



**HAL**  
open science

## Non-parametric estimator of a multivariate madogram for missing-data and extreme value framework

Alexis Boulin, Elena Di Bernardino, Thomas Laloë, Gwladys Toulemonde

### ► To cite this version:

Alexis Boulin, Elena Di Bernardino, Thomas Laloë, Gwladys Toulemonde. Non-parametric estimator of a multivariate madogram for missing-data and extreme value framework. *Journal of Multivariate Analysis*, 2022, *Journal of Multivariate Analysis* 192 (2022), 192, 10.1016/j.jmva.2022.105059 . hal-03502804v2

**HAL Id: hal-03502804**

**<https://hal.science/hal-03502804v2>**

Submitted on 29 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Non-parametric estimator of a multivariate madogram for missing-data and extreme value framework

Alexis Boulin<sup>\*1</sup>, Elena Di Bernardino<sup>1</sup>, Thomas Laloë<sup>1</sup>, and Gwladys Toulemonde<sup>2</sup>

<sup>1</sup>Laboratoire J.A. Dieudonné, UMR CNRS 7351, Université Côte d’Azur, Nice, France

<sup>2</sup>Univ Montpellier, CNRS, Inria, Montpellier, France

## Abstract

The modeling of dependence between maxima is an important subject in several applications in risk analysis. To this aim, the extreme value copula function, characterised via the madogram, can be used as a margin-free description of the dependence structure. From a practical point of view, the family of extreme value distributions is very rich and arise naturally as the limiting distribution of properly normalised component-wise maxima. In this paper, we investigate the nonparametric estimation of the madogram where data are completely missing at random. We provide the functional central limit theorem for the considered multivariate madogram correctly normalized, towards a tight Gaussian process for which the covariance function depends on the probabilities of missing. Explicit formula for the asymptotic variance is also given. Our results are illustrated in a finite sample setting with a simulation study.

**Keywords and phrases** : Madogram, Extreme value copula , Missing Completely At Random (MCAR), Nonparametric estimation.

**MSC2020 subject classifications** : 62D10, 62G05, 62G20, 62G32, 62H10, 62H12.

## 1 Introduction

Management of environmental resources 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], or Section 2 in [Saunders et al., 2021]), 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 variable is meaningful in many applications for financial economics where data cannot be generated as neatly overlapping samples (see [Patton and Wiley, 2006]). Missing values in dependence modeling is of a prime interest as the nonparametric estimation of the empirical copula process has been tackled by [Segers, 2015] under the **Missing Completely At Random (MCAR)** condition. In this paper, we consider nonparametric methods for assessing extremal dependencies involving variables with missing values under **MCAR**

---

<sup>\*</sup>Corresponding author. Email address: [Alexis.Boulin@unice.fr](mailto:Alexis.Boulin@unice.fr)

condition. 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  $F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$ , for all  $\mathbf{x} \in \mathbb{R}^d$  and the copula  $C$  is unique if the marginals are continuous. We will consider in the rest of the paper a  $d$ -dimensional random vector  $\mathbf{X}$  which distribution is a multivariate extreme value distribution  $F$ , *i.e.*, its one dimensional distributions are Generalized Extreme-Value (GEV) distributions and the copula  $C$  is an extreme value copula (see [Gudendorf and Segers, 2010]), defined by

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

with  $\ell : [0, \infty)^d \rightarrow [0, \infty)$  the stable tail dependence function which is convex, homogeneous of order one, namely  $\ell(cx_1, \dots, cx_d) = c\ell(x_1, \dots, x_d)$  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,  $\ell$  is characterized by the *Pickands dependence function*  $A : \Delta^{d-1} \rightarrow [1/d, 1]$ , which is the restriction of  $\ell$  to the unit simplex  $\Delta^{d-1}$ :

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

for  $j \in \{2, \dots, d\}$  and  $w_1 = 1 - w_2 - \dots - w_d$  with  $(x_1, \dots, x_d) \in [0, \infty)^d \setminus \{\mathbf{0}\}$ . Notice that, for every  $\mathbf{w} \in \Delta^{d-1}$  and  $u \in ]0, 1[$

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

Based on the madogram concept from geostatistics, the  $\lambda$ -madogram is introduced in [Naveau et al., 2009] to capture bivariate extremal dependencies. The generalization of the  $\lambda$ -madogram was previously proposed by [Fonseca et al., 2015] and [Marcon et al., 2017], this quantity is defined in the latter as:

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

if  $w_j = 0$  and  $0 < u < 1$ , then  $u^{1/w_j} = 0$  by convention. The  $\mathbf{w}$ -madogram can be interpreted 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 dependence structure between extremes by its relation with the Pickands dependence function as stated by the Proposition 2.2 of [Marcon et al., 2017], namely

$$A(\mathbf{w}) = \frac{\nu(\mathbf{w}) + c(\mathbf{w})}{1 - \nu(\mathbf{w}) - c(\mathbf{w})}, \quad (5)$$

with  $c(\mathbf{w}) = d^{-1} \sum_{j=1}^d w_j / (1 + w_j)$ . 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], [Hall and Tajvidi, 2000]) and extended to the multivariate extreme value copula (see for example [Gudendorf and Segers, 2012]). Also, a test for assessing asymptotic independence in dimension  $d \geq 2$  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  $d \geq 2$  is still in their infancy.

**Main results** The main contribution of this paper is to give an estimator of the  $\mathbf{w}$ -madogram in (4) involving variables with missing values and to study its asymptotic properties. As far as we know, only [Guillou et al., 2014] detailed the variance for the madogram of a bivariate random vector while taking the independent copula and found  $1/90$ . In this paper we propose improvements in three directions : we consider a general multidimensional case ( $d \geq 2$ ), we deal with missing data and we consider a dependence structure given by an extreme value copula. Thus, we present in Theorem 1 a functional central limit theorem that gives the weak convergence for the considered multivariate madogram towards a tight Gaussian process for which the covariance function depends on the probabilities of missing. When the trajectory of our empirical process is fixed, we show in Proposition 1 the asymptotic normality of the estimator of the multivariate madogram where explicit formula for the asymptotic variance is also given. These results are transposed to the estimation of the Pickands dependence function with missing data in Corollary 2 by the use of the functional delta method.

**Notations** The symbol  $\triangleq$  means to be equal to. In order to shorten formulas, notations

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

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  $\mathcal{X}$  an arbitrary set, let  $\ell^\infty(\mathcal{X})$  denote the space of bounded real-valued functions 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  $\xrightarrow{a.s.}$ ,  $\xrightarrow{d}$  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 [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  $\mathbb{E}^* f(X_n) \rightarrow \mathbb{E} f(X)$  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  $X_n \rightsquigarrow X$ .

The paper is organised as follows: We propose in Section 2 estimators of the  $\mathbf{w}$ -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 3, we illustrate the performance of the considered estimator in the finite-sample framework. Section 4 is devoted to apply our method on a dataset with missing data and non-concomittant record

periods of annual maxima rainfall in Central Eastern Canada. A discussion on our assumptions and possible extensions of this work are presented in Section 4. All the proofs are postponed to the A.

## 2 Non parametric estimation of the Madogram with missing data

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  $\otimes_{j=1}^d (\mathbb{R} \cup \{\text{NA}\})$  (where NA denotes a missing data) such as

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

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

$$(\mathbf{I}_i, \tilde{\mathbf{X}}_i), \quad i \in \{1, \dots, n\}, \quad (6)$$

*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 from  $\mathbf{I} = (I_1, \dots, I_d)$  where  $I_j$  is distributed according to a Bernoulli random variable  $\mathcal{B}(p_j)$  with  $p_j = \mathbb{P}(I_j = 1)$  for  $j \in \{1, \dots, d\}$ . 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 write for notational convenience  $\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\} \triangleq \{\tilde{X}_{i,1} \leq x_1, \dots, \tilde{X}_{i,d} \leq x_d\}$  and  $n_j = \sum_{i=1}^n I_{i,j}$ ,

$$\hat{F}_{n,j}(x) = \frac{\sum_{i=1}^n \mathbb{1}_{\{\tilde{X}_{i,j} \leq x\}} I_{i,j}}{n_j}, \quad \forall x \in \mathbb{R}, \quad \hat{F}_n(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbb{1}_{\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\}} \prod_{j=1}^d I_{i,j}}{\sum_{i=1}^n \prod_{j=1}^d I_{i,j}}, \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad (7)$$

where  $\{\tilde{X}_{i,j} \leq x\} = \emptyset$  (*resp.*  $\{\tilde{\mathbf{X}}_i \leq \mathbf{x}\} = \emptyset$ ) if  $\tilde{X}_{i,j} = \text{NA}$  (*resp.* if there exists  $j \in \{1, \dots, d\}$  such that  $\tilde{X}_{i,j} = \text{NA}$ ). The idea raised here is to estimate non parametrically the margins using all available data of the corresponding series. To avoid dealing with points at the boundary of the unit square, it is more convenient to work with scaled ranks (see for example [Genest and Segers, 2009]) defined explicitly by

$$\tilde{U}_{i,j} = \frac{n_j}{n_j + 1} \hat{F}_{n,j}(\tilde{X}_{i,j}) = \frac{1}{n_j + 1} \sum_{k=1}^{n_j} \mathbb{1}_{\{\tilde{X}_{k,j} \leq \tilde{X}_{i,j}\}} I_{i,j}, \quad j \in \{1, \dots, d\}. \quad (8)$$

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

$$\mathbb{C}_n^{\mathcal{H}}(\mathbf{u}) = \sqrt{n} \left( \hat{C}_n^{\mathcal{H}}(\mathbf{u}) - C(\mathbf{u}) \right), \quad \mathbf{u} \in [0, 1]^d. \quad (9)$$

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.**

1. The distribution function  $F$  has continuous margins  $F_1, \dots, F_d$ .
2. For every  $j \in \{1, \dots, d\}$ , the first-order partial derivative  $\dot{\ell}_j$  of  $\ell$  with respect to  $x_j$  exists and is continuous on the set  $\{x \in [0, \infty)^d : x_j > 0\}$ .

