(\mathbf{u}|X_j \leqslant a_j, \forall j \in \{1, \dots, p\}) \prod_{j=1}^p du_j.$$

Corollary 22. *If $p_A > 0$ and Condition 2 holds, then $\sqrt{n}\big(\hat{\rho}_S(\mathbf{a}) - \rho_S(\mathbf{a})\big)$ weakly tends to a r.v $\mathcal{N}\big(0, \sigma^2_S(\mathbf{a})\big)$, where $\sigma^2_S(\mathbf{a}) := \int \mathbb{E}\big[\mathbb{C}_\infty(\mathbf{u}_1|A)\mathbb{C}_\infty(\mathbf{u}_2|A)\big]\, d\mathbf{u}_1\, d\mathbf{u}_2$.*

The analytic formula of $\mathbb{E}\big[\mathbb{C}_\infty(\mathbf{u}_1|A)\mathbb{C}_\infty(\mathbf{u}_2|A)\big]$ is provided in the Supplementary Material, Section C. The asymptotic variance $\sigma^2_S(\mathbf{a})$ can be consistently estimated after replacing the unknown quantities $C(\cdot|A)$, p_A, $D(\cdot, A)$ and its partial derivatives by some empirical counterparts, as in Section 3.1. Alternatively, the limiting law of $\sqrt{n}\big(\hat{\rho}_S(\mathbf{a}) - \rho_S(\mathbf{a})\big)$ can be obtained by several bootstrap schemes, as explained in Section 3. Indeed, since $\sqrt{n}\big(\hat{\rho}_S(\mathbf{a}) - \rho_S(\mathbf{a})\big) = \int \hat{\mathbb{C}}_n(\mathbf{u}|A)\, d\mathbf{u}$, a bootstrap equivalent of the latter statistics is $\int \widetilde{\mathbb{C}}^*_n(\mathbf{u}, A)\, d\mathbf{u}$ or $\int \overline{\mathbb{C}}^*_n(\mathbf{u}, A)\, d\mathbf{u}$, with the same notations as above and conveniently chosen bootstrap weights.

Remark 23. *Theorem 20 is very general as it potentially allows to manage infinite families of ψ functions, i.e. infinite families of (conditional) dependence measures. Even if this situation is non-standard, the latter result could be useful in some circumstances. For instance, consider the family of bivariate maps $\psi_{\alpha,\beta}(u, v) := (\alpha + 1)(\beta + 1)u^\alpha v^\beta$, $(u, v) \in [0, 1]^2$ and positive parameters α and β. For every (α, β), the quantity $\rho(\alpha, \beta) := \mathbb{E}\big[\psi_{\alpha,\beta}(U_1, U_2)\big]$ is of the type (13), with $K = \varnothing$, $K' = I$ and $A = \mathbb{R}^q$. It may be nonparametrically estimated as above, by $\hat{\rho}(\alpha, \beta)$. Even more, it is possible to average the latter measures of association and to focus on the aggregated quantity $\rho := \int \rho(\alpha, \beta)\, \nu(d\alpha, d\beta)$, for* *some finite measure ν on $\mathbb{R}_+^2$. The natural estimator of this "global measure" ρ would be $\hat\rho := \int \hat\rho(\alpha,\beta)\,\nu(d\alpha,d\beta)$. Theorem 20 and the Delta-method would yield the asymptotic behavior of the latter estimator.*

Important practical questions can arise by considering several borelian subsets simultaneously. For instance, is the amount of dependencies among the $\mathbf{X}$'s components the same when $\mathbf{Z}$ belongs to different subsets? This questioning can lead to a way of building relevant subsets A_j, $j \in \{1,\dots,p\}$. Typically, a nice partition of the $\mathbf{Z}$-space is obtained when the copulas $C(\cdot|\mathbf{Z}\in A_j)$ are heterogeneous. In the very general framework of Theorem 20, we will be able to answer such questions.

