Introduction

Context

Definitions and Notation

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathbf{X} = (X_1, \dots, X_d)$ be a d-dimensional random vector of maxima with values in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. This random vector has a joint distribution function G and the margins of G are denoted by $F_i(x) = \mathbb{P}\{X_i \leq x\}$ for all $x \in \mathbb{R}$. A function $C : [0,1]^d \rightarrow [0,1]$ is called a bivariate copula if it is the restriction to $[0,1]^d$ of a bivariate 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 H can be decomposed as $G(\mathbf{x}) = C(F_1(x_d), \dots, F_d(x_d))$, for all $\mathbf{x} \in \mathbb{R}^d$.

Definition 1 ([Gudendorf and Segers, 2010])**.** *A d-dimensional copula C is an extreme-value copula if and only if it admits a representation of the form*

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

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 $\max(x_1, \dots, x_d) \leq \ell(x_1, \dots, x_d) \leq x_1 + \dots + x_d$ for all $(x_1, \dots, x_d) \in [0, \infty)^d$. 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 simple :

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

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

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

Let $\mathbf{X}$ be a random vector with continuous marginal distribution functions $F_1, \dots F_d$. Assume that its copula C, is an extreme-vlue copula with stable tail dependence function ℓ and Pickands dependence function A.

Definition 2 ([Marcon et al., 2017])**.** *The multivariate w-madogram ($w \in \Delta^{d-1}$), denoted by $\nu(\boldsymbol{w})$, is defined as*

$$\nu(\boldsymbol{w}) = \mathbb{E}\left[\bigvee_{i=1}^{d} \{F_i(X_i)\}^{1/w_i} - \frac{1}{d}\sum_{i=1}^{d} \{F_i(X_i)\}^{1/w_i}\right]$$

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

Starting from independent and identically distributed *i.i.d.* copies $\mathbf{X}_1, \dots, \mathbf{X}_n$ of X, suppose we observe a $2d$-tuple such as

$$(\mathbf{I}_m \mathbf{X}_m, \mathbf{X}_m), \quad m \in \{1, \dots, n\}, \tag{4}$$

where $\mathbf{I}_m \mathbf{X}_m = (X_{m,1} I_{m,1}, \dots, X_{m,d} I_{m,d})$ and $I_{m,j} = 0$ if $X_{m,j}$ is missing, otherwise $I_{m,j} = 1$, *i.e.* at each $m \in \{1, \dots, n\}$, several entries may be missing. The probability of observing a realization partially or completely, is denoted by $p_m = \mathbb{P}(I_{m,j} = 1) > 0$, $p = \mathbb{P}(I_{1,j} = 1, \dots, I_{n,j} = 1) > 0$ and we note $\mathbf{p} = (p_1, \dots, p_n, p)$. Let us now define the empirical cumulative distribution of X (resp. Y and (X, Y)) in case of missing data,

$$\hat{F}_{n,i}(x_i) = \frac{\sum_{m=1}^n 1_{\{X_m \leq x\}} I_{m,i}}{\sum_{m=1}^n I_{m,i}}, \quad \forall x_i \in \mathbb{R}. \tag{5}$$

$$\hat{G}_n(\mathbf{x}) = \frac{\sum_{m=1}^n 1_{\{X_{m,1} \leq x_1, \dots, X_{m,d} \leq x_d\}} \Pi_{i=1}^d I_{m,i}}{\sum_{m=1}^n \Pi_{i=1}^d I_{m,i}}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

Here, we weight the estimator by the number of observed data which is a natural estimator if divided by n of probabilities of missing. We have all tools in hand to recall the definition of the *hybrid copula estimator* introduced by [Segers, 2015],

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

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

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

Througout, 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 $t \in \mathbb{N}^*, X, X_t$ are maps from $(\Omega, \mathcal{A}, \mathbb{P})$ into a metric space $\mathcal{X}$ and that X is Borel measurable, $(X_t)_{t \geq 1}$ is said to converge weakly to X if $\mathbb{E}^* f(X_t) \to \mathbb{E} f(X)$ for every bounded continuous real-valued function f defined on $\mathcal{X}$, where $\mathbb{E}^*$ denotes outer expectation in the event that X_t may not be Borel measurable. In what follows, weak convergence is denoted by $X_t \rightsquigarrow X$.

