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. The problem of missing data is present in many fields in particular in environmental research (see [Xia et al., 1999]), usually due to instruments, communication and processing errors. In a time series setting, the observation periods of a multivariate series could be different and overlap only partially. The problem of estimating when unequal amounts of data are available to each variables is of interest in many applications for financial economics where data cannot be generated as neatly overlapping samples (see [Patton and Wiley, 2006]). Also, the machine learning community has to design prediction procedure which handle missing values [Josse et al., 2020]. In this paper, we consider inference methods for assessing extremal dependencies involving variables with missing values. We are particularly interested in the dependence structure of multivariate extreme value distribution. Formally, 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 $\mathbf{x} \in \mathbb{R}^d$ and the copula C is unique if the marginals are continuous. 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 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 $\max(x_1, \dots, x_d) \leq \ell(x_1, \dots, x_d) \leq x_1 + \dots + x_d$, $\forall (x_1, \dots, x_d) \in [0,\infty)^d$. Denote by $\Delta^{d-1} = \{(w_1, \dots, w_d) \in [0,1]^d : w_1 + \dots + w_d = 1\}$ the unit simplex. By homogeneity, ℓ 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 $j \in \{2, \dots, d\}$ and $w_1 = 1 - w_2 - \dots - w_d$ with $(x_1, \dots, x_d) \in [0,\infty)^d \setminus \{0\}$. 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. The $\mathbf{w}$-madogram can be interpred as the L_1-distance between the maximum and the average of the uniform margins $F_1(X_1), \dots, F_d(X_d)$ elevated to the inverse of the corresponding weights $w_1, \dots, w_d$. This quantity describes the possible dependence structure between extremes by its relation with the Pickands dependence function. Through this relation, it contributes to the vast literature of the estimation of the Pickands dependence function for bivariate extreme value copula (see [Pickands, 1981], [Deheuvels, 1991], [Capéraà et al., 1997] or [Hall and Tajvidi, 2000]) but also multivariate extreme value copula, *e.g.* [Gudendorf and Segers, 2012]. Also, a test for assessing asymptotic independence in dimensions $d \geq 2$ has been designed based on the $\mathbf{w}$-madogram (see [Guillou et al., 2018]). Several methods for handling missing values in the framework of extremes have been proposed for univariate time series; see *e.g.*, [Hall and Scotto, 2008, Ferreira et al., 2021]. However, handling missing values in the context of multivariate extreme values with $d \geq 2$ still in their infancy. The main contribution of this paper is to give an estimator of the $\mathbf{w}$-madogram involving variables with missing values.

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 $j < k$. The notation $\mathbf{1}$ (resp. $\mathbf{0}$) corresponds to the d-dimensional vector composed out of 1 (resp. 0). Similarly, we define $\mathbf{1}_j(s)$, $\mathbf{0}_j(s)$, $\mathbf{1}_{jk}(s,t)$ and $\mathbf{0}_{jk}(s,t)$ with the same idea of previous notations of this paragraph.

The following notations are also 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 $f : \mathcal{X} \rightarrow \mathbb{R}$, 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 $\overset{a.s.}{\rightarrow}$, $\overset{d}{\rightarrow}$ 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 $X_n \rightsquigarrow X$.

In this paper, we propose in Section 1 estimators of the **w**-madogram suitable to the missing data framework. We state the weak convergence of the depicted estimators. Explicit formula for the asymptotic variance are also given. In section 2, a simulation study will provides evidence that the conclusions remain valid in the finite-sample framework. All the proofs are postponed to Section 3.

1 Non parametric estimation of the Madogram with missing data

We assume that the copula C is an extreme value copula as in Equation (1). We consider independent and identically distributed *i.i.d.* copies $\mathbf{X}_1, \dots, \mathbf{X}_n$ of $\mathbf{X}$. In presence of missing data, we do not observe a complete vector $\mathbf{X}_i$ for $i \in \{1, \dots, n\}$. We introduce $\mathbf{I}_i \in \{0,1\}^d$ which satisfies, $\forall j \in \{1, \dots, d\}$, $I_{i,j} = 0$ if $X_{i,j}$ is not observed. To formalize incomplete observations, we introduce the incomplete vector $\tilde{\mathbf{X}}_i$ with values in the product space $\bigotimes_{j=1}^d (\mathbb{R} \cup \{\text{NA}\})$. We thus have

$$\tilde{X}_{i,j} = X_{i,j} I_{i,j} + \text{NA}(1 - I_{i,j}), \quad i \in \{1, \dots, n\}, \, j \in \{1, \dots, d\}.$$

We thus suppose that we observe a $2d$-tuple such as

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

i.e. at each $i \in \{1, \dots, n\}$, several entries may be missing. We also suppose that for all $i \in \{1, \dots, n\}$, $\mathbf{I}_i$ are *i.i.d* copies where $I_{i,j}$ are sampled from a Bernoulli random variable $\mathcal{B}(p_j)$ for $j \in \{1, \dots, d\}$ with $p_j = \mathbb{P}(I_j = 1)$. We denote by p the probability of observing completely a realization from $\mathbf{X}$, that is $p = \mathbb{P}(I_1 = 1, \dots, I_d = 1)$. Let us now define the empirical cumulative distribution in case of missing data, we note for notational convenience $\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\} := \{\tilde{X}_{i,1} \leq x_1, \dots, \tilde{X}_{i,d} \leq x_d\}$,

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

The idea raised here is to estimate non parametrically the margins using all available data of the corresponding series. We recall the definition of the *hybrid copula estimator* introduced by [Segers, 2015]

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

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

$$\mathbb{C}_n^{\mathcal{H}}(\mathbf{u}) = \sqrt{n}\left(\hat{C}_n^{\mathcal{H}}(\mathbf{u}) - C(\mathbf{u})\right), \quad \mathbf{u} \in [0,1]^d. \tag{7}$$

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 classical empirical copula process (see [Deheuvels, 1979]). To satisfy this condition, we introduce the following assumption as suggested in [Segers, 2012] (see 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 in Equation (2) *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 (with possible missing data).