To this end, let $\mathcal{A} := \{A_1,\dots,A_m\}$ be a family of borelian subsets, $p_{A_j} > 0$ for every $j \in \{1,\dots,m\}$. Moreover, denote by K_j, K'_j, $j \in \{1,\dots,m\}$ some subsets of indices in $I = \{1,\dots,p\}$. As above, we can deduce the asymptotic law of

$$\sqrt{n}\big(\hat\rho_{K_1,K'_1} - \rho_{K_1,K'_1},\dots,\hat\rho_{K_m,K'_m} - \rho_{K_m,K'_m}\big)$$

from Theorem 20. As an application, let us consider tests of the null assumption

$$\mathcal{H}_0 : C(\cdot|A) \text{ does not depend on } A \in \mathcal{A}, \text{ or equivalently}$$

$$\mathcal{H}_0 : C(\mathbf{u}|A_1) = \cdots = C(\mathbf{u}|A_m) \text{ for every } \mathbf{u}\in[0,1]^p,$$

against its opposite. This can be tackled through any generalized measure of association $\hat\rho_{K,K'}(A)$, for some fixed subsets K and K', and one or several functions ψ. To simplify the discussion, we consider hereafter a single map ψ. Thus, we can build a test statistic in the form of

$$\mathcal{T}_n := \|(i,j) \mapsto \sqrt{n}\big(\hat\rho_{K,K'}(A_i) - \hat\rho_{K,K'}(A_j)\big)\|,$$

where $\|\cdot\|$ is any semi-norm on $\mathbb{R}^{m^2}$. For example, define the Cramer-von Mises type statistic

$$\mathcal{T}_{n,CvM} := n\sum_{i,j=1}^m \big(\hat\rho_{K,K'}(A_i) - \hat\rho_{K,K'}(A_j)\big)^2,$$

or the Kolmogorov-Smirnov type test statistic

$$\mathcal{T}_{n,KS} := \sqrt{n}\max_{i,j=1,\dots,m}\big|\hat\rho_{K,K'}(A_i) - \hat\rho_{K,K'}(A_j)\big|.$$

Note that under the null hypothesis, these test statistics can be rewritten as

$$\begin{aligned}\mathcal{T}_n &= \|(i,j) \mapsto \sqrt{n}\big\{\hat\rho_{K,K'}(A_i) - \rho_{K,K'}(A_i) + \rho_{K,K'}(A_j) - \hat\rho_{K,K'}(A_j)\big\}\| \\ &= \|(i,j) \mapsto \sqrt{n}\big(\hat\rho_{K,K'}(A_i) - \rho_{K,K'}(A_i)\big) - \sqrt{n}\big(\hat\rho_{K,K'}(A_j) - \rho_{K,K'}(A_j)\big)\|.\end{aligned}$$

Therefore, under $\mathcal{H}_0$, Theorem 20 tells us that $\mathcal{T}_n$ (once properly rescaled) is weakly convergent. Since its limiting law is complex, we advise to use bootstrap approximations to evaluate the asymptotic p-values associated with $\mathcal{T}_n$ in practice. A bootstrapped version of such tests statistics is

$$\mathcal{T}^*_n := \|(i,j) \mapsto \sqrt{n}\big\{\hat\rho^*_{K,K'}(A_i) - \hat\rho_{K,K'}(A_i) + \hat\rho_{K,K'}(A_j) - \hat\rho^*_{K,K'}(A_j)\big\}\|,$$

where, in the case of the multiplier bootstrap, we set

$$\hat{\rho}^*_{K,K'}(A) := \int \psi(\mathbf{u})\, \widetilde{C}^*_{n,K}(\mathbf{u}_K|\mathbf{Z} \in A)\widetilde{C}^*_{n,K'}(d\mathbf{u}_{K'}|\mathbf{Z} \in A)\, d\mathbf{u}_{I\setminus K'},$$

and, in the case of the nonparametric bootstrap,

$$\hat{\rho}^*_{K,K'}(A) := \int \psi(\mathbf{u})\, \overline{C}^*_{n,K}(\mathbf{u}_K|\mathbf{Z} \in A)\overline{C}^*_{n,K'}(d\mathbf{u}_{K'}|\mathbf{Z} \in A)\, d\mathbf{u}_{I\setminus K'}.$$

Under the assumptions of Corollary 14 (resp. Theorem 17) and those of Theorem 20, the couple $\big(\mathcal{T}_{n,CvM}, \mathcal{T}^*_{n,CvM}\big)$ weakly converges to a couple of identically distributed vectors when n tends to the infinity, using the exchangeable (resp. nonparametric) bootstrap. And the same result applies to $\mathcal{T}_{n,KS}$.