The Assumption A1 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 A2, 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 propose an estimator of the  $\mathbf{w}$ -madogram defined in Equation (4) under a general context with possible missing data.

**Definition 1.** Let  $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)_{i=1}^n$  be a sample given by Equation (6), we define the hybrid nonparametric estimator of the  $\mathbf{w}$ -madogram in Equation (4) by

$$\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) = \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_{i,j}} \sum_{i=1}^n \left[ \left( \bigvee_{j=1}^d \tilde{U}_{i,j}^{1/w_j} - \frac{1}{d} \sum_{j=1}^d \tilde{U}_{i,j}^{1/w_j} \right) \prod_{j=1}^d I_{i,j} \right], \quad (10)$$

where  $\tilde{U}_{i,j}$  are scaled ranks defined as in Equation (8).

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. Notice that in the complete data framework, *i.e.* when  $p = 1$  we retrieve a variation of the  $\mathbf{w}$ -madogram such as defined in [Marcon et al., 2017], namely

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

with  $\tilde{U}_{i,j}$  in  $\{1/(n+1), \dots, n/(n+1)\}$ .

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

**Remark 1.** Unlike  $\nu$ , the estimator defined in (10) does not verify the endpoints constraints. In addition, the variance at  $\mathbf{e}_j$  does not equal 0. Indeed, suppose that we evaluate this statistic at  $\mathbf{w} = \mathbf{e}_j$  as  $\tilde{U}_{i,j} \in (0, 1)$  for every  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, d\}$  we obtain the following estimator

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

In this situation, the sample  $(\tilde{U}_{i,1}, \dots, \tilde{U}_{i,j-1}, \tilde{U}_{i,j+1}, \dots, \tilde{U}_{i,d})_{i=1}^n$  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 induces a supplementary variance when estimating.

Proceeding as in [Naveau et al., 2009] for the bivariate case and complete data framework, we propose below a modified estimator which satisfies the endpoint constraints in the general multivariate framework with possible missing data.

**Definition 2.** Let  $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)_{i=1}^n$  be a sample given by Equation (6) and  $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$  be as in (10). Given continuous functions  $\lambda_1, \dots, \lambda_d : \Delta^{d-1} \rightarrow \mathbb{R}$  verifying  $\lambda_j(\mathbf{e}_k) = \delta_{jk}$  (the Kronecker delta) for  $j, k \in \{1, \dots, d\}$ , we define the hybrid corrected estimator of the  $\mathbf{w}$ -madogram by

$$\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) = \hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \sum_{j=1}^d \frac{\lambda_j(\mathbf{w})(d-1)}{d} \left[ \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_{i,j}} \sum_{i=1}^n \left( \tilde{U}_{i,j}^{1/w_j} \prod_{j=1}^d I_{i,j} \right) - \frac{w_j}{1+w_j} \right]. \quad (11)$$

**Remark 2.** One has often that endpoint corrections do 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.

In the following we prove a functional central limit theorem (see Theorem 1) concerning the weak convergence of the following processes

$$\sqrt{n} \left( \hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w}) \right)_{\mathbf{w} \in \Delta^{d-1}}, \quad \sqrt{n} \left( \hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - \nu(\mathbf{w}) \right)_{\mathbf{w} \in \Delta^{d-1}}. \quad (12)$$

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

**Assumption B.** We suppose that for all  $i \in \{1, \dots, n\}$ , the vector  $\mathbf{I}_i$  and  $\mathbf{X}_i$  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. The difference being that  $C$  should be continuously differentiable on the closed hypercube. Denoting by  $D([0, 1]^2)$  the Skorohod space, this statement makes use of previous results on the Hadamard differentiability of the map  $\phi : D([0, 1]^2) \rightarrow \ell^\infty([0, 1]^2)$  which transforms the cumulative distribution function  $F$  into its copula function  $C$  (see also Lemma 3.9.28 from [van der Vaart and Wellner, 1996]). 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 (9) (see [Segers, 2015]). We note for convenience marginal distributions and quantile functions into vector valued functions  $\mathbf{F}_d$  and  $\mathbf{F}_d^{\leftarrow}$ :

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

**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

$$\left(\sqrt{n}(\hat{F}_n - F); \sqrt{n}(\hat{F}_{n,1} - F_1), \dots, \sqrt{n}(\hat{F}_{n,d} - F_d)\right) \rightsquigarrow (\alpha \circ \mathbf{F}_d, \beta_1 \circ F_1, \dots, \beta_d \circ F_d),$$

where the stochastic processes  $\alpha$  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  $\mathbb{C}_n^{\mathcal{H}}$  in (9) converges weakly to the tight Gaussian process  $S_C$  defined by

$$S_C(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\beta_j(u_j), \quad \forall \mathbf{u} \in [0, 1]^d. \quad (13)$$

Lemma 1 in **A** states that the estimator  $\hat{F}_n$  of the joint distribution and estimators of margins  $\hat{F}_{n,j}$  defined in Equation (7) verify Assumption **C** (see **A** for details). We now have all tools in hand to consider the weak convergence of the stochastic processes in Equation (12). We note by  $\{\mathbf{X} \leq \mathbf{F}_d^{\leftarrow}(\mathbf{u})\} = \{X_1 \leq F_1^{\leftarrow}(u_1), \dots, X_d \leq F_d^{\leftarrow}(u_d)\}$ .

**Theorem 1.** *Let  $\mathbb{G}$  a tight Gaussian process and continuous functions  $\lambda_1, \dots, \lambda_d : \Delta^{d-1} \rightarrow \mathbb{R}$  verifying  $\lambda_j(\mathbf{e}_k) = \delta_{jk}$ . If  $C$  is an extreme value copula with Pickands dependence function  $A$  and under Assumptions **A** and **B**, we have the weak convergence in  $\ell^\infty(\Delta^{d-1})$  for hybrid estimators defined in Equations (10) and (11), as  $n \rightarrow \infty$ ,*

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

where  $S_C$  is defined in (13),  $\alpha(\mathbf{u}) = p^{-1} \mathbb{G}(\mathbf{1}_{\{\mathbf{X} \leq \mathbf{F}_d^{\leftarrow}(\mathbf{u}), I=1\}} - C(\mathbf{u})\mathbf{1}_{\{I=1\}})$  and  $\beta_j(u_j) = p_j^{-1} \mathbb{G}(\mathbf{1}_{\{X_j \leq F_j^{\leftarrow}(u_j), I_j=1\}} - u_j \mathbf{1}_{\{I_j=1\}})$  for  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$ . For  $(\mathbf{u}, \mathbf{v}, v_k) \in [0, 1]^{2d+1}$ , for  $j \in \{1, \dots, d\}$  and  $j < k$  the covariance functions of the processes  $\alpha$  and  $\beta_j$  are given by

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

and

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

where  $\mathbf{u} \wedge \mathbf{v}$  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 [A.1](#). The following proposition states the asymptotic distribution of the estimators and gives explicit formula for the asymptotic variances for a fixed element of the unit simplex  $\Delta^{d-1}$ .

**Proposition 1.** *Let  $\mathbf{p} = (p_1, \dots, p_d, p)$  and  $\mathbf{w} \in \Delta^{d-1}$ , under the framework of [Theorem 1](#), we have*

$$\sqrt{n} (\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w})) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w})), \quad \sqrt{n} (\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - \nu(\mathbf{w})) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{S}^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w})).$$

Moreover the asymptotic variances are given by

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

and

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

where explicit expressions of the functions  $\sigma_j^2$  for  $j \in \{1, \dots, d\}$ ,  $\sigma_{d+1}^2$ ,  $\sigma_{jk}$  with  $j < k$ ,  $\sigma_j^{(1)}$  with  $j \in \{1, \dots, d\}$ ,  $\sigma_{jk}^{(2)}$  for  $j, k \in \{1, \dots, d\}$  are detailed in the proof for the sake of readability.

Considering the special case of independent copula, [Corollary 1](#) below gives a closed form of the limit variance which no longer depends on the Pickands dependence function.

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

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

and  $\sigma_{jk}$  for  $j < k$ ,  $\sigma_{jk}^{(2)}$  for  $j < k$  and  $\sigma_{kj}^{(2)}$  with  $k < j$  are constants and equal to 0.