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

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

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

The intuitive idea here is to estimate the margins using all available data from the corresponding variables and estimate $\nu(\mathbf{w})$ using only the overlapping data. One may verify that in the complete data framework, *i.e.* when $p = 1$ 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.

Note that the theoretical quantity defined in (4) does verify endpoint coinstraints, *i.e.* $\nu(\mathbf{e}_j) = (d-1)/2d$ for all $j \in \{1,\dots,d\}$ where $\mathbf{e}_j$ is the jth vector of the canonical basis.

Remark 1. *Our estimator, unlike ν, defined in* (8) *does not verify the endpoints constraints. 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 :*

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

In this situation, the sample $(\tilde{X}_{i,1},\dots,\tilde{X}_{i,j-1},\tilde{X}_{i,j+1},\dots,\tilde{X}_{i,d})_{i=1}^{n}$ 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.

An is [Naveau et al., 2009], we propose a slightly modified estimator which satisfy the endpoint constraints. This can be imposed as follows.

Definition 2. *Let $(\boldsymbol{I}_i, \tilde{\boldsymbol{X}}_i)_{i=1}^{n}$ be a sample given by Equation* (5) *and $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$ be as in* (8)*. Given continuous functions $\lambda_1,\dots,\lambda_d : \Delta^{d-1} \to \mathbb{R}$ verifying $\lambda_j(\boldsymbol{e}_k) = \delta_{jk}$ (the Kronecker delta) for $j,k \in \{1,\dots,d\}$. We define the hybrid corrected estimator of the* $\mathbf{w}$*-madogram by*

$$\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) = \hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \sum_{j=1}^{d}\frac{\lambda_j(\mathbf{w})(d-1)}{d}\left[\frac{1}{\sum_{i=1}^{n}\Pi_{j=1}^{d} I_{i,j}}\sum_{i=1}^{n}\left(\left\{\hat{F}_{n,j}(\tilde{X}_{i,j})\right\}^{1/w_j}\Pi_{j=1}^{d} I_{i,j}\right) - \frac{w_j}{1+w_j}\right]. \tag{9}$$

Remark 2. *One has often that endpoint corrections does not have an impact to the asymptotic 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 of interest as discussed in Remark 1.*

We present with Theorem 1 in this Section a functional central limit theorem concerning the weak convergence of the following processes

$$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w})\right), \quad \sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) - \nu(\mathbf{w})\right). \tag{10}$$

Before presenting this result, we introduce a specific assumption on the missing mechanism as detailed below.

Assumption B. *We suppose that for all $i \in \{1, \ldots, n\}$, the vector $\boldsymbol{I}_i$ and $\boldsymbol{X}_i$ are independent, i.e. the data are missing completely at random (**MCAR**).*

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 $\mathbb{C}_n^{\mathcal{H}}$ in (7) (see [Segers, 2015]). We note for convenience the marginal distribution and quantile functions into vector valued functions $\mathbf{F}_d$ and $\mathbf{F}_d^{\leftarrow}$:

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

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

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

where the stochastic processes α and $\beta_j, j \in \{1, \ldots, d\}$ 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 $\mathbb{C}_n^{\mathcal{H}}$ in (7) converges weakly to the tight Gaussian process S_C defined by

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

We have all tools in hand to consider the weak convergence of the stochastic processes in Equation (10). We note by $\{\mathbf{X} \leq \mathbf{F}^{\leftarrow}(\mathbf{u})\} = \{X_1 \leq F_1^{\leftarrow}(u_1), \ldots, X_d \leq F_d^{\leftarrow}(u_d)\}$.

Theorem 1. *Let $\mathbb{G}$ a tight Gaussian process. Under Assumptions A, B, C we have the weak convergence in $\ell^\infty(\Delta^{d-1})$ for the hybrid estimator defined in Equations (8) and (9),* *as* $n \to \infty$,