5. Conclusion

We have made several contributions to the theory of the weak convergence of empirical copula processes, their associated bootstrap schemes and multivariate measures of association. Now, all these concepts and results are stated not only for usual copulas but for conditional copulas too, i.e., for the copula of $\mathbf{X}$ knowing that some vector of covariates $\mathbf{Z}$ (that may be equal to $\mathbf{X}$) belongs to one or several borelian subsets. We only require that the probabilities of the latter events are nonzero. Working with $\mathbf{Z}$-subsets instead of singletons allows to avoid the curse of dimension that rapidly appears when the dimension of $\mathbf{Z}$ is larger than three.

We have proved the weak convergence of the conditional empirical copula process $\hat{\mathbb{C}}_n(\cdot|A)$, indexed by a family of borelian subsets $A \in \mathcal{A}$, and we have explicitly tackled the case of random subsets. Therefore, inference and testing of copula models becomes relatively easy. An interesting avenue for further research will be to use our results to build convenient discretizations of the covariate space (the space of our so-called random vectors $\mathbf{Z}$). There is a need to find efficient algorithms and statistical procedures to build a partition of $\mathbb{R}^q$ with borelian subsets A_j, so that the dependencies across the components of $\mathbf{X}$ are “similar” when $\mathbf{Z}$ belongs to one of theses subsets, but as different as possible from box to box: “maximum homogeneity intra, maximum heterogeneity inter”. A constructive tree-based approach should be feasible, as proposed in [36] in the case of vine modeling, and is left for a further study.

Acknowledgments

The authors warmly thank the Associate Editor and two reviewers whose numerous remarks and suggestions allowed us to strongly improve the paper.

Supplementary Material

Real data application and proofs
(doi: 10.1214/22-EJS2075SUPP; .pdf). In Section A, we show an application to the dependence between financial returns. The proofs for all results are detailed in Section B. Finally, we give the covariance function of $\mathbb{C}_\infty$ in Section C.

References

[1] Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance Math. Econom.*, **44(2)**, 182–198. MR2517884

[2] Abegaz, F., Gijbels, I. and Veraverbeke, N. (2012). Semiparametric estimation of conditional copulas. *J. Multivariate Anal.*, **110**, 43–73. MR2927509

[3] Acar, E.F., Craiu, R.V. and Yao, F. (2011). Dependence Calibration in Conditional copulas: A Nonparametric Approach. *Biometrics*, **67**, 445–453. MR2829013

[4] Acar, E.F., Craiu, R.V. and Yao, F. (2013). Statistical testing of covariate effects in conditional copula models. *Electron. J. Stat.*, **7**, 2822–2850. MR3148369

[5] Bedford, T. and Cooke, R.M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann. Math. Artif. Intell.*, **32(1-4)**, 245–268. MR1859866

[6] Bedford, T. and Cooke, R.M. (2002). Vines: A new graphical model for dependent random variables. *Ann. Statist.*, 1031–1068. MR1926167

[7] Brechmann, E.C., Hendrich, K. and Czado, C. (2013). Conditional copula simulation for systemic risk stress testing. *Insurance Math. Econom.*, **53(3)**, 722–732. MR3130467

[8] Bücher, A., and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statistics & probability letters*, **80**(23-24), 1925–1932. MR2734261

[9] Bücher, A., and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis*, **119**, 61–70. MR3061415

[10] Bücher, A. and Kojadinovic, I. (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *J. Theoret. Probab.*, **32(3)**, 1145–1165. MR3979663

[11] Christoffersen, P., Errunza, V., Jacobs, K. and Langlois, H. (2012). Is the potential for international diversification disappearing? A dynamic copula approach. *The Review of Financial Studies*, **25(12)**, 3711–3751.

[12] Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas*. Lecture Notes in Statistics, Springer. MR3931334

[13] Derumigny, A. and Fermanian, J.-D. (2017). About tests of the "simplifying" assumption for conditional copulas. *Depend. Model.*, **5(1)**, 154–197. MR3694366

[14] Derumigny, A. and Fermanian, J.-D. (2022). Supplement to "Conditional empirical copula processes and generalized measures of association" DOI:10.1214/22-EJS2075SUPP. MR4015246

[15] Derumigny, A. and Fermanian, J.-D. (2019). On kernel-based estimation of conditional Kendall's tau: finite-distance bounds and asymptotic behavior. *Depend. Model.*, **7(1)**, 292–321 MR4015246

[16] Derumigny, A. and Fermanian, J.-D. (2020). On Kendall's regression. To appear in *J. Multivariate Anal.* MR4080864