**Remark 3.** From our knowledge, only [Guillou et al., 2014] gave an explicit value of the variance for the madogram of a bivariate random vector considering the independent copula. The result stated in Corollary 1 is not an extension of this result because the hypothesis  $\mathbf{w} \in \Delta^{d-1}$  is crucial. Nevertheless, the same techniques used to prove Proposition 1 can be applied to show a similar explicit formula of the asymptotic variance for an extension of the madogram in [Guillou et al., 2014] for  $d \geq 2$ .

Weak consistency of our estimators directly comes down from Proposition 1. We are nonetheless able to state the strong consistency only under Assumption B.

**Proposition 2 (Strong consistency).** Let  $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)_{i=1}^n$  an i.i.d sample given by Equation (6). Under Assumption B for a fixed  $\mathbf{w} \in \Delta^{d-1}$ , it holds that

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

For the rest of this section, we use our previous results to state some properties of the Pickands estimator in the missing data framework.

It is a common knowledge that the  $\mathbf{w}$ -madogram is of main interest to construct of the Pickands dependence function. Indeed, given Equation (5), one can define an estimator of the Pickands dependence function by estimating the  $\mathbf{w}$ -madogram and using it as a plug-in estimator. Most interesting properties of the  $\mathbf{w}$ -madogram such as strong consistency and the weak convergence are thus translated for the Pickands estimator using continuous mapping theorem and the Delta method. In the missing data framework we define the following estimator.

**Definition 3.** Let  $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)_{i=1}^n$  be a sample given by (6), the hybrid nonparametric estimator of the Pickands dependence function is defined as

$$\hat{A}_n^{\mathcal{H}^*}(\mathbf{w}) = \frac{\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) + c(\mathbf{w})}{1 - \hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) - c(\mathbf{w})}, \quad (14)$$

where  $\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w})$  defined in Equation (11) 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), Proposition 1 and Proposition 2 of this paper, we state the following corollary.

**Corollary 2.** Let  $\mathbf{p} = (p_1, \dots, p_d, p)$  and  $(\mathbf{I}_i, \tilde{\mathbf{X}}_i)_{i=1}^n$  be a sample given by (6). For  $\mathbf{w} \in \Delta^{d-1}$ , if  $C$  is an extreme value copula with Pickands dependence function and under Assumption B, it holds that

$$\hat{A}_n^{\mathcal{H}^*}(\mathbf{w}) \xrightarrow[n \rightarrow \infty]{a.s.} A(\mathbf{w}).$$

Furthermore, if  $C$  additionally verifies Assumptions A.1 and A.2, we obtain

$$\sqrt{n} \left( \hat{A}_n^{\mathcal{H}^*}(\mathbf{w}) - A(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{V}(\mathbf{p}, \mathbf{w})),$$

where the closed formula of the asymptotic variance is given by  $\mathcal{V}(\mathbf{p}, \mathbf{w}) = (1 + A(\mathbf{w}))^4 \mathcal{S}^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w})$ , with  $\mathcal{S}^{\mathcal{H}^*}(\mathbf{p}, \mathbf{w})$

as in Proposition 1.

### 3 Numerical results

In this section we verify our findings concerning the closed formula of the asymptotic variances through a simulation study. To do so, we compare empirical counterparts of the asymptotic variances computed out with Monte Carlo simulations with the explicit asymptotic variances given by Proposition 1. Our simulation studies are implemented using Python programming language and all the codes are available online in this [github](#) repository.

#### 3.1 Presentation of the models

We present here the six models (**M1** to **M6**) used for this simulation study. The  $d$ -dimensional Gumbel and the asymmetric logistic models are considered in models **M1** and **M2** below, the remaining ones (models **M3** to **M6**) concern only the bivariate case.

**M1 The symmetric logistic, or Gumbel model** [Gumbel, 1960] is defined by the following Pickands dependence function

$$A(w_1, \dots, w_d) = \left( \sum_{j=1}^d w_j^\theta \right)^{1/\theta},$$

with  $\theta \in [1, \infty)$ . We retrieve the independent case when  $\theta = 1$  and the dependence between the variables is stronger as  $\theta$  goes to infinity. The restriction to  $d = 2$  is immediate from the definition.

**M2** Let  $B$  be the set of all nonempty subsets of  $\{1, \dots, d\}$  and  $B_1 = \{b \in B, |b| = 1\}$ , where  $|b|$  denotes the number of elements in the set  $b$ . **The asymmetric logistic model** in [Tawn, 1990] is defined by the following Pickands dependence function

$$A(w_1, \dots, w_d) = \sum_{b \in B} \left( \sum_{j \in b} (\theta_{j,b} w_j)^{\theta_b} \right)^{1/\theta_b},$$

where  $\theta_b \in [1, \infty)$  for all  $b \in B \setminus B_1$ , and the asymmetry parameters  $\theta_{j,b} \in [0, 1]$  for all  $b \in B$  and  $j \in b$ . The model should verify the following constrains  $\sum_{b \in B(j)} \theta_{j,b} = 1$  for  $j \in \{1, \dots, d\}$  where  $B(j) = \{b \in B, j \in b\}$  and if  $\theta_b = 1$  for every  $b \in B \setminus B_1$ , then  $\theta_{j,b} = 0 \forall j \in b$ . The model contains  $2^d - d - 1$  dependence parameters and  $d(2^{d-1} - 1)$  asymmetry parameters. In case of  $d = 2$ , we go back to the asymmetric logistic model in [Tawn, 1988], namely

$$A(w) = (1 - \psi_1)w + (1 - \psi_2)(1 - w) + [(\psi_1 w)^\theta + (\psi_2(1 - w))^\theta]^{1/\theta},$$

with  $\theta \in [1, \infty)$ ,  $\psi_1, \psi_2 \in [0, 1]$ . For  $d = 3$ , the Pickands dependence function is expressed as

$$\begin{aligned} A(\mathbf{w}) = & \alpha_1 w_1 + \psi_1 w_2 + \phi_1 w_3 + ((\alpha_2 w_1)^{\theta_1} + (\psi_2 w_2)^{\theta_1})^{1/\theta_1} + ((\alpha_3 w_2)^{\theta_2} + (\phi_2 w_3)^{\theta_2})^{1/\theta_2} \\ & + ((\psi_3 w_2)^{\theta_3} + (\phi_3 w_3)^{\theta_3})^{1/\theta_3} + ((\alpha_4 w_1)^{\theta_4} + (\psi_4 w_2)^{\theta_4} + (\phi_4 w_3)^{\theta_4})^{1/\theta_4}, \end{aligned}$$

where  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_4)$ ,  $\boldsymbol{\psi} = (\psi_1, \dots, \psi_4)$ ,  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_4)$  are all elements of  $\Delta^3$ .

**M3 The asymmetric negative logistic model** in [Joe, 1990] is defined via

$$A(w) = 1 - [(\psi_1(1-w))^{-\theta} + (\psi_2 w)^{-\theta}]^{-1/\theta},$$

with parameters  $\theta \in (0, \infty)$ ,  $\psi_1, \psi_2 \in (0, 1]$ . The special case  $\psi_1 = \psi_2 = 1$  returns the Galambos model [Oliveira and Galambos, 1977].

**M4 The asymmetric mixed model** in [Tawn, 1988] corresponds to

$$A(w) = 1 - (\theta + \kappa)w + \theta w^2 + \kappa w^3,$$

with parameters  $\theta$  and  $\kappa$  satisfying  $\theta \geq 0$ ,  $\theta + 3\kappa \geq 0$ ,  $\theta + \kappa \leq 1$ ,  $\theta + 2\kappa \leq 1$ . The special case  $\kappa = 0$  and  $\theta \in [0, 1]$  yields the symmetric mixed model. In the symmetric mixed model, when  $\theta = 0$ , we recover the independent copula.

**M5 The model of Hüsler and Reiss** in [Hüsler and Reiss, 1989] is given by the Pickands dependence function

$$A(t) = (1-t)\Phi\left(\theta + \frac{1}{2\theta}\log\left(\frac{1-t}{t}\right)\right) + t\Phi\left(\theta + \frac{1}{2\theta}\log\left(\frac{t}{1-t}\right)\right),$$

where  $\theta \in (0, \infty)$  and  $\Phi$  is the standard normal distribution function. As  $\theta$  goes to  $0^+$ , the dependence between the two variables increases. When  $\theta$  goes to infinity, we are in case of near independence.

**M6 The Student  $t$ -EV model** in [Demarta and McNeil, 2005] is given by

$$A(w) = wt_{\nu+1}(z_w) + (1-w)t_{\nu+1}(z_{1-w}),$$

with  $z_w = (1+\nu)^{1/2}[\{w/(1-w)\}^{1/\nu} - \theta](1-\theta^2)^{-1/2}$ ,

and parameters  $\nu > 0$ , and  $\theta \in (-1, 1)$ , where  $t_{\nu+1}$  is the distribution function of a Student- $t$  random variable with  $\nu + 1$  degrees of freedom.

