Introduction

Management of environmental ressources often requires the analysis of multivariate extreme values. In climate studies, extreme events represent a major challenge due to their consequences. As the volume of data expands, the problem of missing data is present in many fields above all in environmental research (see [Xia et al., 1999]), usually due to instrument errors, communication and processing errors. In a time series setting, the observation periods of a multivariate series could be different and overlap only partially. Rigourous inference methods for assessing extremal depencies which handle missing values are thus in demand. In this paper, we are particularly interested of the dependence structure of multivariate extreme value distribution. This concept is defined as follows.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathbf{X} = (X_1, \dots, X_d)$ be a d-dimensional random vector with values in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $d \geq 2$. This random vector has a joint distribution function F and its margins are denoted by $F_j(x) = \mathbb{P}\{X_j \leq x\}$ for all $x \in \mathbb{R}$ and $j \in \{1, \dots, d\}$. A function $C : [0,1]^d \rightarrow [0,1]$ is called a d-dimensional copula if it is the restriction to $[0,1]^d$ of a distribution function whose margins are given by the uniform distribution on the interval $[0,1]$. Since the work of [Sklar, 1959], it is well known that every distribution function F can be decomposed as $F(\mathbf{x}) = C(F_1(x_d), \dots, F_d(x_d))$, for all $\mathbf{x} \in \mathbb{R}^d$. Under the framework of extreme, the notion of copulas leads to the so-called extreme value copulas (see [Gudendorf and Segers, 2010])

$$C(\mathbf{u}) = exp\left(-\ell(-log(u_1), \dots, -log(u_d)\right), \quad \mathbf{u} \in (0,1]^d, \tag{1}$$

with $\ell : [0,\infty)^d \rightarrow [0,\infty)$ the stable tail dependence function. The tail dependence function ℓ is convex, homogeneous of order one, that is $\ell(cx_1, \dots, cx_d) = c\ell(x_1, \dots, x_d)$ for $c > 0$ and satisfies $\max(x_1, \dots, x_d) \leq \ell(x_1, \dots, x_d) \leq x_1 + \dots + x_d$, $\forall (x_1, \dots, x_d) \in [0,\infty)^d$. By homogeneity, it is characterized by the *Pickands dependence function* $A : \Delta^{d-1} \rightarrow [1/d, 1]$, which is the restriction of ℓ to the unit simplex Δ^{d-1} :

$$\ell(x_1, \dots, x_d) = (x_1 + \dots + x_d)A(w_1, \dots, w_d), \quad w_j = \frac{x_j}{x_1 + \dots + x_d}, \tag{2}$$

for $(x_1, \dots, x_d) \in [0,\infty)^d \setminus \{0\}$. Notice that, for every $\mathbf{w} \in \Delta^{d-1}$

$$C(u^{w_1}, \dots, u^{w_d}) = u^{A(\mathbf{w})}. \tag{3}$$

Based on the madogram concept from geostatistics, [Naveau et al., 2009] introduced the λ-madogram in order to capture bivariate extremal dependencies. This quantity leads to its extension in higher dimension the $\mathbf{w}$-madogram defined in [Marcon et al., 2017]

$$\nu(\mathbf{w}) = \mathbb{E}\left[\bigvee_{j=1}^{d} \{F_j(X_j)\}^{1/w_j} - \frac{1}{d}\sum_{j=1}^{d} \{F_j(X_j)\}^{1/w_j}\right], \tag{4}$$

if $w_j = 0$ and $0 < u < 1$, then $u^{1/w_j} = 0$ by convention.

Paragraphe sur l'apport du papier + insertion dans la littérature (revue de littérature et contribution)

Paragraphe annonce du plan

In order to shorten formulas, notations

$$\begin{aligned}
&\mathbf{u}_j(t) := (u_1, \dots, u_{j-1}, t, u_{j+1}, \dots, u_d), \\
&\mathbf{u}_{jk}(s,t) := (u_1, \dots, u_{j-1}, s, u_{j+1}, \dots, u_{k-1}, t, u_{k+1}, \dots, u_d),
\end{aligned}$$

will be adopted for $s, t \in [0,1]$, $(u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_d) \in [0,1]^{d-1}$ and $j, k \in \{1, \dots, d\}$ with $j < k$.

Also, the following notations are used. Given $\mathcal{X} \subset \mathbb{R}^2$, let $\ell^\infty(\mathcal{X})$ denote the spaces of bounded real-valued function on $\mathcal{X}$. For $f : \mathcal{X} \to \mathbb{R}$, let $||f||_\infty = sup_{x \in \mathcal{X}}|f(x)|$. Here, we use the abbreviation $Q(f) = \int f dQ$ for a given measurable function f and signed measure Q. The arrows $\overset{a.s.}{\to}$, $\overset{d}{\to}$ denote almost sure convergence and convergence in distribution of random vectors. Weak convergence of a sequence of maps will be understood in the sense of J.Hoffman-Jørgensen (see Part 1 in the monograph by [van der Vaart and Wellner, 1996]). Given that $n \in \mathbb{N}^*, X, X_n$ are maps from $(\Omega, \mathcal{A}, \mathbb{P})$ into a metric space $\mathcal{X}$ and that X is Borel measurable, $(X_n)_{n \geq 1}$ is said to converge weakly to X if $\mathbb{E}^* f(X_n) \to \mathbb{E} f(X)$ for every bounded continuous real-valued function f defined on $\mathcal{X}$, where $\mathbb{E}^*$ denotes outer expectation in the event that X_n may not be Borel measurable. In what follows, weak convergence is denoted by $X_n \rightsquigarrow X$.