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

where $S_C(\boldsymbol{u}) = \alpha(\boldsymbol{u}) - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\beta_j(u_j)$, $\alpha(\boldsymbol{u}) = p^{-1}\mathbb{G}(\mathbb{1}_{\{\mathbf{X}\leq \mathbf{F}^{\leftarrow}(\mathbf{u}), \boldsymbol{I}=\boldsymbol{1}\}} - C(\boldsymbol{u})\mathbb{1}_{\{\boldsymbol{I}=\boldsymbol{1}\}})$ *and* $\beta_j(u_j) = p_j^{-1}\mathbb{G}(\mathbb{1}_{\{X_j\leq F_j^{\leftarrow}(u_j), I_j=1\}} - u_j\mathbb{1}_{\{I_j=1\}})$ *for* $j\in\{1,\dots,d\}$ *and* $\boldsymbol{u}\in[0,1]^d$. *The covariance functions of the processes* α *and* β_j *are given by for* $(\boldsymbol{u},\boldsymbol{v},v_k)\in[0,1]^{2d+1}$, *and for* $j\in\{1,\dots,d\}$ *and* $j<k$

$$cov\left(\beta_j(u_j),\beta_j(v_j)\right) = p_j^{-1}\left(u_j\wedge v_j - u_jv_j\right),$$
$$cov\left(\beta_j(u_j),\beta_k(v_k)\right) = \frac{p_{jk}}{p_jp_k}\left(C(\boldsymbol{1}_{j,k}(u_j,v_k)) - u_jv_k\right),$$

and

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

where $\boldsymbol{u}\wedge\boldsymbol{v}$ *denotes the vector of componentwise minima and* $p_{jk} = \mathbb{P}(I_j=1, I_k=1)$.

We use empirical process arguments formulated in [van der Vaart and Wellner, 1996] to establish such a result. Details can be found in Section 3.1.

Let $\mathbf{p} = (p_1,\dots,p_d,p)$. The following proposition state the asymptotic distribution of the estimators and gives explicit formula for the asymptotic variances.

Proposition 1. *For* $\mathbf{w}\in\Delta^{d-1}$, *if* C *is an extreme copula with Pickands dependence function* A, *we have*

$$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}}(\boldsymbol{w}) - \nu(\boldsymbol{w})\right) \xrightarrow{d} \mathcal{N}\left(0,\mathcal{S}^{\mathcal{H}}(\boldsymbol{p},\boldsymbol{w})\right), \quad \sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}*}(\boldsymbol{w}) - \nu(\boldsymbol{w})\right) \xrightarrow{d} \mathcal{N}\left(0,\mathcal{S}^{\mathcal{H}*}(\boldsymbol{p},\boldsymbol{w})\right).$$

Furthermore the asymptotic variances of the random variables $\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w})\right)$ *and*

$\sqrt{n}\left(\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w})-\nu(\mathbf{w})\right)$ *are given by*

$$\begin{aligned}
\mathcal{S}^{\mathcal{H}}(\mathbf{p},\boldsymbol{w}) =& \frac{1}{d^2}\sum_{j=1}^{d}(p^{-1}-p_j^{-1})\sigma_j^2(\mathbf{w})+\sigma_{d+1}^2(\mathbf{p},\mathbf{w}) \\
&+\frac{2}{d^2}\sum_{j<k}\left(p^{-1}-p_j^{-1}-p_k^{-1}+\frac{p_{jk}}{p_jp_k}\right)\sigma_{jk}(\mathbf{w})-\frac{2}{d}\sum_{j=1}^{d}(p^{-1}-p_j^{-1})\sigma_j^{(1)}(\mathbf{w}) \\
&+\frac{2}{d}\sum_{j=1}^{d}\sum_{j<k}\left(p_k^{-1}-\frac{p_{jk}}{p_jp_k}\right)\sigma_{jk}^{(2)}(\mathbf{w})+\frac{2}{d}\sum_{j=1}^{d}\sum_{k<j}\left(p_k^{-1}-\frac{p_{kj}}{p_jp_k}\right)\sigma_{kj}^{(2)}(\mathbf{w}),
\end{aligned}$$

and given continuous functions $\lambda_1,\dots,\lambda_d:\Delta^{d-1}\to\mathbb{R}$ *verifying* $\lambda_j(\boldsymbol{e}_k)=\delta_{jk}$

$$\begin{aligned}
\mathcal{S}^{\mathcal{H}*}(\mathbf{p},\boldsymbol{w}) =& \frac{1}{d^2}\sum_{j=1}^{d}(p^{-1}-p_j^{-1})(1+\lambda_j(\mathbf{w})(d-1))^2\sigma_j^2(\mathbf{w})+\sigma_{d+1}^2(\mathbf{p},\mathbf{w}) \\
&+\frac{2}{d^2}\sum_{j<k}\left(p^{-1}-p_j^{-1}-p_k^{-1}+\frac{p_{jk}}{p_jp_k}\right)(1+\lambda_j(\mathbf{w})(d-1))(1+\lambda_k(\mathbf{w})(d-1))\sigma_{jk}(\mathbf{w}) \\
&-\frac{2}{d}\sum_{j=1}^{d}(p^{-1}-p_j^{-1})(1+\lambda_j(\mathbf{w})(d-1))\sigma_j^{(1)}(\mathbf{w}) \\
&+\frac{2}{d}\sum_{j=1}^{d}\sum_{j<k}\left(p_k^{-1}-\frac{p_{jk}}{p_jp_k}\right)(1+\lambda_j(\mathbf{w})(d-1))\sigma_{jk}^{(2)}(\mathbf{w}) \\
&+\frac{2}{d}\sum_{j=1}^{d}\sum_{k<j}\left(p_k^{-1}-\frac{p_{kj}}{p_jp_k}\right)(1+\lambda_j(\mathbf{w})(d-1))\sigma_{kj}^{(2)}(\mathbf{w}),
\end{aligned}$$

where explicit expressions of the functions $\sigma_j^2(\mathbf{w})$ *for* $j\in\{1,\dots,d\}$, $\sigma_{d+1}^2(\mathbf{w},\boldsymbol{p})$, $\sigma_{jk}(\mathbf{w})$ *with* $j<k$, $\sigma_j^{(1)}(\mathbf{w})$ *with* $j\in\{1,\dots,d\}$, $\sigma_{jk}^{(2)}(\mathbf{w})$ *for* $j<k$ *and* $\sigma_{kj}^{(2)}(\mathbf{w})$ *as* $k<j$ *are postponed to Section 3.1 for the sake of readibility.*