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 Definition 1. Under the weak condition that the first-order partial derivatives of the copula function C exist and are continuous on subsets of the unit hypercube, [Segers, 2012] obtained weak convergence of the normalized estimation error of the empirical copula process. To obtain this condition, we make the following assumption as suggested in [Segers, 2012] in Example 5.3.

Assumption A.

(i) The bivariate distribution function G 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 ans is continuous on set $\{x \in [0, \infty)^d : x_j > 0\}$.

The Assumption A *(i)* guarantees that the representation $H(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$ 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 exists and is continuous on the set $\{\mathbf{u} \in [0,1]^d : 0 < u_j < 1\}$. We now define our estimator of Equation (2) in the general context (allowing missing data).

Definition 3. *Let $(\boldsymbol{I}_m \boldsymbol{X}_m)_{m=1}^n$ be a sample given by Equation (4), we define the hybrid estimator of the $\boldsymbol{w}$-FMadogram by*

$$\hat{\nu}_n^{\mathcal{H}}(\boldsymbol{w}) = \frac{1}{\sum_{m=1}^n \Pi_{i=1}^d I_{m,i}} \sum_{m=1}^n \left[\bigvee_{i=1}^d \left\{ \hat{F}_{n,i}(X_{m,i}) \right\}^{1/w_i} - \frac{1}{d} \sum_{i=1}^d \left\{ \hat{F}_{n,i}(X_{m,i}) \right\}^{1/w_i} \right] \Pi_{i=1}^d I_{m,i}, \tag{6}$$

where $\hat{F}_{n_i}(x_i)$ is defined on Equation (5).

The idea raised 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.* with $\mathbf{p} = \mathbf{1}$ we retrieve the w-FMadogram such as defined in [Marcon et al., 2017], namely

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

with $\hat{F}_{n,i}$ the empirical cumulative distribution function of X_i.

Remark 1. *Our estimator defined in (6) does not verify $\hat{\nu}_T^{\mathcal{H}}(\boldsymbol{e}_i) = (d-1)/2d$ while $\nu(\boldsymbol{e}_i) = (d-1)/2d$. In addition, the variance at $\mathbf{e}_i$ does not equal 0. Indeed, suppose that we evaluate* *this statistic at* $\boldsymbol{w} = 0$*, we thus obtain the following quantity :*

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

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

We can force our estimator as in [Naveau et al., 2009] to satisfy these endpoint conditions. This leads to the following corrected estimator.

Definition 4. *Under the notation of Definition 3, we define the hybrid corrected estimator of the* $\boldsymbol{w}$*-FMadogram by*

$$\hat{\nu}_T^{\mathcal{H}*}(\boldsymbol{w}) = \hat{\nu}_n(\boldsymbol{w}) + \sum_{i=1}^d \left[\frac{w_i(d-1)}{d} \frac{w_i}{1+w_i} - \frac{w_i(d-1)}{d \sum_{m=1}^n \Pi_{i=1}^d I_{m,i}} \sum_{m=1}^n \left\{ \hat{F}_{n,i}(X_{m,i}) \right\}^{1/w_i} \Pi_{i=1}^d I_{m,i} \right]. \tag{7}$$

Let us now introduce a condition on the missing mechanism :

Assumption B. *We suppose for all* $t \in \{1, \dots, T\}$*, the pairs* (I_t, J_t) *and* (X_t, Y_t) *are independent, the data are missing completely at random (**MCAR**). Furthermore, we suppose that there exists at least one* $t \in \{1, \dots, T\}$ *such that* $I_t J_t \neq 0$*.*

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

Proposition 1 (**Strong consistency**)**.** *Let* $(\boldsymbol{I}_m \boldsymbol{X}_m, \boldsymbol{X}_m)_{m=1}^n$ *a i.i.d sample given by Equation (4). We have, under Assumption B for a fixed* $\boldsymbol{w} \in [0,1]$*, as* $T \to \infty$

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

Details on the proof are given in Section 2.