1 Non parametric estimation of the Madogram with missing data

Under the notation of the introduction, we assume that the copula C is of extreme value type as in Equation 1. Starting from independent and identically distributed $i.i.d.$ copies $\mathbf{X}_1, \dots, \mathbf{X}_n$ of $\mathbf{X}$, suppose we observe a $2d$-tuple such as

$$(\mathbf{I}_i \mathbf{X}_i, \mathbf{X}_i), \quad i \in \{1, \dots, n\}, \tag{5}$$

where $\mathbf{I}_i\mathbf{X}_i = (X_{i,1}I_{i,1}, \dots, X_{i,d}I_{i,d})$ and $X_{i,j}$ is missing if $I_{i,j} = 0$, otherwise $I_{i,j} = 1$, *i.e.* at each $i \in \{1, \dots, n\}$, several entries may be missing. For $j \in \{1, \dots, d\}$, we suppose that for all $i \in \{1, \dots, n\}$, $I_{i,j}$ are sampled from a Bernoulli random variable I_j with probability $p_j = \mathbb{P}(I_j = 1)$. We denote by $\mathbf{p}$ the probability of observing completely a realization from $\mathbf{X}$, that is $\mathbf{p} = \mathbb{P}(I_1 = 1, \dots, I_d = 1)$. Let us now define the empirical cumulative distribution of X_j (resp. F) in case of missing data,

$$\hat{F}_{n,j}(x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{X_{i,j} \leq x\}} I_{i,j}}{\sum_{i=1}^n I_{i,j}}, \forall x \in \mathbb{R}, \quad \hat{F}_n(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbb{1}_{\{\mathbf{X}_i \leq \mathbf{x}\}} \Pi_{j=1}^d I_{i,j}}{\sum_{i=1}^n \Pi_{j=1}^d I_{i,j}}, \forall \mathbf{x} \in \mathbb{R}^d. \quad (6)$$

The idea raised here is to estimate separately margins with the complete information given by the realizations of X_j. We thus estimate the $\mathbf{w}$-madogram under the complete database. We recall the definition of the *hybrid copula estimator* introduced by [Segers, 2015]

$$\hat{C}_n^{\mathcal{H}}(\mathbf{u}) = \hat{F}_n(\hat{F}_{n,1}^{\leftarrow}(u_1), \dots, \hat{F}_{n,d}^{\leftarrow}(u_d)), \quad \mathbf{u} \in [0,1]^d.$$

Here, we write the generalized inverse function of F as $F^{\leftarrow}(u) = \inf\{x \in \mathbb{R} | F(x) \geq u\}$ where $0 < u < 1$. The normalized estimation error of the hybrid copula estimator is

$$\mathbb{C}_n^{\mathcal{H}}(\mathbf{u}) = \sqrt{n}\left(\hat{C}_n^{\mathcal{H}}(\mathbf{u}) - C(\mathbf{u})\right), \quad \mathbf{u} \in [0,1]^d.$$

On the condition that the first-order partial derivatives of the copula function C exists and are continuous on a subset of the unit hypercube, [Segers, 2012] obtained weak convergence of the normalized estimation error of the empirical copula process. To satisfy this condition, we introduce the following assumption as suggested in [Segers, 2012] in Example 5.3.

Assumption A.

(i) The distribution function F has continuous margins $F_1, \dots, F_d$.

(ii) For every $j \in \{1, \dots, d\}$, the first-order partial derivative $\dot{\ell}_j$ of ℓ with respect to x_j exists and is continuous on the set $\{x \in [0,\infty)^d : x_j > 0\}$.

The Assumption A *(i)* guarantees that the representation $F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$ is unique on the range of $(F_1, \dots, F_d)$. Under the Assumption A *(ii)*, the first-order partial derivatives of C with respect to u_j denoted as $\dot{C}_j$ exists and are continuous on the set $\{\mathbf{u} \in [0,1]^d : 0 < u_j < 1\}$. We now define our estimator of Equation (4) in the general context (allowing missing data).

Definition 1. *Let $(\boldsymbol{I}_i\boldsymbol{X}_i, \boldsymbol{I}_i)_{i=1}^n$ be a sample given by Equation (5), we define the hybrid* *estimator of the* $\mathbf{w}$*-madogram by*

$$\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) = \frac{1}{\sum_{i=1}^n \Pi_{j=1}^d I_{i,j}} \sum_{i=1}^n \left[\bigvee_{j=1}^d \left\{ \hat{F}_{n,j}(X_{i,j}) \right\}^{1/w_j} - \frac{1}{d} \sum_{j=1}^d \left\{ \hat{F}_{n,j}(X_{i,j}) \right\}^{1/w_j} \right] \Pi_{j=1}^d I_{i,j}, \tag{7}$$

where $\hat{F}_{n,j}(x)$ *are defined on Equation (6).*

The intuitive idea here is to estimate the margins by the complete series for each variables but estimate $\nu(\mathbf{w})$ only based on the time period where all series were recorded simultaneously. One may verify that in the complete data framework, *i.e.* when $p_j = 1, \forall j \in \{1, \dots, d\}$ and $\mathbf{p} = \mathbf{1}$ we retrieve the $\mathbf{w}$-madogram such as defined in [Marcon et al., 2017], namely

$$\hat{\nu}_n(\mathbf{w}) = \frac{1}{n} \sum_{i=1}^n \left[\bigvee_{j=1}^d \left\{ \hat{F}_{n,j}(X_{i,j}) \right\}^{1/w_j} - \frac{1}{d} \sum_{j=1}^d \left\{ \hat{F}_{n,j}(X_{i,j}) \right\}^{1/w_j} \right],$$

with $\hat{F}_{n,j}(x)$ the empirical cumulative distribution function of X_j.

Remark 1. *Our estimator defined in Equation (7) does not verify* $\hat{\nu}_T^{\mathcal{H}}(\boldsymbol{e}_j) = (d-1)/2d$ *while* $\nu(\boldsymbol{e}_j)$ *does where* $\boldsymbol{e}_j$ *is the jth vector of the canonical basis. In addition, the variance at* $\boldsymbol{e}_j$ *does not equal 0. Indeed, suppose that we evaluate this statistic at* $\mathbf{w} = \boldsymbol{e}_j$*, we thus obtain the following quantity :*

$$\hat{\nu}_n^{\mathcal{H}}(\boldsymbol{e}_j) = \frac{1}{\sum_{i=1}^n \Pi_{j=1}^d I_{i,j}} \sum_{i=1}^n \left[\hat{F}_{n,j}(X_{i,j}) - \frac{1}{d} \hat{F}_{n,j}(X_{i,j}) \right] \Pi_{j=1}^d I_{i,j}.$$

In this situation, the sample $(X_{i,1}, \dots, X_{i,j-1}, X_{i,j+1}, \dots, X_{i,d})_{i=1}^n$ *is taken into account through the indicators sequence* $(I_{i,1}, \dots, I_{i,j-1}, I_{i,j+1}, \dots, I_{i,d})_{i=1}^n$ *and induce a supplementary variance when estimating.*

We can force our estimator as in [Naveau et al., 2009] to satisfy these endpoint conditions. Given continuous functions $\lambda_1, \dots, \lambda_d : \Delta^{d-1} \to \mathbb{R}$ verifying $\lambda_j(\mathbf{e}_k) = \delta_{jk}$ (the Kronecker delta) for $j, k \in \{1, \dots, d\}$. This leads to a slightly modified version of the $\mathbf{w}$-madogram.