[17] Durante, F. and Jaworski, P. (2010). Spatial contagion between financial markets: a copula-based approach. *Appl. Stoch. Models Bus. Ind.*, **26(5)**, 551–564. MR2760759

[18] Durante, F., Pappadà, R. and Torelli, N. (2014). Clustering of financial time series in risky scenarios. *Adv. Data Anal. Classif.*, **8(4)**, 359–376. MR3277832

[19] Fang, Y. and Madsen, L. (2013). Modified Gaussian pseudo-copula: Applications in insurance and finance. *Insurance Math. Econom.*, **53(1)**, 292–301. MR3081481

[20] Fermanian, J.-D., Radulovic, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, **10(5)**, 847–860. MR2093613

[21] Fermanian J.-D. and Wegkamp, M. (2012). Time-dependent copulas. *J. Multivariate Anal.*. **110**, 19–29. MR2927507

[22] Fermanian J.-D. and Lopez, O. (2018). Single-index copulas. *J. Multivariate Anal.*, **165**, 27–55. MR3768751

[23] Genest, C., Nešlehová, J. and Ben Ghorbal, N. (2011). Estimators based on Kendall's tau in multivariate copula models. *Aust. N.Z. J. Stat.*, **53**, 157–177. MR2851720

[24] Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulas, association measures and their applications. *Comput. Statist. Data Anal.*, **55**, 1919–1932. MR2765054

[25] Gijbels, I., Veraverbeke, N. and Omelka, M. (2015a). Estimation of a Copula when a Covariate Affects only Marginal Distributions. *Scand. J. Stat.*, **42**, 1109–1126. MR3426313

[26] Gijbels, I., Omelka, M. and Veraverbeke, N. (2017). Nonparametric testing for no covariate effects in conditional copulas. *Statistics*, **51(3)**, 475–509. MR3630461

[27] Goodwin, B.K. and Hungerford, A. (2015). Copula-based models of systemic risk in US agriculture: implications for crop insurance and reinsurance contracts. *American Journal of Agricultural Economics*, **97(3)**, 879–896.

[28] Hesami Afshar, M., Sorman, A.U. and Yilmaz, M.T. (2016). Conditional copula-based spatial–temporal drought characteristics analysis—a case study over Turkey. *Water*, **8(10)**, 426.

[29] Hobæk Haff, I., Aas, K. and Frigessi, A. (2010). On the simplified pair-copula construction–simply useful or too simplistic? *J. Multivariate Anal.*, **101**, 1296–1310. MR2595309

[30] Joe, H. (1990). Multivariate concordance, *J. Multivariate Anal.*, **35**, 12–30. MR1084939

[31] Jondeau, E. and Rockinger, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. *J. Internat. Money Finance*, **25**, 827–853.

[32] Kendall, M.G. and Babington Smith, B. (1940). On the method of paired comparisons. *Biometrika*, **31**, 324–345. MR0002761

[33] Kim, J.Y., Park, C.Y. and Kwon, H.H. (2016). A development of downscaling scheme for sub-daily extreme precipitation using conditional copula model. *Journal of Korea Water Resources Association*, **49(10)**, 863–876.

[34] Kojadinovic, I., Segers, J. and Yan, J. (2011). Large sample tests of extreme value dependence for multivariate copulas. *Canad. J. Statist.*, **39(4)**, 703–720. MR2860835

[35] Kosorok, M.R. (2007). *Introduction to empirical processes and semiparametric inference*. Springer Science. MR2724368

[36] Kurz, M.S. and Spanhel, F. (2017). Testing the simplifying assumption in high-dimensional vine copulas. arXiv:1706.02338 MR4492989

[37] Manner, H. and Reznikova, O. (2012). A survey on time-varying copulas: specification, simulations, and application. *Econometric reviews*, **31(6)**, 654–687. MR2903315

[38] Nelsen, R.B. (1999). *An introduction to copulas*, Lecture Notes in Statistics, vol. **139**. Springer-Verlag, New York. MR1653203

[39] Nelsen, R.B. (2002). Concordance and copulas: A survey. In C. M. Cuadras, J. Fortiana, J. A. Rodriguez-Lallena (Eds.), *Distributions with given marginals and statistical modelling* (pp.16–177) Dordrecht: Kluwer. MR2058990

[40] Oh, D.H. and Patton, A.J. (2018). Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads. *Journal of Business & Economic Statistics*, **36(2)**, 181–195. MR3790207

[41] Omelka, M., Veraverbeke, N. and Gijbels, I. (2013). Bootstrapping the conditional copula. *J. Statist. Plann. Inference*, **143**, 1–23. MR2969007

[42] Palaro, H.P. and Hotta, L.K. (2006). Using conditional copula to estimate value at risk. *Journal of Data Science*, **4**, 93–115.