Proposition 2 (**Concentration inequality**)**.** *Under the framework of Proposition 1, we have with probability* $1 - \eta$ *where* $\eta \in (0,1)$*,*

$$\left| \hat{\nu}_n^{\mathcal{H}}(\boldsymbol{w}) - \nu(\boldsymbol{w}) \right| \leq \sqrt{\frac{2}{N} log \left(\frac{d+1}{\eta} \right)}.$$

we present with Theorem 1 our main result concerning the weak convergence of the following processes

$$\sqrt{T} \left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda) \right), \quad \sqrt{T} \left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda) \right). \tag{8}$$

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. The difference being that C should be continuously differentiable on the closed cube. This statement make use of previous results on the Hadamard differentiability of the map $\phi : D([0,1]^2) \to \ell^\infty([0,1]^2)$ which transforms the cumulative distribution function H into its copula function C (see Lemma 3.9.28 from [van der Vaart and Wellner, 1996]). With the hybrid copula estimator, we need a following technical assumption in order to guarantee the weak convergence of the process $\mathbb{C}_T^{\mathcal{H}}$ (see [Segers, 2015]),

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

$$\begin{aligned}\Big(\sqrt{n}(\hat{G}_n - G); \sqrt{n}(\hat{F}_{n,1} - F)_1, \ldots, \sqrt{n}(\hat{F}_{n,d} - F_d)\Big) \\ \rightsquigarrow (\alpha \circ \boldsymbol{F}, \beta_1 \circ F_1, \ldots, \beta_d \circ F_d)\,.\end{aligned}$$

The stochastic processes α and $\beta_j, j \in \{1, \ldots, d\}$ take values in $l^\infty([0,1]^d)$ and $l^\infty([0,1])$ 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,

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

Considering the same statistical framework and missing mechanism, [Segers, 2015] shows (in Example 3.5) that the processes α, β_1 and β_2 take the following closed form

$$\begin{aligned}\beta_i(u_i) &= p_i^{-1}\mathbb{G}\left(\mathbb{1}_{X_i \le F_i^{\leftarrow}(u_i), I_i=1} - u_i \mathbb{1}_{I_i=1}\right),\\ \alpha(\mathbf{u}) &= p^{-1}\mathbb{G}\left(\mathbb{1}_{\mathbf{X} \le \mathbf{F}^{\leftarrow}(\mathbf{u})}\mathbb{1}_{\mathbf{I}=\mathbf{1}} - C(\mathbf{u})\mathbb{1}_{\mathbf{I}=\mathbf{1}}\right),\end{aligned}$$

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 $\beta_i(u_i)$, $\alpha(\boldsymbol{u})$ are, for $(\boldsymbol{u}, u_j, \boldsymbol{v}, v_j) \in$*

$[0,1]^{2d+2}$,

$$cov\left(\beta_i(u_i), \beta_i(u_j)\right) = p_i^{-1}\left(u_i \wedge u_j - u_i u_j\right),$$
$$cov\left(\beta_i(u_i), \beta_j(v_j)\right) = \frac{p_{ij}}{p_i p_j}\left(C(1, \ldots, 1, u_i, 1, \ldots, 1, v_j, 1, \ldots, 1) - uv\right),$$

and

$$cov\left(\alpha(\boldsymbol{u}), \alpha(\boldsymbol{v})\right) = p^{-1}\left(C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})\right),$$
$$cov\left(\alpha(\boldsymbol{u}), \beta_i(v_i)\right) = p_i^{-1}\left(C(u_1, \ldots, u_i \wedge v_i, \ldots, u_d) - C(\boldsymbol{u})v_i\right).$$

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 (8).

Theorem 1 (**Functional central limit theorem with missing data**). *Under Assumptions A, B, C we have the weak convergence in $\ell^\infty([0,1])$ for the hybrid estimator defined in (6) and (7), as $T \to \infty$,*

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

$$\sqrt{n}\left(\hat{\nu}_T^{\mathcal{H}*}(\boldsymbol{w}) - \nu(\boldsymbol{w})\right) \rightsquigarrow \left(\frac{1}{d}\sum_{i=1}^{d}(1 + w_i(d-1))\int_{[0,1]} \alpha(1, \ldots, x^{w_i}, \ldots, 1) - \beta_i(x^{w_i})dx \right.$$
$$\left. - \int_{[0,1]} S_C(x^{w_1}, \ldots, x^{w_d})dx\right)_{\lambda \in [0,1]}.$$

Proof is deferred 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 $w \in \Delta^{d-1}$ fixé, la loi de $\sqrt{n}(\nu_n(\mathbf{w}) - \nu(\mathbf{w})$ suit une Gaussienne centré (car transformation linéaire continue d'un processus Gaussien tendue) et sa variance est donnée par :

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

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

Proposition 3 (**Boulin, 2021**)**.** *Je pense avoir une forme close de la variance et celle-ci est décomposée comme suit :*