Technical details are available on Section 3.1. Considering the special case of independent copula, Corollary 1 below gives a closed form of the limit variance which no longer depends of the integral of the Pickands dependence function.

Corollary 1. *In the framework of Theorem 1 and if* $C(\boldsymbol{u})=\Pi_{j=1}^d u_j$, *then the functions* $\sigma_{d+1}^2(\mathbf{w},\boldsymbol{p})$, $\sigma_j^{(1)}(\mathbf{w})$ *with* $j\in\{1,\dots,d\}$, *has the following forms, for* $\mathbf{w}\in\Delta^{d-1}$ *:*

$$\begin{aligned}
\sigma_{d+1}^2(\boldsymbol{p},\mathbf{w}) &= \frac{1}{4}\left(\frac{1}{3p}-\sum_{j=1}^{d}p_j^{-1}\frac{w_j}{4-w_j}\right), \\
\sigma_j^{(1)}(\mathbf{w}) &= \frac{1}{2}\left[\frac{1}{3}-\frac{1}{1+w_j}\right]+\frac{w_j}{3(1+w_j)(3+w_j)},
\end{aligned}$$

and $\sigma_{jk}(\mathbf{w}) = 0$ *for* $j < k$, $\sigma_{jk}^{(2)}(\mathbf{w}) = 0$ *for* $j < k$ *and* $\sigma_{kj}^{(2)}(\mathbf{w}) = 0$ *with* $k < j$.

Remark 3. *From our knowledge, only [Guillou et al., 2014] detailed the variance for the madogram of a bivariate random vector while taking the independent copula and found* $1/90$. *The result stated in Corollary 1 is not an extension of this result because the hypothesis* $\boldsymbol{w} \in \Delta^{d-1}$ *is crucial. Nevertheless, the same techniques used to prove Proposition 1 could be of interest to show a similar explicit formula of the asymptotic variance for an extension of the madogram with* $d \geq 2$.

It is a common knowledge that the $\mathbf{w}$-madogram is of main interest to construct of the Pickands dependence function. In the missing data framework we define the following estimator.

Definition 3. *Let* $(\boldsymbol{I}_i, \tilde{\boldsymbol{X}}_i)_{i=1}^n$ *be a samble given by* (5), *the hybrid nonparametric estimator of the Pickands dependence function is defined as*

$$\hat{A}_n^{\mathcal{H}*}(\mathbf{w}) = \frac{\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) + c(\mathbf{w})}{1 - \hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) - c(\mathbf{w})}, \tag{11}$$

where $\hat{\nu}_n^{\mathcal{H}*}(\mathbf{w})$ *defined in Equation* (9) *and* $c(\mathbf{w}) = d^{-1}\sum_{j=1}^d w_j/(1+w_j)$.

Using the results of [Marcon et al., 2017] (namely, Theorem 2.4) and Proposition 1 of this paper, we state the following corollary.

Corollary 2. *For* $\mathbf{w} \in \Delta^{d-1}$, *if* C *is an extreme value copula with Pickands dependence function, we have :*

$$\sqrt{n}\left(\hat{A}_n^{\mathcal{H}*}(\mathbf{w}) - A(\mathbf{w})\right) \underset{n\to\infty}{\overset{d}{\to}} \mathcal{N}\left(0, \mathcal{V}(\boldsymbol{p}, \mathbf{w})\right),$$

where $\mathcal{V}(\boldsymbol{p}, \mathbf{w})$ *is the asymptoptic variance and its closed formula is given by*

$$\mathcal{V}(\boldsymbol{p}, \mathbf{w}) = (1 + A(\mathbf{w}))^4 \mathcal{S}^{\mathcal{H}*}(\boldsymbol{p}, \mathbf{w}).$$

Proof of this result can be find in Section 3.1. Weak consistency of our estimators directly comes down from Proposition 1. We are nonetheless able to state the strong consistency with only the help of Assumption B.

Proposition 2 (Strong consistency). *Let* $(\boldsymbol{I}_i, \tilde{\boldsymbol{X}}_i)_{i=1}^n$ *a i.i.d sample given by Equation (5). We have, under Assumption B for a fixed* $\mathbf{w} \in \Delta^{d-1}$

$$\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) \underset{n\to\infty}{\overset{a.s.}{\to}} \nu(\mathbf{w}), \quad \hat{\nu}_n^{\mathcal{H}*}(\mathbf{w}) \underset{n\to\infty}{\overset{a.s.}{\to}} \nu(\mathbf{w}).$$

Details of the proof are given in Section 3.1.