Definition 2. *Under the notation of Definition 1, we define the hybrid corrected estimator of the* $\mathbf{w}$*-madogram by*

$$\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) = \hat{\nu}_n(\mathbf{w}) - \sum_{j=1}^d \frac{\lambda_j(\mathbf{w})(d-1)}{d} \left[\frac{1}{\sum_{i=1}^n \Pi_{j=1}^d I_{i,j}} \sum_{i=1}^n \left\{ \hat{F}_{n,j}(X_{i,j}) \right\}^{1/w_j} \Pi_{j=1}^d I_{i,j} - \frac{w_j}{1+w_j} \right]. \tag{8}$$

Remark 2. *One has often that endpoint corrections does not have an impact to the asymp-*

totic behavior with complete data framework and unknown margins (see Section 2.3 and 2.4 of [Genest and Segers, 2009]). That is not always the case in the missing data framework and this feature is somehow wanted as we have discussed in Remark 1.

Let us now introduce a condition on the missing mechanism :

Assumption B. *We suppose that for all $i \in \{1, \dots, n\}$, the vector $\boldsymbol{I}_i$ and $\boldsymbol{X}_i$ are independent, the data are missing completely at random (**MCAR**). Furthermore, we suppose that there exists at least one $i \in \{1, \dots, n\}$ such that $\Pi_{j=1}^d I_{i,j} \neq 0$.*

Under this Assumption, we state the strong consistency of our hybrid estimator of the $\mathbf{w}$-madogram.

Proposition 1 (**Strong consistency**). *Let $(\boldsymbol{I}_i \boldsymbol{X}_i, \boldsymbol{X}_i)_{i=1}^n$ a i.i.d sample given by Equation (5). We have, under Assumption B for a fixed $\mathbf{w} \in \Delta^{d-1}$, as $n \rightarrow \infty$*

$$\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) \xrightarrow{a.s.} \nu(\mathbf{w}), \quad \hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) \xrightarrow{a.s.} \nu(\mathbf{w}).$$

Details of the proof are given in Section 2. We present with Theorem 1 our main result concerning the weak convergence of the following processes

$$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}}(\lambda) - \nu(\lambda)\right), \quad \sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right). \tag{9}$$

Without missing data, the weak convergence of the normalized estimation error of the empirical copula process has been proved by [Fermanian et al., 2004] under a more restrictive condition than Assumption A. With the hybrid copula estimator, we need a technical assumption in order to guarantee the weak convergence of the process $\mathbb{C}_T^{\mathcal{H}}$ (see [Segers, 2015]). Before spelled it, we note for convenience the marginal distribution and quantile functions into vector valued functions $\mathbf{F}$ and $\mathbf{F}^{\leftarrow}$:

$$\mathbf{F}(\mathbf{x}) = (F_1(x_1), \dots, F_d(x_d)), \quad \mathbf{x} \in \mathbb{R}^d, \quad \mathbf{F}^{\leftarrow}(\mathbf{u}) = (F_1^{\leftarrow}(u_1), \dots, F^{\leftarrow}(u_d)), \quad \mathbf{u} \in [0,1]^d.$$

Assumption C. *In the space $\ell^\infty(\mathbb{R}^d) \otimes (\ell^\infty(\mathbb{R}), \dots, \ell^\infty(\mathbb{R}))$ equipped with the topology of uniform convergence, we have the joint weak convergence*

$$\begin{aligned}&\left(\sqrt{n}(\hat{F}_n - F); \sqrt{n}(\hat{F}_{n,1} - F_1), \dots, \sqrt{n}(\hat{F}_{n,d} - F_d)\right) \\ &\qquad \rightsquigarrow \left(\alpha \circ \boldsymbol{F}, \beta_1 \circ F_1, \dots, \beta_d \circ F_d\right).\end{aligned}$$

The stochastic processes α and $\beta_j, j \in \{1, \dots, d\}$ take values in $\ell^\infty([0,1]^d)$ and $\ell^\infty([0,1])$ respectively, and are such that $\alpha \circ F$ and $\beta_j \circ F_j$ have continuous trajectories on $[-\infty, \infty]^d$ and $[-\infty, \infty]$ almost surely.

Under Assumptions A and C, the stochastic process $\mathbb{C}_T^{\mathcal{H}}$ converges weakly to the tight Gaussian process S_C defined by,

$$S_C(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{j=1}^{d} \dot{C}_j(\mathbf{u})\beta_j(u_j), \quad \forall \mathbf{u} \in [0,1]^d.$$

Considering the same statistical framework and missing mechanism as [Segers, 2015] (in Example 3.5) but in higher dimension, we show that the processes α, β_j takes the following closed form

$$\begin{aligned}
\beta_j(u_j) &= p_j^{-1}\mathbb{G}\left(\mathbb{1}_{\{X_j \leq F_j^{\leftarrow}(u_j), I_j=1\}} - u_j \mathbb{1}_{\{I_j=1\}}\right), \quad j \in \{1,\dots,d\}, \\
\alpha(\mathbf{u}) &= \mathbf{p}^{-1}\mathbb{G}\left(\mathbb{1}_{\{\mathbf{X} \leq \mathbf{F}^{\leftarrow}(\mathbf{u}), \mathbf{I}=\mathbf{1}\}} - C(\mathbf{u})\mathbb{1}_{\{\mathbf{I}=\mathbf{1}\}}\right),
\end{aligned}$$

where $\mathbb{G}$ is a tight Gaussian process. Furthermore, we are able to compute their covariance functions given in the following lemma.