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

2 Proof

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

$$\left| \bigvee_{i=1}^d \{F_{n,i}(X_{m,i})\}^{1/w_i} - \bigvee_{i=1}^d \{F_i(X_i)\}^{1/w_i} \right| \leq \sup_{i \in \{1,\dots,d\}} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_i)^{1/w_i} \right|.$$

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

$$\begin{aligned} \{F_{n,i}(X_{m,i})\}^{1/w_i} &= F_i(X_i)^{1/w_i} + F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_i)^{1/w_i}, \\ &\leq F_i(X_i)^{1/w_i} + \sup_{i \in \{1,\dots,d\}} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_i)^{1/w_i} \right|, \\ &\leq \bigvee_{i=1}^d \{F_i(X_i)^{1/w_i}\}^{1/w_i} + \sup_{i \in \{1,\dots,d\}} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_i)^{1/w_i} \right|. \end{aligned}$$

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

$$\bigvee_{i=1}^d \{F_{n,i}(X_{m,i})\}^{1/w_i} - \bigvee_{i=1}^d \{F_i(X_i)\}^{1/w_i} \leq \sup_{i \in \{1,\dots,d\}} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_i)^{1/w_i} \right|.$$

Moreover, by symmetry of $F_{n,i}$ and F_i, the second ones follows similarly. □

Proof of Proposition 1 We write, for notational convenience $n_m = \Pi_{i=1}^d I_{m,i}$ and $N = \sum_{m=1}^n n_m$. We prove it for $\hat{\nu}_T^{\mathcal{H}}(\lambda)$ as the strong consistency for $\hat{\nu}_T^{\mathcal{H}*}(\lambda)$ use the same arguments. The estimator $\hat{\nu}_T(\lambda)$ 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

$$\nu_n(\mathbf{w}) = \frac{1}{N} \sum_{m=1}^n \left(\bigvee_{i=1}^d \{F_i(X_{m,i})\}^{1/w_i} - \frac{1}{d} \sum_{i=1}^d \{F_i(X_{m,i})\}^{1/w_i} \right) n_m$$

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 :

$$\begin{aligned}
|\hat{\nu}_n(\mathbf{w}) - \nu(\mathbf{w})| \leq & \frac{1}{N} \sum_{m=1}^{n} \left| \bigvee_{i=1}^{d} F_{n,i}(X_{m,i})^{1/w_i} - \bigvee_{i=1}^{d} F_i(X_{m,i})^{1/w_i} \right| n_m \\
& + \frac{1}{2nd} \sum_{m=1}^{n} \sum_{i=1}^{d} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_{m,i})^{1/w_i} \right| n_m \\
\leq & 2 \sup_{i \in \{1,\dots,d\}} \sup_{m \in \{1,\dots,n\}} \left| F_{n,i}(X_{m,i})^{1/w_i} - F_i(X_{m,i})^{1/w_i} \right|,
\end{aligned}$$

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 $\mathbf{x} \in \mathbb{R}^d$,

$$\begin{aligned}
f_i(\mathbf{I}, \mathbf{X}) = \mathbb{1}_{\{I_i = 1\}}, \quad g_{i,x_i}(\mathbf{I}, \mathbf{X}) \mathbb{1}_{\{X_i \leq x_i, I_i = 1\}}, \\
f_{d+1} = \Pi_{i=1}^{d} f_i, \quad g_{d+1,\mathbf{x}} = \Pi_{i=1}^{d} g_{i,x_i}.
\end{aligned}$$

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

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

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

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

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

$$\begin{aligned}
\hat{F}_{n,i}(x_i) &= \frac{p_i F_i(x_i) + n^{-1/2} \mathbb{G}_n g_{i,x_i}}{p_i + n^{-1/2} \mathbb{G}_n f_i}, \\
\hat{G}_n(\mathbf{x}) &= \frac{p G(\mathbf{x}) + n^{-1/2} \mathbb{G}_n g_{d+1,\mathbf{x}}}{p + n^{-1/2} \mathbb{G}_n f_{d+1}}
\end{aligned}$$

We obtain for the second one

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

We thus have

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

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 $\hat{G}_n(\mathbf{x}) - G(\mathbf{x}) = \circ_{\mathbb{P}}(1)$. Using Slutsky's lemma gives us