### 3.2 Description of numerical experiments

For each numerical experiment, the endpoint-corrected  $\mathbf{w}$ -madogram estimator in (11) is computed using  $\lambda_j(\mathbf{w}) = w_j$ . The study consists in three different experiments (**E1**, **E2** and **E3**). For all experiments, the empirical counterpart of the asymptotic variance given by Proposition 1 is computed out through a given grid of the simplex  $\Delta^{d-1}$ . For a given element  $\mathbf{w}$  of this grid,  $n_{iter} \in \mathbb{N} \setminus \{0\}$  random samples of size  $n$  are generated from the models **M1** to **M6** given above. By using these samples we estimate the associated  $\mathbf{w}$ -madogram. We thus compute the empirical variance of the normalized estimation error namely,

$$\mathcal{E}_n^{\mathcal{H}}(\mathbf{w}) \triangleq \widehat{Var}(\sqrt{n}(\hat{\boldsymbol{\nu}}_n^{\mathcal{H}}(\mathbf{w}) - \boldsymbol{\nu}(\mathbf{w}))), \quad \mathcal{E}_n^{\mathcal{H}^*}(\mathbf{w}) \triangleq \widehat{Var}(\sqrt{n}(\hat{\boldsymbol{\nu}}_n^{\mathcal{H}^*}(\mathbf{w}) - \boldsymbol{\nu}(\mathbf{w}))), \quad (15)$$

where  $\hat{\nu}_n^{\mathcal{H}}$  and  $\hat{\nu}_n^{\mathcal{H}*}$  are the vectors composed out of the  $n_{iter}$  hybrid and corrected estimators (see Equations (10) and (11)) of the  $\mathbf{w}$ -madogram, respectively. We also define the Mean Integrated Squared Error (MISE) between  $\mathcal{E}_n^{\mathcal{H}}$  and  $\mathcal{S}^{\mathcal{H}}$  the asymptotic variance computed in Proposition 1 (*resp.* between  $\mathcal{E}_n^{\mathcal{H}*}$  and  $\mathcal{S}^{\mathcal{H}*}$ ), that is

$$MISE^{\mathcal{H}} \triangleq \mathbb{E} \left[ \int_{\Delta^{d-1}} (\mathcal{E}_n^{\mathcal{H}}(\mathbf{w}) - \mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w}))^2 d\mathbf{w} \right], \quad MISE^{\mathcal{H}*} \triangleq \mathbb{E} \left[ \int_{\Delta^{d-1}} (\mathcal{E}_n^{\mathcal{H}*}(\mathbf{w}) - \mathcal{S}^{\mathcal{H}*}(\mathbf{p}, \mathbf{w}))^2 d\mathbf{w} \right]. \quad (16)$$

**E1** We set  $d = 2$ . A Monte Carlo study is implemented here to illustrate Proposition 1 in finite-sample setting with missing data. We consider **M2**, **M3**, **M4**, **M5** and **M6** where we fix  $n_{iter} = 300$  and  $n = 1024$ . The chosen grid is  $\{1/200, \dots, 199/200\}$  and we take  $p_1 = p_2 = 0.75$ . We estimate  $MISE^{\mathcal{H}}$  in (16) by

$$\widehat{MISE}_n^{\mathcal{H}} = \frac{1}{10} \sum_{l=1}^{10} \frac{1}{199} \sum_{k=1}^{199} \left( \mathcal{E}_{n,l}^{\mathcal{H}} \left( \frac{k}{200} \right) - \mathcal{S}^{\mathcal{H}} \left( \mathbf{p}, \frac{k}{200} \right) \right)^2,$$

with  $\mathcal{E}_{n,l}^{\mathcal{H}}, l \in \{1, \dots, 10\}$  is the empirical counterpart of  $\mathcal{S}^{\mathcal{H}}$  taking the empirical variance of 30 estimators  $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$  where  $\mathbf{w} = (k/200, 1 - k/200)$  and  $k \in \{1, \dots, 199\}$ . Each estimator of the  $\mathbf{w}$ -madogram is computed out through a random sample with  $n = 1024$ . By using the second equation in (16), the estimator  $\widehat{MISE}_n^{\mathcal{H}*}$  is defined similarly.

**E2** We fix  $d = 3$  and we consider **M1** and **M2** with  $n_{iter} = 100$  and  $n = 512$ . We set the dependence parameter as  $\theta = 1$  and  $\theta = 2$  for the first model. For the second one we take  $\boldsymbol{\alpha} = (0.4, 0.3, 0.1, 0.2)$ ,  $\boldsymbol{\psi} = (0.1, 0.2, 0.4, 0.3)$ ,  $\boldsymbol{\phi} = (0.6, 0.1, 0.1, 0.2)$  and  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_4) = (0.6, 0.5, 0.8, 0.3)$  as the dependence parameter. We take  $p_1 = p_2 = p_3 = 0.9$  and thus  $p = 0.729$ ,  $p_{ij} = 0.81$  with  $i, j \in \{1, 2, 3\}$  and  $i < j$ . We grid the  $[0, 1]^2$  cube into 10000 points at same distance from each other and we only keep those with  $w_2 + w_3 < 1.0$  where  $w_2$  and  $w_3$  are in the grid of the cube, we set  $w_1 = 1 - w_2 - w_3$ . Let  $\Delta_n^{d-1}$  be 199 points uniformly sampled from  $\Delta^2$  and  $n_{iter} = 300$ , Equation (16) is estimated with

$$\widehat{MISE}_n^{\mathcal{H}} = \frac{1}{10} \sum_{l=1}^{10} \frac{1}{199} \sum_{k \in \Delta_n^{d-1}} (\mathcal{E}_{n,l}^{\mathcal{H}}(k) - \mathcal{S}^{\mathcal{H}}(\mathbf{p}, k))^2,$$

where  $\mathcal{E}_{n,l}^{\mathcal{H}}, l \in \{1, \dots, 10\}$  is the empirical counterpart of  $\mathcal{S}^{\mathcal{H}}$  taking the empirical variance of 30 estimators  $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$  with  $\mathbf{w} \in \Delta_n^{d-1}$ . Each estimator of the  $\mathbf{w}$ -madogram is computed out through a random sample with  $n = 512$ . Again,  $\widehat{MISE}_n^{\mathcal{H}*}$  is defined in a similar way.

**E3** In this experiment, we aim to show that our conclusions are verified in a high dimension setting. We compute empirical counterpart of the asymptotic variance for a varying dimension  $d$  and we compare its value to the theoretical one given by Proposition 1. Furthermore, as the probability of observing a complete row decrease quickly with respect to the dimension  $d$ , *i.e.*  $p = p_1^{-d}$ , we set that there is no missing data. We consider the symmetric logistic model with dependence parameter  $\theta = 2$ . We sample 300 points from the unit simplex  $\Delta^{d-1}$  and we compute the following quantity

$$\delta_n^{\mathcal{H}}(\mathbf{w}) \triangleq \frac{|\mathcal{E}_n^{\mathcal{H}}(\mathbf{w}) - \mathcal{S}^{\mathcal{H}}(\mathbf{1}, \mathbf{w})|}{\mathcal{S}^{\mathcal{H}}(\mathbf{1}, \mathbf{w})}, \quad (17)$$

where  $\mathcal{E}_n^{\mathcal{H}}$  is computed from  $n_{iter} = 100$  estimators of the  $\mathbf{w}$ -madogram with sample size  $n \in \{216, 512, 1024\}$ . The results are collected for several values of  $d \in \{5, 10, \dots, 40\}$ .

Note that for Experiments **E1** and **E2**, the missing mechanism is such as  $I_1, \dots, I_d$  are pairwise independent and  $p_j = p_1, \forall j \in \{1, \dots, d\}$ . The independence setup corresponds to the worst scenario where the missingness of one variable does not influence the missingness of the other variables. *A contrario*, if we suppose that  $I_1, \dots, I_d$  are strongly dependent, *i.e.* none or all entries are missing, we then estimate a statistic on a sample of average length  $p \times n$  and we are turning back to inference in a complete data framework with a reduced sample size. This is also readily seen from the closed formula in Proposition 1, indeed in a strongly dependent setting we have  $p = p_1$ , so the asymptotic variance is reduced to the complete data framework up to a multiplicative factor.

### 3.3 Results of experiments

Results of Experiment **E1** are depicted in Figure 1. For all panels, empirical counterparts given by Equation (15) (points) fit the theoretical values exhibited from Proposition 1 (solid lines). For the hybrid estimator, as discussed in Remark 1, both empirical and theoretical values of the asymptotic variance are different from zero for each  $w \in \{\{0\}, \{1\}\}$ . The corrected version provides this feature and also modifies the shape of the curve (see Remark 2). Indeed the asymptotic behavior of the hybrid and the corrected estimators are different in the missing data framework. Notice that, in terms of variance, we do not have a strict dominance from one estimator to another.

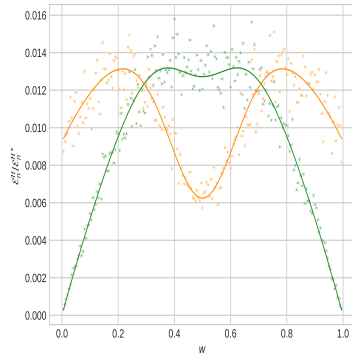
Results for Experiment **E2** are depicted in Figures 2 and 3. In Figure 2, empirical counterparts given by Equation (15) are depicted with points and closed expressions of the asymptotic variance given by Proposition 1 are drawn by a surface. Figure 3 presents the same studies differently by showing the level sets associated to the surfaces of Figure 2. As in Experiment **E1**, empirical counterparts given by the points fits the surface. Also, for the first row of Figure 2, we see that if  $\mathbf{w} \in \{\{\mathbf{e}_1\}, \{\mathbf{e}_2\}, \{\mathbf{e}_3\}\}$  then both theoretical and empirical counterparts are different from zero while this feature no longer applies in the second row with the introduction of the corrected version. In this two figures, we see that  $\mathcal{E}_n^{\mathcal{H}}$  and  $\mathcal{E}_n^{\mathcal{H}^*}$  and their empirical counterparts are close.