Lemma 1. *The covariance function of the process $\beta_j(u_j)$, $\alpha(\boldsymbol{u})$ are, for $(\boldsymbol{u}, \boldsymbol{v}, v_k) \in [0,1]^{2d+1}$, and for $j \in \{1,\dots,d\}$ and $j < k$*

$$\begin{aligned}
cov\left(\beta_j(u_j), \beta_j(v_j)\right) &= p_j^{-1}\left(u_j \wedge v_j - u_j v_j\right), \\
cov\left(\beta_j(u_j), \beta_k(v_k)\right) &= \frac{p_{jk}}{p_j p_k}\left(C(\boldsymbol{1}_{j,k}(u_j, v_k)) - u_j v_k\right),
\end{aligned}$$

and

$$\begin{aligned}
cov\left(\alpha(\boldsymbol{u}), \alpha(\boldsymbol{v})\right) &= \mathbf{p}^{-1}\left(C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})\right), \\
cov\left(\alpha(\boldsymbol{u}), \beta_j(v_j)\right) &= p_j^{-1}\left(C(\boldsymbol{u}_j(u_j \wedge v_j)) - C(\boldsymbol{u})v_j\right).
\end{aligned}$$

Where we denote by $\mathbf{u} \wedge \mathbf{v}$ the vector of componentwise minima and $p_{jk} = \mathbb{P}(I_j = 1, I_k = 1)$. Proof of Lemma 1 is deferred to Section 2.

We have all tools in hand to consider the weak convergence of the stochastic processes in Equation (9).

Theorem 1 (**Functional central limit theorem with missing data**). *Under Assumptions A, B, C we have the weak convergence in $\ell^\infty(\Delta^{d-1})$ for the hybrid estimator defined* *in Equations (7) and (8), as $n \to \infty$,*

$$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w})\right) \rightsquigarrow \left(\frac{1}{d}\sum_{j=1}^{d}\int_{[0,1]} \alpha(\boldsymbol{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx \right.$$
$$\left. - \int_{[0,1]} S_C(x^{w_1}, \dots, x^{w_d})dx\right)_{\lambda\in[0,1]},$$

$$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) - \nu(\mathbf{w})\right) \rightsquigarrow \left(\frac{1}{d}\sum_{j=1}^{d}(1+\lambda_j(\mathbf{w})(d-1))\int_{[0,1]} \alpha(\boldsymbol{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx \right.$$
$$\left. - \int_{[0,1]} S_C(x^{w_1}, \dots, x^{w_d})dx\right)_{\lambda\in[0,1]}.$$

Proof is postponed in Section 2.

Ici[1], nous nous posons dans le cas de données complètes. Le cas général peut être déduit ensuite, mais il faut d'abord voir si le raisonnement est correct. Pour un $\mathbf{w} \in \Delta^{d-1}$ fixé, la loi de $\sqrt{n}(\hat{\nu}_n(\mathbf{w}) - \nu(\mathbf{w}))$ (estimateur sans données manquantes) suit une Gaussienne centrée (car transformation linéaire continue d'un processus Gaussien tendu) et sa variance est donnée par :

$$Var\left(\int_{[0,1]} N_C(u^{w_1}, \dots, u^{w_d})du\right)$$

Proposition 2. *Je pense avoir une forme close de la variance et celle-ci est décomposée comme suit :*

$$Var\left(\int_{[0,1]} N_C(u^{w_1}, \dots, u^{w_d})du\right) = \sigma_1^2(\boldsymbol{w}) + \sum_{j=1}^{d}\gamma_i^2(\boldsymbol{w}) - 2\sum_{j=1}^{d}\sigma_{1i}(\boldsymbol{w}) + 2\sum_{j<k}\gamma_{jk}(\boldsymbol{w}).$$

Technical details are available on Section 2.

2 Proof

Lemma 2. *We have, $\forall i \in \{1, \dots, n\}$*

$$\left|\bigvee_{j=1}^{d}\left\{\hat{F}_{n,j}(X_{i,j})\right\}^{1/w_j} - \bigvee_{j=1}^{d}\left\{F_j(X_j)\right\}^{1/w_j}\right| \le \sup_{j\in\{1,\dots,d\}}\left|\hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_j)^{1/w_j}\right|.$$

[1] J'écris en français tout paragraphes qui vont être modifiés

Proof The lemma becomes trivial once we write, $\forall i \in \{1,\dots,n\}$ and $j \in \{1,\dots,d\}$

$$\begin{aligned}\left\{\hat{F}_{n,j}(X_{i,j})\right\}^{1/w_j} &= F_j(X_j)^{1/w_j} + \hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_j)^{1/w_j},\\ &\leq F_j(X_j)^{1/w_j} + \sup_{j\in\{1,\dots,d\}} \left|\hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_j)^{1/w_j}\right|,\\ &\leq \bigvee_{j=1}^{d} \{F_j(X_j)\}^{1/w_j} + \sup_{j\in\{1,\dots,d\}} \left|\hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_j)^{1/w_j}\right|.\end{aligned}$$

Taking the max over $j \in \{1,\dots,d\}$ gives

$$\bigvee_{j=1}^{d}\left\{\hat{F}_{n,j}(X_{i,j})\right\}^{1/w_j} - \bigvee_{j=1}^{d}\{F_j(X_j)\}^{1/w_j} \leq \sup_{j\in\{1,\dots,d\}}\left|F_{n,j}(X_{i,j})^{1/w_j} - F_j(X_j)^{1/w_j}\right|.$$

Moreover, by symmetry of $\hat{F}_{n,j}$ and F_j, the second ones follows similarly. □

Proof of Proposition 1 We write, for notational convenience $n_i = \Pi_{j=1}^{d} I_{i,j}$ and $N = \sum_{i=1}^{n} n_i$. We prove it for $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$ as the strong consistency for $\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w})$ use the same arguments. The estimator $\hat{\nu}_n(\mathbf{w})$ is strongly consistent since it holds

$$\begin{aligned}|\hat{\nu}_n(\mathbf{w}) - \nu(\mathbf{w})| &= \left|\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu_n(\mathbf{w}) + \nu_n(\mathbf{w}) - \nu(\mathbf{w})\right|,\\ &\leq \left|\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu_n(\mathbf{w})\right| + |\nu_n(\mathbf{w}) - \nu(\mathbf{w})|,\end{aligned}$$

where

$$\nu_n(\mathbf{w}) = \frac{1}{N}\sum_{i=1}^{n}\left(\bigvee_{j=1}^{d}\{F_j(X_{i,j})\}^{1/w_j} - \frac{1}{d}\sum_{j=1}^{d}\{F_j(X_{i,j})\}^{1/w_j}\right)n_i,$$