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

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

$$
\begin{aligned}
\beta_i(u_i) &= p_i^{-1}\mathbb{G}\left(g_{i,F_i^{\leftarrow}(u_i)} - u_i f_i\right), \\
\alpha(\mathbf{u}) &= p^{-1}\mathbb{G}\left(g_{d+1,\mathbf{F}^{\leftarrow}(\mathbf{u})} - C(\mathbf{u})f_{d+1}\right).
\end{aligned}
$$

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

$$
\begin{aligned}
cov(\beta_i(u_i), \beta_j(v_j)) &= \mathbb{E}\left[p_i^{-1}\mathbb{G}\left(g_{i,F_i^{\leftarrow}(u_i)} - u_i f_i\right) p_j^{-1}\mathbb{G}\left(g_{j,F_j^{\leftarrow}(v_j)} - v_j f_j\right)\right], \\
&= \frac{1}{p_i p_j}\mathbb{E}\left[\mathbb{G}\left(g_{i,F_i^{\leftarrow}(u_i)} - u_i f_i\right)\mathbb{G}\left(g_{j,F_j^{\leftarrow}(v_j)} - v_j f_j\right)\right], \\
&= \frac{1}{p_i p_j}\mathbb{P}\left\{X_i \leq F_i^{\leftarrow}(u_i), X_j \leq F_j^{\leftarrow}(v_j), I_i = 1, I_j = 1\right\} - \frac{p_{ij}}{p_i p_j}u_i v_j, \\
&= \frac{1}{p_i p_j}\mathbb{P}\left\{X_i \leq F_i^{\leftarrow}(u_i), X_j \leq F_j^{\leftarrow}(v_j)\right\}\mathbb{P}\left\{I_i = 1, I_j = 1\right\} - \frac{p_{ij}}{p_i p_j}u_i v_j, \\
&= \frac{p_{ij}}{p_i p_j}\left(C(1,\dots,1,u_i,1,\dots,1,v_j,1,\dots,1) - u_i v_j\right).
\end{aligned}
$$

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 $\mathbb{E}[F_i(X_i)^\alpha] = (1+\alpha)^{-1}$ for $\alpha \neq 1$, we can write $\nu(\mathbf{w})$ as :

$$\begin{aligned}\nu(\mathbf{w}) =&\mathbb{E}\left[\bigvee_{i=1}^d \{F_i(X_i)\}^{1/w_i} - \frac{1}{d}\sum_{i=1}^d \{F_i(X_i)\}^{1/w_i}\right] + \\ &\sum_{i=1}^d \left(\frac{w_i(d-1)}{d}\frac{w_i}{1+w_i} - \frac{w_i(d-1)}{d}\mathbb{E}\left[F_i(X_i)^{1/w_i}\right]\right), \\ &= \mathbb{E}\left[\bigvee_{i=1}^d \{F_i(X_i)\}^{1/w_i}\right] - \frac{1}{d}\sum_{i=1}^d (1+w_i(d-1))\mathbb{E}\left[F_i(X_i)^{1/w_i}\right] + c(\mathbf{w}),\end{aligned}$$

with $c(\mathbf{w}) = d^{-1}\sum_{i=1}^d w_i/(1+w_i)$. Let us note by $g_\mathbf{w}$ the function defined as

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

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:

$$\nu_n^{\mathcal{H}*}(\mathbf{w}) = \frac{1}{N}\sum_{m=1}^n g_\mathbf{w}\left(\hat{F}_{n,1}(X_{m,1}), \dots, \hat{F}_{n,d}(X_{m,d})\right) + c(\mathbf{w}) = \int_{[0,1]^d} g_\mathbf{w}(\mathbf{u})\, d\hat{C}_n^{\mathcal{H}}(\mathbf{u}) + c(\mathbf{w}),$$

$$\nu(\mathbf{w}) = \int_{[0,1]^d} g_\mathbf{w}(\mathbf{u})\, dC(\mathbf{u}) + c(\mathbf{w}).$$

We thus have, proceeding as in Theorem 2.4 of [Marcon et al., 2017] :

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

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

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

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 $S_C(1,\dots,1,u_i,1,\dots,1) = \alpha(1,\dots,1,u_i,1,\dots,1) - \beta_i(u_i)$ and we obtain our statement. $\square$

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 3 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_i^2 is defined by the following

$$\begin{aligned}\gamma_i^2 =& \mathbb{E}\Bigg[\int_{[0,1]} B_C(1,\dots,u^{w_i},\dots,1)\dot{C}_i(u^{w_1},\dots,u^{w_d})du\\ &\times \int_{[0,1]} B_C(1,\dots,v^{w_i},\dots,1)\dot{C}_i(v^{w_1},\dots,v^{w_d})dv\Bigg].\end{aligned}$$