In order to quantify errors in Figures 1 and 2, in Table 1 are displayed  $\widehat{MISE}_n^{\mathcal{H}}$  and  $\widehat{MISE}_n^{\mathcal{H}^*}$  for the corresponding models in Experiments **E1** and **E2** to appreciate the proximity between the terms  $\mathcal{E}_n^{\mathcal{H}}$  and  $\mathcal{S}^{\mathcal{H}}$  (respectively for the corrected terms  $\mathcal{E}_n^{\mathcal{H}^*}$  and  $\mathcal{S}^{\mathcal{H}^*}$ ). As indicated by Figures 1 and 2, errors in Table 1 are close to zero.

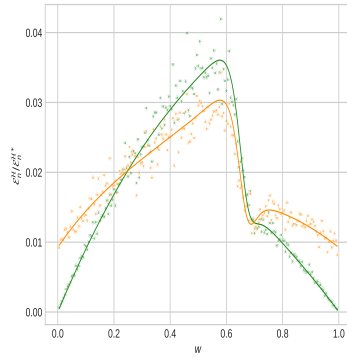
MISE ( $\times 10^{-5}$ )	E1						E2		
	GAL	ANL	ASL	ASM	HR	tEV	IND	LOG	ASL
$\widehat{MISE}_n^{\mathcal{H}}$	2.49	8.10	2.43	1.85	1.89	1.93	2.93	1.31	3.40
$\widehat{MISE}_n^{\mathcal{H}^*}$	2.77	7.02	2.04	1.94	1.96	1.93	1.95	1.57	2.91

Table 1: Estimation of  $MISE^{\mathcal{H}}$  and  $MISE^{\mathcal{H}^*}$  ( $\times 10^{-5}$ ) defined in (16) for Experiment **E1** in the sixth first columns and **E2** in the last three columns.

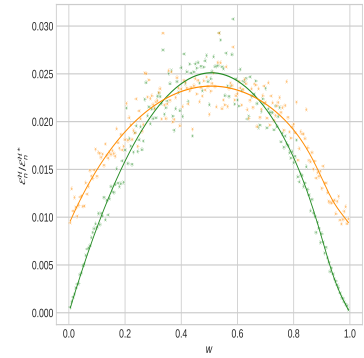
Figure 4 illustrates the results of Experiment **E3** where we have drawn boxplots for Equation (17). Not surprisingly, we observe that both the size of the boxplots and the median value are increasing with  $d$ . However, this augmentation drops as the sample size  $n$  increases and seems to appear as reasonable and able to handle the case of rather high



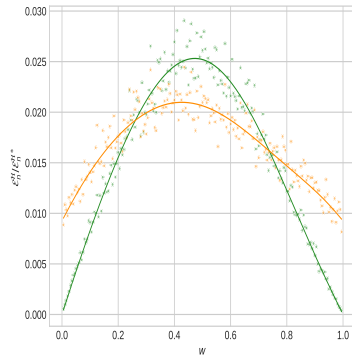
(a) GAL (M3,  $\theta = 2.5$ )



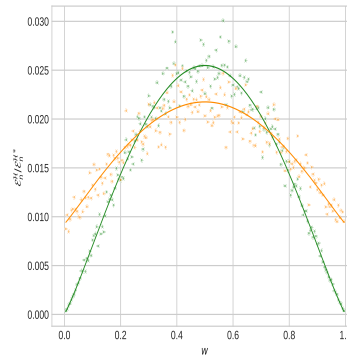
(b) ANL (M3,  $\theta = 10, \psi_1 = .5, \psi_2 = 1$ )



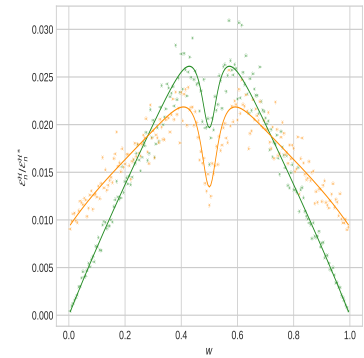
(c) ASL (M2,  $\theta = \frac{5}{2}, \psi_1 = .1, \psi_2 = 1$ )



(d) ASM (M4,  $\theta = \frac{4}{3}, \kappa = -\frac{1}{3}$ )



(e) HR (M5,  $\theta = 1.0$ )



(f) tEV (M6,  $\theta = 0.8, \nu = 0.2$ )

Figure 1:  $\mathcal{E}_n^{\mathcal{H}}$  in red and  $\mathcal{E}_n^{\mathcal{H}^*}$  in green (see (15)) as a function of  $w$ , of the asymptotic variances of the estimators of the  $\mathbf{w}$ -madogram for six extreme-value copula models. The empirical variances are based on 300 samples of size  $n = 1024$ . Solid lines are the theoretical value given by Proposition 1.

(a) IND (M1,  $\theta = 1.0$ )(b) LOG (M1,  $\theta = 2.0$ )

(c) ASL (M2)

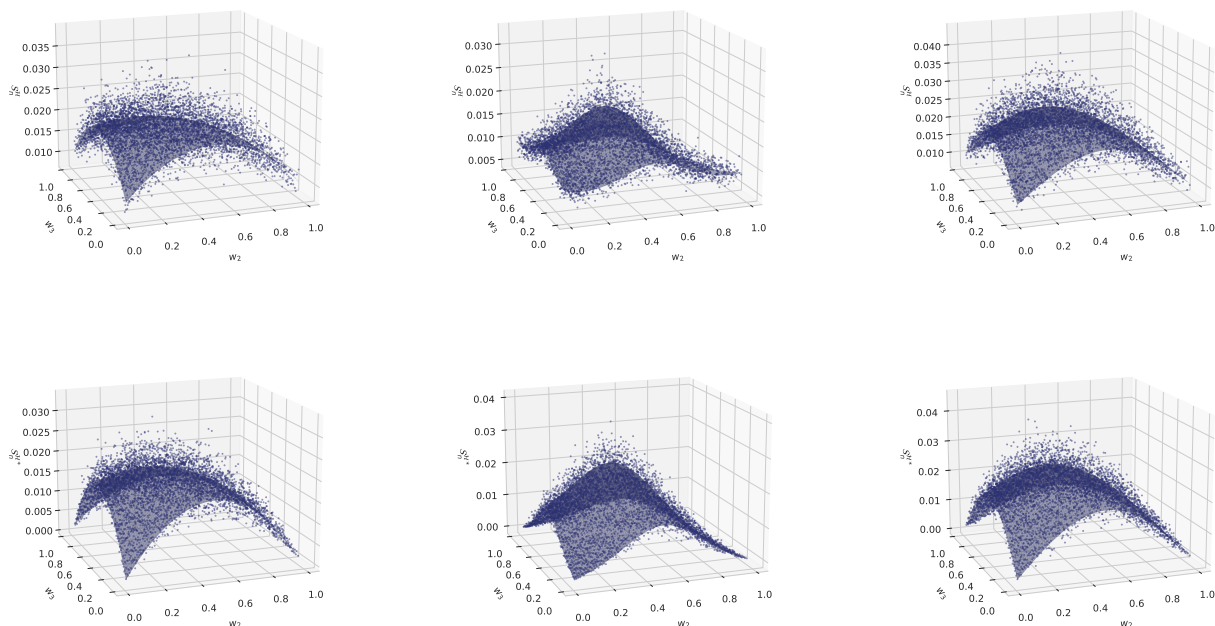


Figure 2:  $\mathcal{E}_n^{\mathcal{H}}$  (first row) and  $\mathcal{E}_n^{\mathcal{H}*}$  (second row) given by (15) as a function of  $\mathbf{w}$ -madogram. The empirical variances are based on 100 samples of size  $n = 512$ . Empirical counterparts are represented with points and theoretical values given by Proposition 1 are drawn by a surface.