By direct application of Assumption B and the law of large number, we have that

$$|\nu_n(\mathbf{w}) - \nu(\mathbf{w})| \overset{a.s.}{\to} 0$$

For the second term, we write :

$$\begin{aligned}|\hat{\nu}_n(\mathbf{w}) - \nu(\mathbf{w})| \leq& \frac{1}{N}\sum_{i=1}^{n}\left|\bigvee_{j=1}^{d}\hat{F}_{n,j}(X_{i,j})^{1/w_j} - \bigvee_{j=1}^{d}F_j(X_{i,j})^{1/w_j}\right| n_i\\ &+ \frac{1}{Nd}\sum_{i=1}^{n}\sum_{j=1}^{d}\left|\hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_{i,j})^{1/w_j}\right| n_j\\ \leq& 2\sup_{j\in\{1,\dots,d\}}\sup_{i\in\{1,\dots,n\}}\left|\hat{F}_{n,j}(X_{i,j})^{1/w_j} - F_j(X_{i,j})^{1/w_j}\right|,\end{aligned}$$

Where we used Lemma 2 to obtain the second inequality. The right term converges almost surely to zero by Glivencko-Cantelli. □

Proof of Lemma 1 Following [Segers, 2015] Example 3.5, we consider the function from $\{0,1\}^d \times \mathbb{R}^d$ into $\mathbb{R}$: for $\mathbf{x} \in \mathbb{R}^d$, and $j \in \{1, \dots, d\}$

$$f_j(\mathbf{I}, \mathbf{X}) = \mathbb{1}_{\{I_j=1\}}, \quad g_{j,x_j}(\mathbf{I}, \mathbf{X})\mathbb{1}_{\{X_j \le x_j, I_j=1\}},$$
$$f_{d+1} = \Pi_{j=1}^d f_j, \quad g_{d+1,\mathbf{x}} = \Pi_{j=1}^d g_{j,x_j}.$$

Let P denote the common distribution of the tuple $(\mathbf{I}, \mathbf{X})$. The collection of functions

$$\mathcal{F} = \{f_1, \dots, f_d, f_{d+1}\} \cup \bigcup_{j=1}^d \{g_{j,x_j}, x_j \in \mathbb{R}\} \cup \{g_{d+1,\mathbf{x}}, \mathbf{x} \in \mathbb{R}^d\}$$

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

$$\mathbb{G}_n(f) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n f(\mathbf{I}_i, \mathbf{X}_i) - \mathbb{E}[f(\mathbf{I}_i, \mathbf{X}_i)]\right), \quad f \in \mathcal{F},$$

converges in $\ell^\infty(\mathcal{F})$ to a P-browian bride $\mathbb{G}$. For $\mathbf{x} \in \mathbb{R}^d$,

$$\hat{F}_{n,j}(x_j) = \frac{p_j F_j(x_j) + n^{-1/2}\mathbb{G}_n g_{j,x_j}}{p_j + n^{-1/2}\mathbb{G}_n f_j},$$
$$\hat{F}_n(\mathbf{x}) = \frac{\mathbf{p}F(\mathbf{x}) + n^{-1/2}\mathbb{G}_n g_{d+1,\mathbf{x}}}{\mathbf{p} + n^{-1/2}\mathbb{G}_n f_{d+1}}$$

We obtain for the second one

$$\begin{aligned}\mathbf{p}\left(\hat{F}_n(\mathbf{x}) - F(\mathbf{x})\right) &= n^{-1/2}\left(\mathbb{G}_n(g_{d+1,\mathbf{x}}) - \hat{F}_n(\mathbf{x})\mathbb{G}_n(f_{d+1})\right), \\ &= n^{-1/2}\left(\mathbb{G}_n(g_{d+1,\mathbf{x}} - F(\mathbf{x})f_{d+1})\right) - n^{-1/2}\mathbb{G}_n(f_{d+1})(\hat{F}_n(\mathbf{x}) - F(\mathbf{x})).\end{aligned}$$

We thus have

$$\sqrt{n}\left(\hat{F}_n(\mathbf{x}) - F(\mathbf{x})\right) = \mathbf{p}^{-1}\left(\mathbb{G}_n(g_{d+1,\mathbf{x}} - F(\mathbf{x})f_{d+1})\right) - \mathbf{p}^{-1}\mathbb{G}_n(f_{d+1})(\hat{F}_n(\mathbf{x}) - F(\mathbf{x})).$$

Applying the central limit theorem gives that $\mathbb{G}_n(f_{d+1}) \xrightarrow{d} \mathcal{N}(0, \mathbb{P}(f_{d+1} - \mathbb{P}f_{d+1})^2)$, the law of large numbers gives also $\hat{F}_n(\mathbf{x}) - F(\mathbf{x}) = \circ_{\mathbb{P}}(1)$. Using Slutsky's lemma gives us

$$\sqrt{n}\left(\hat{F}_n(\mathbf{x}) - F(\mathbf{x})\right) = p^{-1}\left(\mathbb{G}_n(g_{d+1,\mathbf{x}} - F(\mathbf{x})f_{d+1})\right) + \circ_{\mathbb{P}}(1).$$

Similar reasoning might be applied to the margins, as a consequence, Condition B is fulfilled with for $\mathbf{u} \in [0,1]^d$,

$$\beta_j(u_j) = p_j^{-1}\mathbb{G}\left(g_{j,F_j^{\leftarrow}(u_j)} - u_j f_j\right),$$
$$\alpha(\mathbf{u}) = \mathbf{p}^{-1}\mathbb{G}\left(g_{d+1,\mathbf{F}^{\leftarrow}(\mathbf{u})} - C(\mathbf{u})f_{d+1}\right).$$