It is clear that

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

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

$$\sigma_{1i} = \mathbb{E}\Big[\int_{[0,1]} B_C(u^{w_1},\dots,u^{w_d})du \int_{[0,1]} B_C(1,\dots,v^{w_i},\dots,1)\dot{C}_i(v^{w_1},\dots,v^{w_d})dv\Big],$$
$$= \int_{[0,1]^2} \left(C(u^{w_1},\dots,(u\wedge v)^{w_i},\dots,u^{w_d}) - u^{A(\mathbf{w})}v^{w_i}\right)\dot{C}_i(v^{w_1},\dots,v^{w_d})duv.$$

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

$$\sigma_{1i} = \int_{[0,1]\times[0,v]} \left(C(u^{w_1},\dots,u^{w_i},\dots,u^{w_d}) - u^{A(\mathbf{w})}v^{w_i}\right)\dot{C}_i(v^{w_1},\dots,v^{w_d})duv,$$
$$= \int_{[0,1]\times[0,v]} u^{A(\mathbf{w})}(1-v^{w_i})v^{A(\mathbf{w})-w_i}\mu_i(\mathbf{w})duv = \frac{\mu_i(\mathbf{w})}{2(1+A(\mathbf{w})}\frac{w_i}{2A(\mathbf{w})+1+1-w_i}.$$

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

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

For the left term, by definition, we have

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

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

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

Let us compute the quantity

$$\dot{C}_i(x^{w_1/w_i},\dots,x^{w_d/w_i}) = \frac{C(x^{w_1/w_i},\dots,x^{w_d/w_i})}{x}\mu_i(\mathbf{w}).$$

Using Equation (1), we have

$$C(x^{w_1/w_i},\dots,x^{w_d/w_i}) = exp\left(-\ell\left(-\frac{log(x)}{w_i}w_1,\dots,\frac{log(x)}{w_i}w_d\right)\right)$$
$$= exp\left(-\frac{log(x)}{w_i}\ell\left(-w_1,\dots,-w_d\right)\right) = u^{A(\mathbf{w})/w_i}$$

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

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

We now compute the quantity

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

Using the same methods as above, we have

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

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

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

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

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

No further simplifications can be obtained. For $i < j$, let us define the quantity γ_{ij} such as

$$\begin{aligned}
\gamma_{ij} = \mathbb{E}\Bigg[&\int_{[0,1]} B_C(1, \dots, u^{w_i}, \dots, 1)\dot{C}_i(u^{w_1}, \dots, u^{w_d})du \\
&\times B_C(1, \dots, v^{w_j}, \dots, 1)\dot{C}_j(v^{w_1}, \dots, v^{w_d})dv\Bigg]
\end{aligned}$$

Again, we have

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

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

$$\begin{aligned}\gamma_{ij} = & \frac{1}{w_i(1-w_j)} \int_{[0,1]^2} C(1,\dots,x,\dots,y,\dots,1) \\ & \times \dot{C}_i(x^{w_1/w_i},\dots,x^{w_d/w_i})\dot{C}_j(y^{w_1/w_j},\dots,y^{w_d/w_j})x^{(1-w_i)/w_i}y^{(1-w_j)/w_j}dxy \\ & = \frac{\mu_i(\mathbf{w})\mu_j(\mathbf{w})}{w_iw_j} \int_{[0,1]^2} C(1,\dots,x,\dots,y,\dots,1)x^{A_i(\mathbf{w})+(1-w_i)/w_i-1}y^{A_j(\mathbf{w})+(1-w_j)/w_j-1}dxy\end{aligned}$$

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

$$\begin{aligned}\gamma_{ij} = & \frac{\mu_i(\mathbf{w})\mu_j(\mathbf{w})}{w_iw_j} \int_{[0,1]} \Bigg[A(0,\dots,s,\dots,0) + (1-s)\left(A_i(\mathbf{w}) + \frac{1-w_i}{w_i} - 1\right) \\ & + s\left[A_j(\mathbf{w}) + \frac{1-w_j}{w_j} + 1\right]^{-2}\Bigg]ds.\end{aligned}$$

Where we set $1-s$ at the ith component of the Pickands. The right side of the expression is given by

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

Hence the result. □

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.

[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.