(a) IND (M1,  $\theta = 1.0$ )(b) LOG (M1,  $\theta = 2.0$ )

(c) ASL (M2)

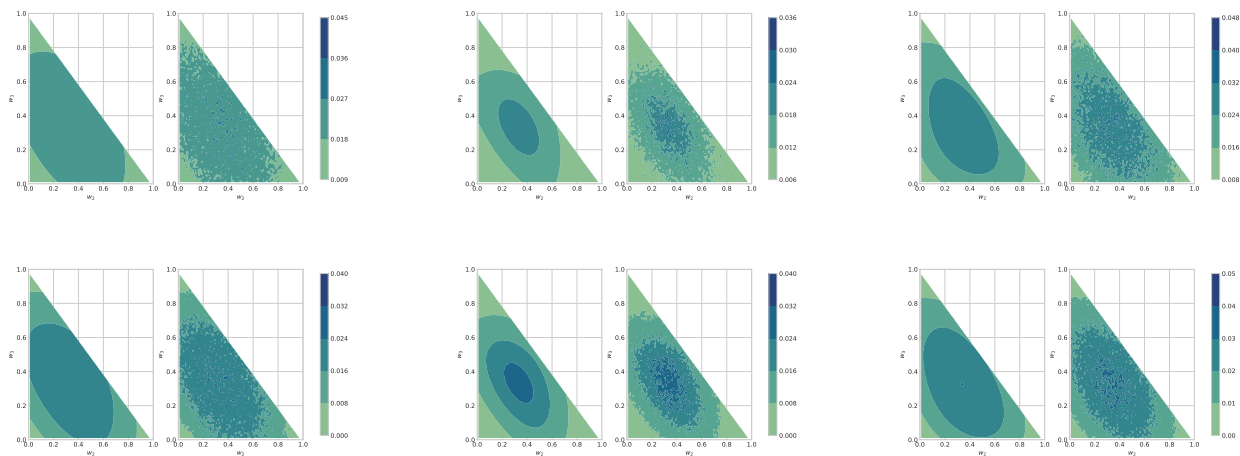


Figure 3: Level sets of  $\mathcal{E}_n^{\mathcal{H}}$  and  $\mathcal{E}_n^{\mathcal{H}*}$ , as a function of  $\mathbf{w}$ , of the asymptotic variances of the estimators of the  $\mathbf{w}$ -madogram. We present the level sets corresponding to surfaces of Fig 2. On the left panel is represented the theoretical value given by Proposition 1 while on the right the empirical counterpart is given.



dimensional data. A limitation (due to computation time issues) of this figure is that the number of points on the simplex is constant (= 300) as a function of the dimension.

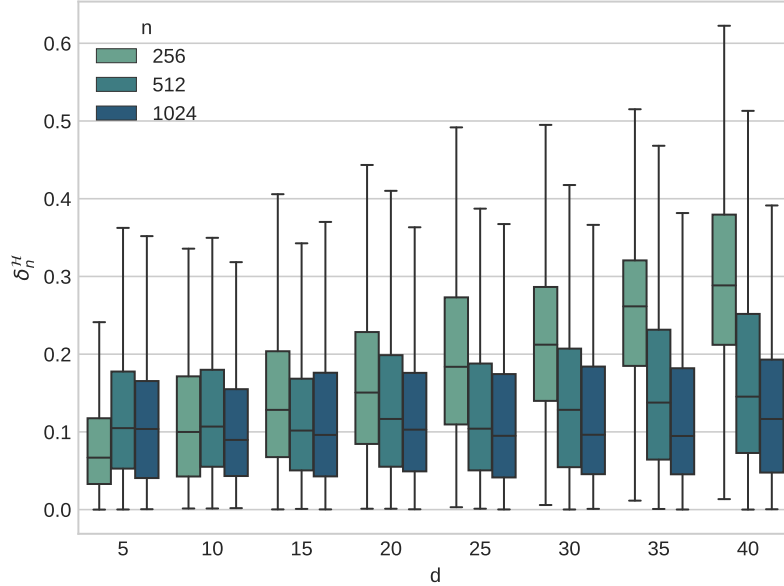


Figure 4: Boxplots for  $\delta_n^H$  for different values of  $d$  and  $n$ .

## 4 Extremal dependence rainfall analysis via hybrid madogram

In climate studies, extreme events such as heavy precipitations represent major challenge since damages from extreme weather events may have heavy consequences in both economic and human terms. Their spatial characteristics are of a prime interest and  $\mathbf{w}$ -madogram and its estimator studied in this paper (see Equation (10)) are able to capture those characteristics. A seminal application which bridges extreme value theory and geostatistics is the study of extreme rainfall since we expect spatial dependence among the recording weather stations. Precisely, we observe daily precipitation at station  $j \in \{1, \dots, d\}$  over  $n$  years. Concerning extreme events, one cannot use directly the observation for inference and we focus on block maxima. The block maxima approach is based on the observation of a sample of block maxima  $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,d})$  where  $X_{i,j}$  corresponds to the maximum at station  $j \in \{1, \dots, d\}$  within the  $i$ th disjoint block of observation. A block could be either hourly, daily or annual for example. Consistent to our approach, we do not observe  $\mathbf{X}_i$  but an incomplete vector  $\tilde{\mathbf{X}}_i \in \otimes_{j=1}^d (\mathbb{R}_+ \cup \text{NA})$ . Our main goal is to estimate the extremal dependence between maxima of groups of station. This will be done for several clusters within which similar climate characteristics are envisaged leading to dependence among extremes.

For each cluster, we compute the corrected hybrid madogram in Equation (10). This quantity is used to estimate the extremal coefficient (see for instance [Smith, 1990]), using the relation between the Pickands and the madogram given in Equation (5), defined by

$$\theta = dA \left( \frac{1}{d}, \dots, \frac{1}{d} \right). \quad (18)$$

This satisfies the condition  $1 \leq \theta \leq d$ , where the lower and upper bounds represent the case of complete positive dependence and independence among the extremes, respectively. Since its upper bound depends on  $d$ , the extremal coefficient can, alas, only be used to compare clusters of the same size. In each cluster, the extremal coefficient in (18) is estimated by  $\hat{\theta}_n = d \hat{A}_n^{\mathcal{H}*}$  where  $\hat{A}_n^{\mathcal{H}*}$  is given in Definition 3.

We illustrate the proposed methodology on rainfall data measured in millimeter registered in 95 stations in Center Eastern Canada for a duration of 24 hours publicly available in the section [engineering climate datasets](#) of the Government of Canada website. Annual maxima precipitations for a 24-hour duration are recorded from 1914 to 2017. The location of stations in Fig. 5 are given in the WGS84 coordinate space in order to have Euclidean distance between the stations and taking account of the geodesic geometry of the Earth. A specific characteristic of the considered rainfall data is the sparsity of the recorded data, *i.e.* a lot of recordings are missing (see [Palacios-Rodriguez et al., 2021] for details). Four stations were removed of the analysis due to a tiny coverage of the observation period. As the measurements are maxima over a long period of time, it is reasonable to assume that they come from a multivariate extreme value distribution see Equation (1). The dataset we consider in this section and codes are available in the [github repository](#).

With the remaining 91 stations, we compare the extremal dependence between several groups of stations as it has been done by [Marcon et al., 2017] (see Section 5) for France using a dataset with complete observations. We emphasize that the comparison of the extremal coefficient is solely relevant when clusters are of the same size. Thus, clusters were obtained by running the constrained  $k$ -means algorithm on the station coordinates (see for instance [Bradley et al., 2000]) by forcing clusters of the same size :  $d = 7$  or  $d = 13$ , *i.e.* 13 groups of 7 stations and *vice versa*. As overlapping data naturally decrease as the size of clusters increases, the case of cluster size  $d = 13$  cannot be considered here. Among the 13 clusters of size  $d = 7$ , we only keep those having at least 10 overlapping annual maxima within the cluster which results on 7 remaining clusters depicted in Figure 5a. The estimated coefficient range is between 3.88, indicating strong dependence, and 5.07, indicating medium dependence (see Figure 5b).

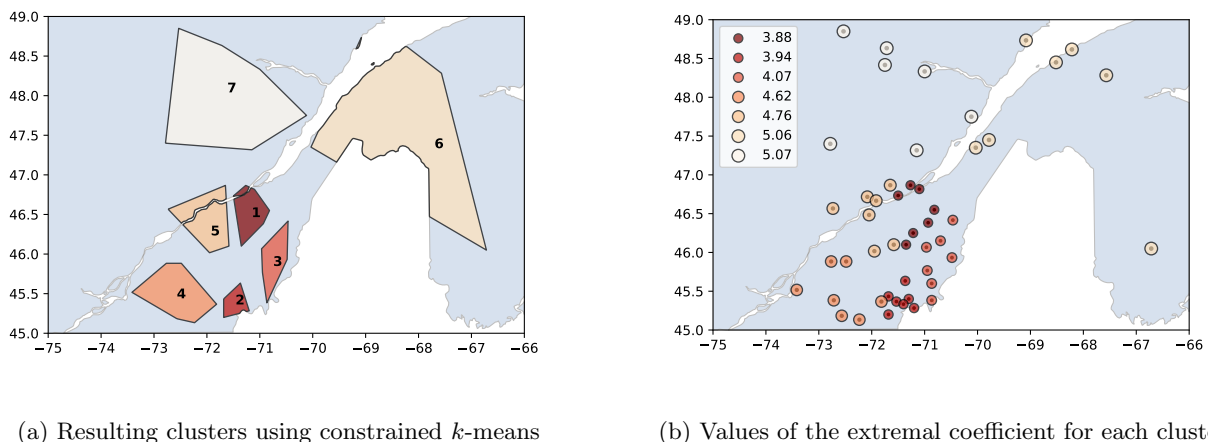


Figure 5: Analysis of Canadian annual rainfall maxima in the period 1914-2017. **(a)** Spatial representation of the 7 selected clusters obtained via the constrained  $k$ -means algorithm. **(b)** Clusters of 49 weather stations and their estimated extremal coefficients (with  $d = 7$ ) obtained with the corrected version of the hybrid madogram.