Let us compute one covariance function, the method still the same for the others, without loss of generality, suppose that $j < k$, we have for $u_j, v_k \in [0,1]$

$$\begin{aligned}
cov(\beta_j(u_j), \beta_k(v_k)) &= \mathbb{E}\left[p_j^{-1}\mathbb{G}\left(g_{j,F_j^{\leftarrow}(u_j)} - u_j f_j\right) p_k^{-1}\mathbb{G}\left(g_{k,F_k^{\leftarrow}(v_k)} - v_k f_k\right)\right], \\
&= \frac{1}{p_j p_k}\mathbb{E}\left[\mathbb{G}\left(g_{j,F_j^{\leftarrow}(u_i)} - u_j f_j\right)\mathbb{G}\left(g_{k,F_k^{\leftarrow}(v_j)} - v_k f_k\right)\right], \\
&= \frac{1}{p_j p_k}\mathbb{P}\left\{X_j \leq F_j^{\leftarrow}(u_j), X_k \leq F_k^{\leftarrow}(v_k), I_j = 1, I_k = 1\right\} - \frac{p_{jk}}{p_j p_k} u_j v_k, \\
&= \frac{1}{p_j p_k}\mathbb{P}\left\{X_j \leq F_j^{\leftarrow}(u_j), X_k \leq F_k^{\leftarrow}(v_k)\right\}\mathbb{P}\left\{I_j = 1, I_k = 1\right\} - \frac{p_{jk}}{p_j p_k} u_j v_k, \\
&= \frac{p_{jk}}{p_j p_k}\left(C(\mathbf{1}_{jk}(u_j, v_k)) - u_j v_k\right).
\end{aligned}$$

Hence the result. □

Proof of Theorem 1 We do the proof for $\nu_n^{\mathcal{H}*}$ as the proof for $\nu_n^{\mathcal{H}}$ is similar. Using that $\mathbb{E}[F_j(X_j)^\alpha] = (1+\alpha)^{-1}$ for $\alpha \neq 1$, we can write $\nu(\mathbf{w})$ as :

$$\begin{aligned}
\nu(\mathbf{w}) =& \mathbb{E}\left[\bigvee_{j=1}^d \{F_j(X_j)\}^{1/w_j} - \frac{1}{d}\sum_{j=1}^d \{F_j(X_j)\}^{1/w_j}\right] + \\
& \sum_{j=1}^d \frac{\lambda_j(\mathbf{w})(d-1)}{d}\left(\frac{w_j}{1+w_j} - \mathbb{E}\left[F_j(X_j)^{1/w_j}\right]\right), \\
=& \mathbb{E}\left[\bigvee_{j=1}^d \{F_j(X_j)\}^{1/w_j}\right] - \frac{1}{d}\sum_{j=1}^d (1+\lambda_j(\mathbf{w})(d-1))\mathbb{E}\left[F_j(X_j)^{1/w_j}\right] + a(\mathbf{w}),
\end{aligned}$$

with $a(\mathbf{w}) = (d-1)d^{-1}\sum_{j=1}^d \lambda_j(\mathbf{w})w_j/(1+w_j)$. Let us note by $g_{\mathbf{w}}$ the function defined as

$$g_{\mathbf{w}} : [0,1]^d \to [0,1], \quad \mathbf{u} \mapsto \bigvee_{j=1}^d u_j^{1/w_j} - \frac{1}{d}\sum_{j=1}^d (1+\lambda_j(\mathbf{w})(d-1))u_j^{1/w_j}.$$

We are to write our estimator of the $\mathbf{w}$-madogram and the $\mathbf{w}$-madogram in missing data framework as an integral with respect to the hybrid copula estimator and the copula function. We thus have:

$$\nu_n^{\mathcal{H}*}(\mathbf{w}) = \frac{1}{N}\sum_{i=1}^{n} g_{\mathbf{w}}\left(\hat{\mathbf{F}}_n(\mathbf{X}_i)\right)\Pi_{j=1}^{d} I_{i,j} + c(\mathbf{w}) = \int_{[0,1]^d} g_{\mathbf{w}}(\mathbf{u})\, d\hat{C}_n^{\mathcal{H}}(\mathbf{u}) + a(\mathbf{w}),$$
$$\nu(\mathbf{w}) = \int_{[0,1]^d} g_{\mathbf{w}}(\mathbf{u})\, dC(\mathbf{u}) + a(\mathbf{w}).$$

Where $\hat{\mathbf{F}}_n(\mathbf{X}_i) = (\hat{F}_{n,1}(X_{i,1}), \dots, \hat{F}_{n,d}(X_{i,d}))$. We obtain, proceeding as in Theorem 2.4 of [Marcon et al., 2017] :

$$\begin{aligned}\sqrt{n}\left(\nu_n^{\mathcal{H}*}(\mathbf{w}) - \nu(\mathbf{w})\right) =& \frac{1}{d}\sum_{j=1}^{d}\left(1+\lambda_j(\mathbf{w})(d-1)\right)\int_{[0,1]} \mathbb{C}_n^{\mathcal{H}}(\mathbf{1}_j(x^{w_j}))dx \\ &- \int_{[0,1]} \mathbb{C}_n^{\mathcal{H}}\left(x^{w_1}, \dots, x^{w_d}\right) dx.\end{aligned}$$

Consider the function $\phi : \ell^\infty([0,1]^d) \to \ell^\infty(\Delta^{d-1}), f \mapsto \phi(f)$, defined by

$$(\phi)(f)(\mathbf{w}) = \frac{1}{d}\sum_{j=1}^{d}(1+\lambda_j(\mathbf{w})(d-1))\int_{[0,1]} f(\mathbf{1}_j(x^{w_j}))dx - \int_{[0,1]} f(x^{w_1}, \dots, x^{w_d})dx.$$

this function is linear and bounded thus continuous. The continous mapping theorem (Theorem 1.3.6 of [van der Vaart and Wellner, 1996]) implies, as $n \to \infty$

$$\sqrt{n}(\hat{\nu}_n^{\mathcal{H}*} - \nu) = \phi(\mathbb{C}_n^{\mathcal{H}}) \rightsquigarrow \phi(S_C),$$

in $\ell^\infty(\Delta^{d-1})$. We note that $S_C(\mathbf{1}_j(x^{w_j})) = \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(u_j)$ and we obtain our statement. □

Lemma 3. *If $\ell(x_1, \dots, x_d)$ is homogeneous of degree 1, then for any $i \in \{1, \dots, d\}$ the partial derivative $\dot{\ell}_j(x_1, \dots, x_d)$ is homogeneous of degree 0.*

Proof of Proposition 2 We have $\forall j \in \{1, \dots, d\}$

$$\dot{C}_j(\mathbf{u}) = \frac{C(\mathbf{u})}{u_j}\dot{\ell}_j(-log(u_1), \dots, -log(u_d)).$$

Furthermore, using Lemma 3, we have

$$\begin{aligned}\dot{C}_j(u^{w_1}, \dots, u^{w_d}) &= \frac{u^{A(\mathbf{w})}}{u^{w_j}}\dot{\ell}_j(-w_1 log(u), \dots, -w_d log(u)) = \frac{u^{A(\mathbf{w})}}{u^{w_j}}\dot{\ell}_j(-w_1, \dots, -w_d) \\ &= \frac{u^{A(\mathbf{w})}}{u^{w_j}}\mu_j(\mathbf{w}).\end{aligned}$$