2 Numerical results

3 Proof

3.1 Proof of the main results

For the rest of this section, we will write, for notational convenience $n_i = \Pi_{j=1}^{d} I_{i,j}$ and $N = \sum_{i=1}^{n} n_i$. The following proof gives arguments used to establish the functional central limit theorem of our processes defined in Equation (10). Before going into the details, we need a intermediary lemma to assert that the empirical cumulative distributions functions in case of missing data does verify Assumption C and give covariance functions of the asymptotic processes α and β_j with $j \in \{1, \ldots, d\}$. This result comes down from [Segers, 2015] (see Example 3.5) where the result was proved for bivariate random variables but the higher dimension is directly obtained using same arguments.

Lemma 1. *The vector $(\sqrt{n}(\hat{F}_n - F); \sqrt{n}(\hat{F}_{n,1} - F_1), \ldots, \sqrt{n}(\hat{F}_{n,d} - F_d))$ where $\hat{F}_n$ and $\hat{F}_{n,j}$ for $j \in \{1, \ldots, d\}$ are defined in* (6) *does verify Assumption C with*

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

where $\mathbb{G}$ is a tight Gaussian process. Furthermore the covariance functions of the processes $\beta_j(u_j)$, $\alpha(\boldsymbol{u})$ are, for $(\boldsymbol{u}, \boldsymbol{v}, v_k) \in [0,1]^{2d+1}$, and for $j \in \{1, \ldots, d\}$ 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

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

Where $\boldsymbol{u} \wedge \boldsymbol{v}$ denotes the vector of componentwise minima and $p_{jk} = \mathbb{P}(I_j = 1, I_k = 1)$.

Proof of Lemma 1 is postponed to Section 3.2.

Proof of Theorem 1 We do the proof for $\hat{\nu}_n^{\mathcal{H}*}$ as the proof for $\hat{\nu}_n^{\mathcal{H}}$ 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:

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

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

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

where $\mathbf{1}_j(u)$ denotes the vector composed out of 1 except for the jth component where u does stand. 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 $\ell^\infty(\Delta^{d-1})$. Recall that S_C is the asymptotic process where $\mathbb{C}_n^{\mathcal{H}}$ does converge in the sense of the weak convergence in $\ell^\infty(\Delta^{d-1})$ and is defined by $S_C(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{j=1}^d \beta_j(u_j)\dot{C}_j(\mathbf{u})$ with $\mathbf{u} \in [0,1]^d$ and α and β_j are processes defined in Lemma 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. $\square$

The asymptotic normality of our estimators directly comes down from being a linear transformation of a tight Gaussian process for $\mathbf{w} \in \Delta^{d-1}$. The proof below use technical arguments to exhibits the closed expressions of the asymptotic variances of the Gaussians limit law of our estimators defined in Equation (8) and (9). Two tools make the computation feasible, the first one is the form exhibited by Equation (2) which transform a double integrals with respect to the trajectory of the copula function as the double integrals of a power function. When this trick is not possible, again the expression of the extreme value copula with respect to the Pickands dependence function is of main interest. Indeed, with some substitutions, we are able to express the double integrals as a the integrals with respect to the Pickands dependence function using the following equality :

$$-\int_{[0,1]} w^\alpha log(w) dw = \frac{1}{(\alpha+1)^2},$$

where $\alpha \neq 1$.

Proof of Proposition 1 By definition the asymptotic variance $\mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w})$ is given for a fixed $\mathbf{w} \in \Delta^{d-1}$ by

$$\mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w}) := Var\left(\frac{1}{d}\sum_{j=1}^d \int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx - \int_{[0,1]} S_C(x^{w_1}, \dots, x^{w_d})dx\right).$$

Using the property of the variance, we thus obtain

$$
\begin{aligned}
\mathcal{S}^{\mathcal{H}}(\mathbf{p},\mathbf{w}) = & \frac{1}{d^2}\sum_{j=1}^{d} Var\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx\right) \\
& + Var\left(\int_{[0,1]} S_C(x^{w_1},\dots,x^{w_d})dx\right) \\
& + \frac{2}{d^2}\sum_{j<k} cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \alpha(\mathbf{1}_k(x^{w_k})) - \beta_k(x^{w_k})dx\right) \\
& - \frac{2}{d}\sum_{j=1}^{d} cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \alpha(x^{w_1},\dots,x^{w_d})dx\right) \\
& + \frac{2}{d}\sum_{j=1}^{d}\sum_{k=1}^{d} cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \beta_k(x^{w_k})\dot{C}_k(x^{w_1},\dots,x^{w_d})dx\right).
\end{aligned}
$$

By order, we write

$$
\begin{aligned}
& Var\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx\right) = \left(p^{-1} - p_j^{-1}\right)\sigma_j^2(\mathbf{w}), \\
& Var\left(\int_{[0,1]} S_C(x^{w_1},\dots,x^{w_d})dx\right) = \sigma_{d+1}^2(\mathbf{p},\mathbf{w}), \\
& cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \alpha(\mathbf{1}_k(x^{w_k})) - \beta_k(x^{w_k})dx\right) = \\
& \left(p^{-1} - p_j^{-1} - p_k^{-1} + \frac{p_{jk}}{p_j p_k}\right)\sigma_{jk}(\mathbf{w}), \\
& cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \alpha(x^{w_1},\dots,x^{w_d})dx\right) = \left(p^{-1} - p_j^{-1}\right)\sigma_j^{(1)}(\mathbf{w}), \\
& cov\left(\int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx, \int_{[0,1]} \beta_k(x^{w_k})\dot{C}_k(x^{w_1},\dots,x^{w_d})dx\right) = \\
& \left(p_k^{-1} - \frac{p_{jk}}{p_j p_k}\right)\sigma_{jk}^{(2)}(\mathbf{w})\mathbb{1}_{\{j<k\}} - \left(p_k^{-1} - \frac{p_{jk}}{p_j p_k}\right)\sigma_{kj}^{(2)}(\mathbf{w})\mathbb{1}_{\{k<j\}}.
\end{aligned}
$$