Our estimations suggest an acute dependence among extremes in clusters 1-3 in Figure 5a. We can observe in Figure 5b that extreme precipitations are more likely to be dependent in the central coastal Atlantic region, *a contrario*, one can notice a weak dependence among extreme values in the scattered clusters in the north of the region.

## 5 Conclusions

A method based on madograms to estimate multivariate extremal dependencies with allowing missing data has been developed in this paper. Under the **MCAR** hypothesis, we studied the asymptotic behaviour for the proposed estimators. This approach is of interest to study spatio-temporal process punctually observed as observations may not overlap. Moreover, we have derived closed expressions of their respective asymptotic variances for a fixed element in the simplex. Numerical results in a finite sample setting give further evidences to our theoretical results and on performances of the proposed estimators of the madogram in the missing data setting. Finally, we applied our approach to the study of extremal dependencies of annual maxima of daily rainfall in Central Eastern Canada.

As for future work, an interesting improvement could be to lower the **MCAR** assumption on the missing data. Indeed, estimating nonparametrically the empirical copula process with missing data outside this framework is still unexplored. As a starting point, semiparametric inference for copula and copula based-regression allowing missing data under **Missing At Random (MAR)** mechanism have been studied by [Hamori et al., 2019] and [Hamori et al., 2020].

Another interesting direction could also be to build a dissimilarity measure based on the bivariate  $\mathbf{w}$ -madogram for clustering. This approach was already tackled by [Bernard et al., 2013], [Bador et al., 2015] and [Saunders et al., 2021] to partition respectively France, Europe and Australia with respect to extreme observations using the sole madogram. The idea here could be to use the infimum or the integral over  $\mathbf{w} \in (0, 1)$  of the bivariate  $\mathbf{w}$ -madogram as a dissimilarity measure and to show its strong consistency in the sense formulated by [Pollard, 1981]. One limitation of our application is that clusters of same size is mandatory to compare the estimated extremal coefficient between clusters in Equation (18). This feature stems from the bounds of the Pickands dependence function which depends on the dimension of the extremal random vector. Further investigations are thus needed to interpret extremal coefficient between clusters of different sizes, *e.g.* to assess asymptotic independence between two extremal random vectors.

## Acknowledgments

This work has been supported by the project ANR McLaren (ANR-20-CE23-0011). This work has been partially supported by the French government, through the 3IA Côte d’Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. This work was also supported by the french national programme LEFE/INSU.

# A Proofs

## A.1 Proofs of main results

For the rest of this section, we will write, for notational convenience,  $n_i = \prod_{j=1}^d I_{i,j}$  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 (12). Before going into details, we need an intermediary lemma to assert that the empirical cumulative distribution functions in case of missing data verify Assumption C and give covariance functions of the asymptotic processes  $\alpha$  and  $\beta_j$  with  $j \in \{1, \dots, d\}$ . 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.** *Let  $(\sqrt{n}(\hat{F}_n - F); \sqrt{n}(\hat{F}_{n,1} - F_1), \dots, \sqrt{n}(\hat{F}_{n,d} - F_d))$  with  $\hat{F}_n$  and  $\hat{F}_{n,j}$  for  $j \in \{1, \dots, d\}$  as in (7). Then Assumption C is satisfied with*

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

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

$$\begin{aligned}\text{cov}(\beta_j(u_j), \beta_j(v_j)) &= p_j^{-1} (u_j \wedge v_j - u_j v_j), \\ \text{cov}(\beta_j(u_j), \beta_k(v_k)) &= \frac{p_{jk}}{p_j p_k} (C(\mathbf{1}_{j,k}(u_j, v_k)) - u_j v_k), \\ \text{cov}(\alpha(\mathbf{u}), \alpha(\mathbf{v})) &= p^{-1} (C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})C(\mathbf{v})), \\ \text{cov}(\alpha(\mathbf{u}), \beta_j(v_j)) &= p_j^{-1} (C(\mathbf{u}_j(u_j \wedge v_j)) - C(\mathbf{u})v_j),\end{aligned}$$

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

Proof of Lemma 1 is postponed to A.2.

**Proof of Theorem 1** First, let us define the rank-corrected hybrid copula process suited with our estimator and its associated empirical copula process by

$$\hat{C}_n^{\mathcal{R}}(\mathbf{u}) = \frac{1}{\sum_{i=1}^n \prod_{j=1}^d I_{i,j}} \sum_{i=1}^n \prod_{j=1}^d \mathbb{1}_{\{\tilde{U}_{i,j} \leq u_j\}} I_{i,j}, \quad \mathbb{C}_n^{\mathcal{R}} = \sqrt{n} (\hat{C}_n^{\mathcal{R}} - C).$$

One can show that

$$\sup_{\mathbf{u} \in [0,1]^d} \left| \hat{C}_n^{\mathcal{H}}(\mathbf{u}) - C_n^{\mathcal{R}}(\mathbf{u}) \right| \leq \frac{2d}{n\hat{p}_n},$$

with  $\hat{p}_n = n^{-1} \sum_{i=1}^n \prod_{j=1}^d I_{i,j}$ . Note that  $\hat{p}_n$  converges in probability to  $p \in ]0, 1]$  which implies that the difference between  $\hat{C}_n^{\mathcal{H}}$  and  $\hat{C}_n^{\mathcal{R}}$  is asymptotically negligible. Details for the proof are given solely for the estimator  $\hat{\nu}_n^{\mathcal{H}*}$  as the weak convergence for  $\hat{\nu}_n^{\mathcal{H}}$  is obtained similarly via an adequate continuous transformation of  $\hat{\nu}_n^{\mathcal{H}}$  with  $\hat{C}_n^{\mathcal{R}}$ . 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

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

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

$$\begin{aligned}\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w}) &= \frac{1}{N} \sum_{i=1}^n g_{\mathbf{w}}(\tilde{U}_{i,1}, \dots, \tilde{U}_{i,d}) \prod_{j=1}^d I_{i,j} + a(\mathbf{w}) = \int_{[0,1]^d} g_{\mathbf{w}}(\mathbf{u}) d\hat{C}_n^{\mathcal{R}}(\mathbf{u}) + a(\mathbf{w}), \\ \nu(\mathbf{w}) &= \int_{[0,1]^d} g_{\mathbf{w}}(\mathbf{u}) dC(\mathbf{u}) + a(\mathbf{w}).\end{aligned}$$

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

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

where  $\mathbf{1}_j(u)$  denotes the vector composed out of 1 except for the  $j$ th component where  $u$  does stand and with  $\mathbb{C}_n^{\mathcal{H}}$  in (9). Consider the function  $\phi : \ell^\infty([0, 1]^d) \rightarrow \ell^\infty(\Delta^{d-1})$ ,  $f \mapsto \phi(f)$ , 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 continuous mapping theorem (see, *e.g.*, Theorem 1.3.6 of [van der Vaart and Wellner, 1996]) implies, as  $n \rightarrow \infty$

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

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  $\alpha$  and  $\beta_j$  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 uses technical arguments to exhibit the closed expressions of the

asymptotic variances of the Gaussian limit distributions of our estimators in Equation (10) and (11). Furthermore, this proof strengthens our choice of the definition of the corrected estimator. Indeed, the chosen form of the corrected estimator makes computations more tractable as we only have to compute terms for the hybrid estimator and to multiply those by different factors. Two tools make the computation feasible. The first one is the form exhibited by Equation (2) which transforms a double integral with respect to the trajectory of the copula function as the double integral 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 the integral with respect to the Pickands dependence function using the following equality :

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

where  $\alpha \neq -1$ .

**Proof of Proposition 1** Recall that  $\mathbf{p} = (p_1, \dots, p_d, p)$ . By definition the asymptotic variance  $\mathcal{S}^{\mathcal{H}}(\mathbf{p}, \mathbf{w})$  for a fixed  $\mathbf{w} \in \Delta^{d-1}$  is given 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 properties of the variance operator, 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 definition of the covariance functions of  $\alpha$ ,  $\beta_j$  with  $j \in \{1, \dots, d\}$  given in Lemma 1, we have for the variance terms

$$\begin{aligned} Var \left( \int_{[0,1]} \alpha(\mathbf{1}_j(x^{w_j})) - \beta_j(x^{w_j}) dx \right) &= (p^{-1} - p_j^{-1}) \sigma_j^2(\mathbf{w}), \\ Var \left( \int_{[0,1]} S_C(x^{w_1}, \dots, x^{w_d}) dx \right) &= \sigma_{d+1}^2(\mathbf{p}, \mathbf{w}). \end{aligned}$$

We obtain similarly for the covariance terms

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

We first show in details the closed form for  $\sigma_{d+1}^2$ , the other forms are given without explanations as the technical tools used are those used for  $\sigma_{d+1}^2$ . Proceeding as before, we decompose this quantity as a linear combination of the variance (the squared term  $\gamma_1^2$  and  $\gamma_j^2$  for  $j \in \{1, \dots, d\}$ ) and the covariance terms ( $\gamma_{1j}$  and  $\tau_{jk}$ ) with the probabilities of missing. The explicit formula of these quantities will be defined below. We set

$$\sigma_{d+1}^2(\mathbf{p}, \mathbf{w}) = p^{-1} \gamma_1^2(\mathbf{w}) + \sum_{j=1}^d p_j^{-1} \gamma_j^2(\mathbf{w}) - 2 \sum_{j=1}^d p_j^{-1} \gamma_{1j}(\mathbf{w}) + 2 \sum_{j < k} \frac{p_{jk}}{p_j p_k} \tau_{jk}(\mathbf{w}). \quad (19)$$