[43] Patton, A. (2006a) Modelling Asymmetric Exchange Rate Dependence, *Internat. Econom. Rev.*, **47**, 527–556. MR2216591

[44] Patton, A. (2006b) Estimation of multivariate models for time series of possibly different lengths. *J. Appl. Econometrics*, **21**, 147–173. MR2226593

[45] Patton, A.J. (2009). Copula–based models for financial time series. In *Handbook of financial time series* (pp. 767–785). Springer, Berlin, Heidelberg.

[46] Patton, A.J. (2012). A review of copula models for economic time series. *J. Multivariate Anal.*, **110**, 4–18. MR2927506

[47] Portier, F. and Segers, J. (2018). On the weak convergence of the empirical conditional copula under a simplifying assumption. *J. Multivariate Anal.*, **166**, 160–181. MR3799641

[48] Radulović, D., Wegkamp M. and Zhao, Y. (2017). Weak convergence of empirical copula processes indexed by functions. *Bernoulli*, **23(8)**, 3346–3384. MR3654809

[49] Rémillard, B. and Scaillet, O. (2009). Testing for Equality between Two copulas. *J. Multivariate Anal.*, **100**, 377–386. MR2483426

[50] Ruymgaart, F.H. and van Zuijlen, M.C.A. (1978). Asymptotic normality of multivariate linear rank statistics in the non-iid case. *Ann. Statist.*, 588–602. MR0464489

[51] Scarsini, M. (1984). On measures of concordance. *Stochastica*, **8(3)**, 201–218. MR0796650

[52] Schellhase, C. and Spanhel, F. (2018). Estimating non-simplified vine copulas using penalized splines. *Stat. Comput.*, **28(2)**, 387–409. MR3747570

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

[54] Schmid, F. and Schmidt, R. (2007). Nonparametric inference on multivariate versions of Blomqvist's beta and related measures of tail dependence. *Metrika*, **66(3)**, 323–354. MR2336484

[55] Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *J. Multivariate Anal.*, **98(6)**, 1123–1140. MR2326243

[56] Schmid, F., Schmidt, R., Blumentritt, T., Gaißer, S. and Ruppert, M. (2010). Copula-based measures of multivariate association. In *Copula theory and its applications* (pp. 209–236). Springer, Berlin, Heidelberg. MR3051270

[57] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, **18(3)**, 764–782. MR2948900

[58] Shorack, G.R. and Wellner, J.A. (2009). *Empirical processes with applications to statistics*. Society for Industrial and Applied Mathematics. MR3396731

[59] Spanhel, F. and Kurz, M.S. (2017). The partial vine copula: A dependence measure and approximation based on the simplifying assumption. arXiv:1510-06971. MR3935849

[60] Spanhel, F. and Kurz, M.S. (2019). Simplified vine copula models: Approximations based on the simplifying assumption. *Electron. J. Stat.*, **13(1)**, 1254–1291. MR3935849

[61] Taylor, M.D. (2007). Multivariate measures of concordance. *Ann. Inst. Statist. Math.*, **59(4)**, 789–806. MR2397737

[62] van der Vaart, A. (1998). *Asymptotic statistics*. Cambridge University Press. MR1652247

[63] van der Vaart, A., and Wellner, J. (1996). *Weak convergence and empirical processes*. Springer. MR1385671

[64] Vatter, T. and Chavez-Demoulin, V. (2015). Generalized additive models for conditional dependence structures. *J. Multivariate Anal.*, **141**, 147–167. MR3390064

[65] Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a Conditional Copula and Association Measures. *Scand. J. Stat.*, **38**, 766–780. MR2859749

[66] Wellner, J., and van der Vaart, A. (2007). Empirical processes indexed

by estimated functions. In *Asymptotics: particles, processes and inverse problems* (pp. 234–252). Institute of Mathematical Statistics. MR2459942

[67] Wolff, E.F. (1980). N-dimensional measures of dependence. *Stochastica*, **4(3)**, 175–188. MR0611502

[68] Wuertz, D. et al. (2020). fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling. R package version 3042.83.2. `https://CRAN.R-project.org/package=fGarch`.