We first show in details the closed form for $\sigma_{d+1}^2(\mathbf{p},\mathbf{w})$, the other forms are given without explanations as the technical tools used are those used for $\sigma_{d+1}^2(\mathbf{p},\mathbf{w})$. Proceding as before, we decompose this quantity as its sum of the variance (the squared term γ_1^2 and γ_j^2 for $j \in \{1,\dots,d\}$) and the covariance terms (γ_{1j} and τ_{jk}). The explicit formula of these quantities will be defined below. We thus explicit σ_{d+1}^2 as a linear combination of the probabilities of observing and the variance / covariances terms such as :

$$\sigma_{d+1}^2(\mathbf{p},\mathbf{w}) = p^{-1}\gamma_1^2(\mathbf{w}) + \sum_{j=1}^d p_j^{-1}\gamma_j^2(\mathbf{w}) - 2\sum_{j=1}^d p_j^{-1}\gamma_{1j}(\mathbf{w}) + 2\sum_{j<k}\frac{p_{jk}}{p_j p_k}\tau_{jk}(\mathbf{w}). \tag{12}$$

Let us before exhibit a useful form of the partial derivatives of the extreme value copula. 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, as $\ell(x_1,\dots,x_d)$ is homogeneous of degree 1, the partial derivative $\dot{\ell}_j(x_1,\dots,x_d)$ is homogeneous of degree 0 for $j \in \{1,\dots,d\}$, we thus obtain a suitable form of the partial derivatives of the extreme value copula for $u \in]0,1[$ and $\mathbf{w} \in \Delta^{d-1}$:

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

Now, let us compute

$$\begin{aligned}p^{-1}\gamma_1^2(\mathbf{w}) &\triangleq \mathbb{E}\left[\int_{[0,1]}\alpha(u^{w_1},\dots,u^{w_d})du\int_{[0,1]}\alpha(v^{w_1},\dots,v^{w_d})dv\right],\\ &= \frac{2}{p}\int_{[0,1]}\int_{[0,v]}u^{A(\mathbf{w})}(1-v^{A(\mathbf{w})})duv.\end{aligned}$$

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

$$\gamma_1^2(\mathbf{w}) = 2\int_{[0,1]}\int_{[0,v]}u^{A(\mathbf{w})}(1-v^{A(\mathbf{w})})duv = \frac{1}{(1+A(\mathbf{w}))^2}\frac{A(\mathbf{w})}{2+A(\mathbf{w})}.$$

The quantity $\gamma_j^2(\mathbf{w})$ is defined by the following

$$\begin{aligned}p_j^{-1}\gamma_j^2(\mathbf{w}) &\triangleq \mathbb{E}\bigg[\int_{[0,1]}\beta_j(u^{w_j})\dot{C}_j(u^{w_1},\dots,u^{w_d})du\int_{[0,1]}\beta_j(u^{w_j})\dot{C}_j(v^{w_1},\dots,v^{w_d})dv\bigg],\\ &= \frac{2}{p_j}\int_{[0,1]}\int_{[0,v]}u^{w_j}(1-v^{w_j})\mu_j(\mathbf{w})\mu_j(\mathbf{w})u^{A(\mathbf{w})-w_j}v^{A(\mathbf{w})-w_j}duv.\end{aligned}$$

It is clear that

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

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

$$p_j^{-1}\gamma_{1j}(\mathbf{w}) \triangleq \mathbb{E}\bigg[\int_{[0,1]}\alpha(u^{w_1},\ldots,u^{w_d})du\int_{[0,1]}\beta_j(v^{w_j})\dot{C}_j(v^{w_1},\ldots,v^{w_d})dv\bigg],$$
$$= p_j^{-1}\int_{[0,1]^2}\left(C(u^{w_1},\ldots,(u\wedge v)^{w_j},\ldots,u^{w_d}) - u^{A(\mathbf{w})}v^{w_j}\right)\dot{C}_j(v^{w_1},\ldots,v^{w_d})duv.$$

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

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

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

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

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

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

Let us compute the quantity

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

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

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

Where we note by $A_j(\mathbf{w}) := A(\mathbf{w})/w_j$ with $j \in \{1, \dots, d\}$. 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)