Let us exhibit a useful form of the partial derivatives of the extreme value copula. We have  $\forall j \in \{1, \dots, d\}$  :

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

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  $\mathbf{w} \in \Delta^{d-1}$  :

$$\dot{C}_j(u^{w_1}, \dots, u^{w_d}) = \frac{u^{A(\mathbf{w})}}{u^{w_j}} \dot{\ell}_j(-w_1 \ln(u), \dots, -w_d \ln(u)) = \frac{u^{A(\mathbf{w})}}{u^{w_j}} \dot{\ell}_j(-w_1, \dots, -w_d) = \frac{u^{A(\mathbf{w})}}{u^{w_j}} \mu_j(\mathbf{w}),$$

where  $\mu_j(\mathbf{w}) \triangleq \dot{\ell}_j(-w_1, \dots, -w_d)$ . Now, using linearity of the integral and the definition of the covariance function of  $\alpha$ , we obtain

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

Let us compute

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

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

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

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

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

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

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

Under the rectangle  $[0, 1] \times [0, u]$ , 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

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

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

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

where we use the homogeneity of order one of  $\ell$  and that  $-\ell(-w_1, \dots, -w_d) = A(\mathbf{w})$  as stated by the identity 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  $-\ln(w)$ , we have

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

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 techniques as above, we have

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

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)

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

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

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

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

$$\frac{p_{jk}}{p_j p_k} \tau_{jk}(\mathbf{w}) \triangleq \mathbb{E} \left[ \int_{[0,1]} \beta_j(u^{w_j}) \dot{C}_j(u^{w_1}, \dots, u^{w_d}) du \int_{[0,1]} \beta_k(v^{w_k}) \dot{C}_k(v^{w_1}, \dots, v^{w_d}) dv \right]. \quad (20)$$

Again, we have

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

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

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

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

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

The right side of Equation (20) is given by

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

Hence the result for  $\sigma_{d+1}^2(\mathbf{w})$ . Using the same techniques, we show that for  $j \in \{1, \dots, d\}$

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

For  $j < k$ , we compute

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

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

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

Now, for  $\sigma_{jk}^{(2)}$ , we have to consider three cases :

- if  $j = k$ , we directly have

$$\sigma_{jk}^{(2)}(\mathbf{w}) = 0,$$

- if  $j < k$ , we obtain

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

- if  $j > k$ , we have

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

Hence the statement. □

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

**Proof of Corollary 1** In the term  $\sigma_{d+1}^2$  given in Equation (19), only the terms  $\gamma_1^2$ ,  $\gamma_j^2$  and  $\gamma_{1j}$  matter because, in the independent case :

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

For  $\gamma_{1j}$ , 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  $\gamma_1^2$  and  $\gamma_j^2$ , we just have to set  $A(\mathbf{w}) = 1$  in their respective expressions to obtain :

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

We thus have

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

Other computations follow from the same arguments. □

We are now 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

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

is zero almost surely. Before going into the main arguments, we need the following lemma.

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

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

The proof of Lemma 2 is postponed to A.2.

**Proof of Proposition 2** We prove it for  $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$  as the strong consistency for  $\hat{\nu}_n^{\mathcal{H}^*}(\mathbf{w})$  uses the same arguments. The estimator  $\hat{\nu}_n^{\mathcal{H}}(\mathbf{w})$  in (10) is strongly consistent since it holds

$$|\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu(\mathbf{w})| = |\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu_n(\mathbf{w}) + \nu_n(\mathbf{w}) - \nu(\mathbf{w})| \leq |\hat{\nu}_n^{\mathcal{H}}(\mathbf{w}) - \nu_n(\mathbf{w})| + |\nu_n(\mathbf{w}) - \nu(\mathbf{w})|,$$

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

$$|\nu_n(\mathbf{w}) - \nu(\mathbf{w})| \xrightarrow[n \rightarrow \infty]{a.s.} 0$$

For the second term, we write :

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

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

Finally, we give some elements to establish Corollary 2. The strong consistency follows directly from the stability of the almost surely convergence through a continuous function. The weak convergence comes down from the functional Delta method (see, *e.g.*, Theorem 3.9.4 of [van der Vaart and Wellner, 1996]) and from result in Proposition 1.

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

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

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

$$\sqrt{n} \left( \hat{A}_n^{\mathcal{H}^*}(\mathbf{w}) - A(\mathbf{w}) \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \mathcal{V}(\mathbf{p}, \mathbf{w})),$$

with,  $\mathcal{V}(\mathbf{p}, \mathbf{w})$  equals by definition

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

where we used Proposition 1 to conclude. □

## A.2 Proofs of auxiliary results

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

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

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 (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  $P$ -brownian bridge  $\mathbb{G}$ . For  $\mathbf{x} \in \mathbb{R}^d$ ,

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

We obtain for the second one

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

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 and Assumption B 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, Assumption C is fulfilled with for  $\mathbf{u} \in [0, 1]^d$ ,

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

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  $u_j, v_k \in [0, 1]$

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

Hence the result. □

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

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

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

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

Moreover, by symmetry of  $\tilde{U}_{i,j}$  and  $F_j$ , the second one follows similarly.  $\square$

## References

- [Bador et al., 2015] Bador, M., Naveau, P., Gilleland, E., Castellà, M., and Arivelo, T. (2015). Spatial clustering of summer temperature maxima from the cnrm-cm5 climate model ensembles & e-obs over europe. *Weather and Climate Extremes*, 9:17–24.
- [Bernard et al., 2013] Bernard, E., Naveau, P., Vrac, M., and Mestre, O. (2013). Clustering of Maxima: Spatial Dependencies among Heavy Rainfall in France. *Journal of Climate*, 26(20):7929–7937.
- [Bradley et al., 2000] Bradley, P. S., Bennett, K. P., and Demiriz, A. (2000). Constrained k-means clustering. *Microsoft Research, Redmond*, 20(0):0.
- [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.
- [Demarta and McNeil, 2005] Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas. *International Statistical Review*, 73(1):111–129.
- [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.
- [Fonseca et al., 2015] Fonseca, C., Pereira, L., Ferreira, H., and Martins, A. P. (2015). Generalized madogram and pairwise dependence of maxima over two regions of a random field. *Kybernetika*, 51(2):193–211.
- [Genest and Segers, 2009] Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value cop-

- ulas. *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.
- [Gumbel, 1960] Gumbel, E. (1960). Distributions de valeurs extrêmes en plusieurs dimensions. *Publications de l’institut de Statistique de l’Université de Paris*, 9:171–173.
- [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.
- [Hamori et al., 2019] Hamori, S., Motegi, K., and Zhang, Z. (2019). Calibration estimation of semiparametric copula models with data missing at random. *Journal of Multivariate Analysis*, 173:85–109.
- [Hamori et al., 2020] Hamori, S., Motegi, K., and Zhang, Z. (2020). Copula-based regression models with data missing at random. *Journal of Multivariate Analysis*, 180:104654.
- [Hüsler and Reiss, 1989] Hüsler, J. and Reiss, R.-D. (1989). Maxima of normal random vectors: Between independence and complete dependence. *Statistics & Probability Letters*, 7(4):283–286.
- [Joe, 1990] Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistics & Probability Letters*, 9:75–81.
- [Marcon et al., 2017] Marcon, G., Padoan, S., Naveau, P., Muliere, P., and Segers, J. (2017). Multivariate non-parametric 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.



- [Oliveira and Galambos, 1977] Oliveira, J. D. T. and Galambos, J. (1977). The asymptotic theory of extreme order statistics. *International Statistical Review*, 47:230.
- [Palacios-Rodriguez et al., 2021] Palacios-Rodriguez, F., Di Bernardino, E., and Mailhot, M. (2021). Smooth Copula-based Generalized Extreme Value model and Spatial Interpolation for Sparse Extreme Rainfall in Central Eastern Canada. Preprint <https://hal.archives-ouvertes.fr/hal-03355026>.
- [Patton and Wiley, 2006] Patton, A. J. and Wiley, J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, pages 147–173.
- [Pickands, 1981] Pickands, J. (1981). Multivariate extreme value distribution. *Proceedings 43th, Session of International Statistical Institution, 1981*.
- [Pollard, 1981] Pollard, D. (1981). Strong Consistency of  $K$ -Means Clustering. *The Annals of Statistics*, 9(1):135 – 140.
- [Saunders et al., 2021] Saunders, K., Stephenson, A., and Karoly, D. (2021). A regionalisation approach for rainfall based on extremal dependence. *Extremes*, 24.
- [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.
- [Smith, 1990] Smith, R. L. (1990). Max-stable processes and spatial extremes. Unpublished Manuscript.
- [Tawn, 1988] Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, 75(3):397–415.
- [Tawn, 1990] Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika*, 77(2):245–253.
- [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.