Now, let us compute

$$\sigma_1^2(\mathbf{w}) = \mathbb{E}\left[\int_{[0,1]} B_C(u^{w_1},\dots,u^{w_d})du \int_{[0,1]} B_C(v^{w_1},\dots,v^{w_d})dv\right].$$

Using linearity of the integral and the definition of the covariance function of B_C, we obtain

$$\sigma_1^2(\mathbf{w}) = 2\int_{[0,1]}\int_{[0,v]} u^{A(\mathbf{w})}(1-v^{A(\mathbf{w})})duv = \frac{1}{(1+A(\mathbf{w}))^2}\frac{A(\mathbf{w})}{2+A(\mathbf{w})}.$$

The quantity γ_j^2 is defined by the following

$$\gamma_j^2 = \mathbb{E}\bigg[\int_{[0,1]} B_C(\mathbf{1}_j(u^{w_j}))\dot{C}_j(u^{w_1},\dots,u^{w_d})du \int_{[0,1]} B_C(\mathbf{1}_j(v^{w_j}))\dot{C}_j(v^{w_1},\dots,v^{w_d})dv\bigg].$$

It is clear that

$$\begin{aligned}\gamma_j^2 &= 2\int_{[0,1]}\int_{[0,v]} u^{w_j}(1-v^{w_j})\mu_j(\mathbf{w})\mu_j(\mathbf{w})u^{A(\mathbf{w})-w_j}v^{A(\mathbf{w})-w_j}duv,\\ &= \left(\frac{\mu_j(\mathbf{w})}{1+A(\mathbf{w})}\right)^2\frac{w_j}{2A(\mathbf{w})+1+1-w_j}.\end{aligned}$$

We now deal with cross product terms, the first we define is

$$\begin{aligned}\sigma_{1j} &= \mathbb{E}\bigg[\int_{[0,1]} B_C(u^{w_1},\dots,u^{w_d})du \int_{[0,1]} B_C(\mathbf{1}_j(v^{w_j}))\dot{C}_j(v^{w_1},\dots,v^{w_d})dv\bigg],\\ &= \int_{[0,1]^2}\left(C(u^{w_1},\dots,(u\wedge v)^{w_j},\dots,u^{w_d}) - u^{A(\mathbf{w})}v^{w_j}\right)\dot{C}_j(v^{w_1},\dots,v^{w_d})duv.\end{aligned}$$

Under the cube $[0,1]\times[0,v]$, we have

$$\begin{aligned}\sigma_{1j} &= \int_{[0,1]\times[0,v]}\left(C(u^{w_1},\dots,u^{w_j},\dots,u^{w_d}) - u^{A(\mathbf{w})}v^{w_j}\right)\dot{C}_j(v^{w_1},\dots,v^{w_d})duv,\\ &= \int_{[0,1]\times[0,v]} u^{A(\mathbf{w})}(1-v^{w_j})v^{A(\mathbf{w})-w_j}\mu_j(\mathbf{w})duv = \frac{\mu_j(\mathbf{w})}{2(1+A(\mathbf{w}))}\frac{w_j}{2A(\mathbf{w})+1+1-w_j}.\end{aligned}$$

Under the cube $[0,1]\times[0,u]$, we have for the right term

$$\int_{[0,1]\times[0,u]} u^{A(\mathbf{w})}v^{w_j}v^{A(\mathbf{w})-w_j}\mu_j(\mathbf{w})dvu = \frac{\mu_j(\mathbf{w})}{2(1+A(\mathbf{w}))^2}.$$

For the left term, by definition, we have

$$\int_{[0,1]\times[0,u]} C(u^{w_1},\dots,v^{w_j},\dots,u^{w_d})\dot{C}_j(v^{w_1},\dots,v^{w_d})dvu.$$

Let us consider the substitution $x = v^{w_j}$ and $y = u^{1-w_j}$, we obtain

$$\frac{1}{w_j(1-w_j)} \int_{[0,1]} \int_{[0,y^{w_j/(1-w_j)}]} C\left(y^{w_1/(1-w_j)}, \dots, x, \dots, y^{w_d/(1-w_j)}\right) \\ \times \dot{C}_j\left(x^{w_1/w_j}, \dots, x^{w_d/w_j}\right) x^{(1-w_j)/w_j} y^{w_j/(1-w_j)} dxy.$$

Let us compute the quantity

$$\dot{C}_j(x^{w_1/w_j}, \dots, x^{w_d/w_j}) = \frac{C(x^{w_1/w_j}, \dots, x^{w_d/w_j})}{x} \mu_j(\mathbf{w}).$$

Using Equation (1), we have

$$\begin{aligned} C(x^{w_1/w_j}, \dots, x^{w_d/w_j}) &= exp\left(-\ell\left(-\frac{log(x)}{w_j}w_1, \dots, \frac{log(x)}{w_j}w_d\right)\right) \\ &= exp\left(-\frac{log(x)}{w_j}\ell\left(-w_1, \dots, -w_d\right)\right) = x^{A(\mathbf{w})/w_j} =: x^{A_j(\mathbf{w})}. \end{aligned}$$

Where we use the homogeneity of order one of ℓ and that $-\ell(-w_1, \dots, -w_d) = A(\mathbf{w})$ because of Equation (2) and that $\mathbf{w} \in \Delta^{d-1}$. Now, consider the substitution $x = w^{1-s}$ and $y = w^s$, the jacobian of this transformation is given by $-log(w)$, we have

$$-\frac{\mu_j(\mathbf{w})}{w_j(1-w_j)} \int_{[0,1]} \int_{[0,1-w_j]} C\left(w^{sw_1/(1-w_j)}, \dots, w^{1-s}, \dots, w^{sw_d/(1-w_j)}\right) \\ \times w^{(1-s)\left[A_j(\mathbf{w}) + \frac{1-w_j}{w_j} - 1\right] + s\frac{w_j}{1-w_j}} log(w) dsw.$$