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

where $\mathbf{z} = (sw_1/(1-w_j), \dots, sw_d/(1-w_j))$. So we have

$$
\begin{aligned}
\gamma_{1j}(\mathbf{w}) &= -\frac{\mu_j(\mathbf{w})}{w_j(1-w_j)} \int_{[0,1-w_j]}\int_{[0,1]} w^{A(\mathbf{z}_j(1-s))+(1-s)\left(A_j(\mathbf{w})+\frac{1-w_j}{w_j}-1\right)+s\frac{w_j}{1-w_j}} log(w) dw s, \\
&= \frac{\mu_j(\mathbf{w})}{w_j(1-w_j)} \int_{[0,1-w_j]} \left[A\left(\mathbf{z}_j(1-s)\right) + (1-s)\left(A_j(\mathbf{w}) + \frac{1-w_j}{w_j} - 1\right) + s\frac{w_j}{1-w_j} + 1\right]^{-2} ds.
\end{aligned}
$$

No further simplifications can be obtained. For $j < k$, let us define the quantity τ_{jk} such as

$$\frac{p_{jk}}{p_j p_k}\tau_{jk}(\mathbf{w}) \triangleq \mathbb{E}\Big[\int_{[0,1]} \beta_j(u^{w_j})\dot{C}_j(u^{w_1},\dots,u^{w_d})du \int_{[0,1]} \beta_k(v^{w_k})\dot{C}_k(v^{w_1},\dots,v^{w_d})dv\Big]. \quad (13)$$

Again, we have

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

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

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

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

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

The right side of Equation (13) 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 for $\sigma^2_{d+1}(\boldsymbol{w})$. Using the same techniques, we show that for $j \in \{1,\dots,d\}$

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

For $j < k$, we compute

$$\begin{aligned}\sigma_{jk}(\mathbf{w}) = & \int_{[0,1]^2} C(1_{jk}(u^{w_j},v^{w_k})) - u^{w_j}v^{w_k}duv, \\ & = \frac{1}{w_j w_k}\int_{[0,1]} \left[A(0_{jk}(1-s,s)) + (1-s)\frac{1-w_j}{w_j} + s\frac{1-w_k}{w_k} + 1\right]^{-2} ds \\ & - \frac{1}{1+w_j}\frac{1}{1+w_k}.\end{aligned}$$

Let $j \in \{1, \ldots, d\}$, thus

$$\sigma_j^{(1)}(\mathbf{w}) = \int_{[0,1]^2} C\left(u^{w_1}, \ldots, (u \wedge v)^{w_j}, \ldots, u^{w_d}\right) - C(u^{w_1}, \ldots, u^{w_d}) v^{w_j} ds,$$
$$= \frac{1}{w_j(1-w_j)} \int_{[0,1]} \left[A(\mathbf{z}_j(1-s) + (1-s)\frac{1-w_j}{w_j} + s\frac{w_j}{1-w_j} + 1 \right]^{-2} ds$$
$$+ \frac{1}{1+A(\mathbf{w})} \left[\frac{1}{2+A(\mathbf{w})} - \frac{1}{1+w_j} \right].$$

Now, if $j < k$, we have :

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

If $k < j$, we obtain

$$\sigma_{kj}^{(2)}(\mathbf{w}) = \frac{\mu_k(\mathbf{w})}{w_j w_k} \int_{[0,1]} \left[A(0_{kj}(1-s, s)) + s\frac{1-w_j}{w_j} \right.$$
$$\left. + (1-s)\left(A_k(\mathbf{w}) + \frac{1-w_k}{w_k} - 1 \right) + 1 \right]^{-2} ds - \frac{\mu_k(\mathbf{w})}{1+A(\mathbf{w})} \frac{1}{1+w_j}.$$

Hence the statement. $\square$

The following lines will gives some details to establish the explicit formula of the asymptotic variance when we suppose that each components of the random vector $\mathbf{X}$ are independent. In this framework, we have that $\mu_j(\mathbf{w}) = 1$ for every $j \in \{1, \ldots, d\}$ and thus $\dot{C}_j(u^{w_1}, \ldots, u^{w_d}) = u^{1-w_j}$. Furthermore, in the independent case, most of the integrals are reduced to zero.

Proof of Corollary 1 In the term σ_{d+1}^2 given in (12), only the terms γ_1^2, γ_j^2 and γ_{1j} matters because, in the independent case :

$$\tau_{jk}(\mathbf{w}) = \int_{[0,1]^2} \left(u^{w_j} v^{w_k} - u^{w_j} v^{w_k}\right) \dot{C}_j(u^{w_1}, \ldots, u^{w_d}) \dot{C}_k(v^{w_1}, \ldots, v^{w_d}) duv = 0.$$

For γ_{1j}, we have to compute

$$\gamma_{1j}(\mathbf{w}) = 2\int_{[0,1]\times[0,v]} u(1-v^{w_j})v^{1-w_j} duv = \frac{1}{4}\frac{w_j}{4-w_j}.$$

For γ_1^2 and γ_j^2, we just have to set $A(\mathbf{w})$ in their expressions, we thus have

$$\gamma_1^2(\mathbf{w}) = \frac{1}{12}, \quad \gamma_j^2 = \frac{1}{4}\frac{w_j}{4-w_j}.$$

We thus have

$$\sigma_{d+1}^2(\mathbf{p},\mathbf{w}) = \frac{1}{4}\left(\frac{1}{3p} - \sum_{j=1}^d p_j^{-1}\frac{w_j}{4-w_j}\right).$$

The rest is left to the reader as the other computations follows from the same arguments.

□

We now give some element to establish Corollary 2. This result follows from the functional Delta method (Theorem 3.9.4 of [van der Vaart and Wellner, 1996]) and the result establish in Proposition 1.

Proof of Corollary 2 Applying the functional Delta method, we have as $n \to \infty$,

$$\sqrt{n}\left(\hat{A}_n^{\mathcal{H}*}(\mathbf{w}) - A(\mathbf{w})\right) \rightsquigarrow -\left(1+A(\mathbf{w})\right)^2\Bigg\{\int_{[0,1]}\alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx \\ - \int_{[0,1]} S_C(x^{w_1},\dots,x^{w_d})dx\Bigg\}_{\boldsymbol{w}\in\Delta^{d-1}}.$$

For a fixed $\mathbf{w}\in\Delta^{d-1}$, as a linear transformation of a tight Gaussian process, it follows that

$$\sqrt{n}\left(\hat{A}_n^{\mathcal{H}*}(\mathbf{w}) - A(\mathbf{w})\right) \underset{n\to\infty}{\overset{d}{\to}} \mathcal{N}\left(0,\mathcal{V}(\mathbf{p},\mathbf{w})\right),$$

where, by definition :

$$\begin{aligned}\mathcal{V}(\mathbf{p},\mathbf{w}) &\triangleq Var\left(-\left(1+A(\mathbf{w})\right)^2\left\{\int_{[0,1]}\alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j})dx - \int_{[0,1]} S_C(x^{w_1},\dots,x^{w_d})\right\}dx\right)\\ &= (1+A(\mathbf{w}))^4\mathcal{S}^{\mathcal{H}*}(\mathbf{p},\mathbf{w}),\end{aligned}$$

where we used Proposition 1 to conclude. □

Finally, we are going to prove Proposition 2. The strong consistency of the our estimators will be established in a two-step process : first, we prove the strong consistency of the estimator $\nu_n(\mathbf{w})$ which is the nonparametric estimator of the $\mathbf{w}$-madogram with known margins and, second, we show that the limit of

$$\sup_{j\in\{1,\dots,d\}} \sup_{i\in\{1,\dots,n\}} \left| \left\{ \hat{F}_{n,j}(\tilde{X}_{i,j}) \right\}^{1/w_j} - \left\{ F_j(\tilde{X}_{i,j}) \right\}^{1/w_j} \right|,$$

is zero almost surely.

Proof of Proposition 2 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. Before going into the main arguments, we need the following lemma

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

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

The proof of Lemma 2 can be fin in Section 3.2. Thus, the estimator $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$ is strongly consistent since it holds

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

where

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

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

$$\left| \nu_n(\mathbf{w}) - \nu(\mathbf{w}) \right| \underset{n\to\infty}{\overset{a.s.}{\to}} 0$$

For the second term, we write :

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

Where we used Lemma 2 to obtain the second inequality. The right term converges almost surely to zero by Glivencko-Cantelli and the uniform continuity of $x \mapsto x^{1/w_j}$ on $[0,1]$. $\square$

3.2 Proof of lemmata

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 $j \in \{1, \dots, d\}$

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

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

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

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

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

We obtain for the second one

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

We thus have

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

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

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

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

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

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 Lemma 2 The lemma becomes trivial once we write, $\forall i \in \{1,\ldots,n\}$ and $j \in \{1,\ldots,d\}$

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

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

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

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

References

[Capéraà et al., 1997] Capéraà, P., Fougères, A.-L., and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84:567–577.

[Deheuvels, 1979] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. *Bulletin Royal Belge de L'Académie des sciences, 65, 274-292*.

[Deheuvels, 1991] Deheuvels, P. (1991). On the limiting behavior of the pickands estimator for bivariate extreme-value distributions. *Statistics & Probability Letters*, 12(5):429–439.

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

[Ferreira et al., 2021] Ferreira, H., Martins, A., and Temido, M. (2021). Extremal behaviour of a periodically controlled sequence with imputed values. *Statistical Papers*, 62:1–23.

[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.

[Gudendorf and Segers, 2012] Gudendorf, G. and Segers, J. (2012). Nonparametric estimation of multivariate extreme-value copulas. *Journal of Statistical Planning and Inference*, 142(12):3073–3085.

[Guillou et al., 2014] Guillou, A., Naveau, P., and Schorgen, A. (2014). Madogram and asymptotic independence among maxima. *REVSTAT*, 12(2):119–134.

[Guillou et al., 2018] Guillou, A., Padoan, S. A., and Rizzelli, S. (2018). Inference for asymptotically independent samples of extremes. *Journal of Multivariate Analysis*, 167:114–135.

[Hall and Scotto, 2008] Hall, A. and Scotto, M. (2008). On the extremes of randomly sub-sampled time series. *REVSTAT – Statistical Journal Volume*, 6:151–164.

[Hall and Tajvidi, 2000] Hall, P. and Tajvidi, N. (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli*, 6(6):835–844.

[Josse et al., 2020] Josse, J., Prost, N., Scornet, E., and Varoquaux, G. (2020). On the consistency of supervised learning with missing values.

[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.

[Patton and Wiley, 2006] Patton, A. J. and Wiley, J. (2006). Estimation of multivariate models for time series of possibly different lenghts. *Journal of Applied Econometrics*, pages 147–173.

[Pickands, 1981] Pickands, J. (1981). Multivariate extreme value distribution. *Proceedings 43th, Session of International Statistical Institution, 1981*.

[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.