We now compute the quantity

$$C\left(w^{sw_1/(1-w_j)}, \dots, w^{1-s}, \dots, w^{sw_d/(1-w_j)}\right).$$

Using the same methods as above, we have

$$\begin{aligned} &C\left(w^{sw_1/(1-w_j)}, \dots, w^{1-s}, \dots, w^{sw_d/(1-w_j)}\right) \\ &= exp\left(-\ell\left(-\frac{sw_1}{1-w_j}log(w), \dots, -(1-s)log(w), \dots, -\frac{sw_d}{1-w_j}log(w)\right)\right), \\ &= exp\left(-log(w)\ell\left(-\frac{sw_1}{1-w_j}, \dots, -(1-s), \dots, -\frac{sw_d}{1-w_j}\right)\right). \end{aligned}$$

Now, using that $\mathbf{w} \in \Delta^{d-1}$, remark that $s\sum_{i \neq j} w_i/(1-w_j) = s$, we have, using Equation (2)

$$-\ell\left(-\frac{sw_1}{1-w_j}, \dots, -(1-s), \dots, -\frac{sw_d}{1-w_j}\right) = A\left(\frac{sw_1}{1-w_j}, \dots, \frac{sw_d}{1-w_j}\right).$$

Where we set $1-s$ in the j-th components of the Pickands dependence function A. So we have

$$
\begin{aligned}
\sigma_{1i} &= -\frac{\mu_j(\mathbf{w})}{w_j(1-w_j)} \int_{[0,1-w_j]} \int_{[0,1]} w^{A\left(\frac{sw_1}{1-w_j}, \dots, \frac{sw_d}{1-w_j}\right) + (1-s)\left(A_j(\mathbf{w}) + \frac{1-w_j}{w_j} - 1\right) + s\frac{w_j}{1-w_j}} log(w) dws, \\
&= \frac{\mu_j(\mathbf{w})}{w_j(1-w_j)} \int_{[0,1-w_j]} \left[A\left(\mathbf{z}_j(1-s)\right) + (1-s)\left(A_j(\mathbf{w}) + \frac{1-w_j}{w_j} - 1\right) + s\frac{w_j}{1-w_j} + 1 \right]^{-2} ds,
\end{aligned}
$$

where $\mathbf{z} = (sw_1/(1-w_j), \dots, sw_d/(1-w_j))$. No further simplifications can be obtained. For $j < k$, let us define the quantity γ_{jk} such as

$$
\gamma_{jk} = \mathbb{E}\left[\int_{[0,1]} B_C(\mathbf{1}_j(u^{w_j})) \dot{C}_j(u^{w_1}, \dots, u^{w_d}) du B_C(\mathbf{1}_k(v^{w_k})) \dot{C}_k(v^{w_1}, \dots, v^{w_d}) dv \right]. \tag{10}
$$

Again, we have

$$
\gamma_{ij} = \int_{[0,1]^2} \left(C(\mathbf{1}_{jk}(u^{w_j}, v^{w_j})) - u^{w_j} v^{w_j} \right) \dot{C}_j(u^{w_1}, \dots, u^{w_d}) \dot{C}_k(v^{w_1}, \dots, v^{w_d}) duv
$$

We set $x = u^{w_j}$ and $y = v^{w_k}$, the left side become

$$
\begin{aligned}
\gamma_{ij} &= \frac{1}{w_j(1-w_k)} \int_{[0,1]^2} C(\mathbf{1}_{jk}(x,y)) \\
&\times \dot{C}_j(x^{w_1/w_j}, \dots, x^{w_d/w_j}) \dot{C}_k(y^{w_1/w_k}, \dots, y^{w_d/w_k}) x^{(1-w_j)/w_j} y^{(1-w_k)/w_k} dxy, \\
&= \frac{\mu_j(\mathbf{w})\mu_k(\mathbf{w})}{w_j w_k} \int_{[0,1]^2} C(\mathbf{1}_{jk}(x,y)) x^{A_j(\mathbf{w}) + (1-w_j)/w_j - 1} y^{A_k(\mathbf{w}) + (1-w_k)/w_k - 1} dxy.
\end{aligned}
$$

Now, we set $x = w^{1-s}$ and $y = w^s$ and we obtain

$$
\begin{aligned}
\gamma_{jk} = &\frac{\mu_j(\mathbf{w})\mu_k(\mathbf{w})}{w_j w_k} \int_{[0,1]} \left[A(\mathbf{0}_{jk}(1-s)) + (1-s)\left(A_j(\mathbf{w}) + \frac{1-w_j}{w_j} - 1\right) \right. \\
&\left. + s\left(A_k(\mathbf{w}) + \frac{1-w_k}{w_k} + 1\right) \right]^{-2} ds.
\end{aligned}
$$

The right side of Equation (10) is given by

$$
\int_{[0,1]^2} u^{w_j} v^{w_k} \dot{C}_j(u^{w_1}, \dots, u^{w_d}) \dot{C}_k(v^{w_1}, \dots, v^{w_d}) duv = \frac{\mu_j(\mathbf{w})\mu_k(\mathbf{w})}{(1+A(\mathbf{w}))^2}.
$$

Hence the result. $\square$

References

[Fermanian et al., 2004] Fermanian, J.-D., Radulović, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860.

[Genest and Segers, 2009] Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *The Annals of Statistics*, 37(5B):2990 – 3022.

[Gudendorf and Segers, 2010] Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. In *Copula theory and its applications*, volume 198 of *Lect. Notes Stat. Proc.*, pages 127–145. Springer, Heidelberg.

[Marcon et al., 2017] Marcon, G., Padoan, S., Naveau, P., Muliere, P., and Segers, J. (2017). Multivariate nonparametric estimation of the pickands dependence function using bernstein polynomials. *Journal of Statistical Planning and Inference*, 183:1–17.

[Naveau et al., 2009] Naveau, P., Guillou, A., Cooley, D., and Diebolt, J. (2009). Modelling pairwise dependence of maxima in space. *Biometrika*, 96(1):1–17.

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

[Segers, 2015] Segers, J. (2015). Hybrid copula estimators. *J. Statist. Plann. Inference*, 160:23–34.

[Sklar, 1959] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231.

[van der Vaart and Wellner, 1996] van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Process: With Applications to Statistics*. Springer.

[Xia et al., 1999] Xia, Y., Fabian, P., Stohl, A., and Winterhalter, M. (1999). Forest climatology: estimation of missing values for bavaria, germany. *Agricultural and Forest Meteorology*, 96(1):131–144.