BOULIN Alexis

ENSAE 3ème année
Stage de fin d'études
Année scolaire 2020 - 2021

Non parametric estimation of the Madogram in missing data and outliers setting

LJAD
Parc Valrose, Nice

Maîtres de stage : Elena DI BERNARDINO
Thomas LALOË
Gwladys TOULEMONDE

17/05/2021 - 30/09/2021

Résumé (Français)

Dans ce travail, nous nous intéressons à l'estimation statistique du Madogramme en présence de données manquantes et d'*outliers*. Dans le premier cadre, nous avons à notre disposition un échantillon bivarié de T observations, indépendantes et de même loi mais certaines observations peuvent manquer dans l'une des composantes ou les deux. Dans le second, on observe un échantillon bivarié de T pour lequel une fraction des réalisations est corrompue. Dans le premier modèle, notre objectif est de construire un estimateur du Madogramme pour lequel le processus d'estimation doit extraire le maximum d'informations du jeu de données tronqué. Dans le second modèle, il s'agit de construire un estimateur qui prenne en compte la présence des données corrompues et, dans son processus d'estimation, diminue leur influence.

Dans ces deux modèles, nous nous intéressons plus particulièrement à l'estimation non paramétrique du Madogramme. Un théorème central limite fonctionnel est démontré en présence de données manquantes. Lorsque la trajectoire du processus limite est fixée, nous démontrons la convergence en distribution vers une Gausienne centrée dont nous donnons une expression fermée de la variance limite. Pour le second cadre, nous démontrons une inégalité de concentration de notre estimateur assurant sa robustesse face aux outliers.

Dans chaque modèle, des simulations sont apportées afin d'attester nos résultats lorsque l'échantillon est de taille finie. Ces simulations sont réalisées sur différentes copules à valeurs extrêmes.

Abstract (English)

In this work, we are interested in statistical inference of the Madogram with missing data and outliers setting. Two models are investigated. The first one consists of the observations of T independent and distributed accordingly from a bivariate random vector while some observations might be missing in each or both coordinate. The second one, we observe a bivariate sample of length T in which some observations might be corrupted. In the first model, our goal is to estimate the Madogram by getting as much information as possible from the truncated dataset. In the second model, we aim to construct an estimation process that is less sensitive to the presence of outliers.

In both models, we investigate the case of non-parametric estimation of the Madogram. A central limit functional theorem is proven in the missing data framework. When the path of the limit process is fixed, we show the convergence in law towards a centered Gaussian where we give a closed expression of the limit variance. For the second framework, we prove a concentration inequality that our estimator does verify that assert the robustness against outliers.

In each model, simulations are made to assert our results in finite sample settings. These simulations are performed on various copulas with extreme values.

Contents

Introduction

Context

Management of environmental resources often requires the analysis of multivariate extreme values. In climate studies, extreme events such as heavy precipitation and record temperatures represent a major challenges due to their consequences. In the classical statistical theory, one is often interested in the behavior of the mean of a random variable X defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. This average will then be described through the expected value $\mathbb{E}[X]$ of the distribution. But in case of extreme events, it can be just as important to estimate tails probabilities. Furthermore, what if the second moment $\mathbb{E}[X^2]$ or even the mean is not finite? Then the central limit theorem does not apply and the classical theory, carried by the normal distribution, is no longer relevant.

Also, inference methods for assessing dependence is increasingly in demand. The most popular approach is based on second moment of the underlying random variables, the covariance. It is well known that only linear dependence can be captured by the covariance that it is characterizing only for a few special classes of distribution. As a beneficial alternative of dependence modeling, the concept of copulas going back to [Sklar, 1959]. The copula $C : [0,1]^2 \rightarrow [0,1]$ of a random vector (X,Y) allows us to separate the effect of dependence from the effects of the marginal distribution such as :

$$\mathbb{P}\left\{X \leq x, Y \leq y\right\} = C\left(\mathbb{P}\left\{X \leq x\right\}, \mathbb{P}\left\{Y \leq y\right\}\right), \quad \forall (x,y) \in \mathbb{R}^2.$$

The main consequence of this identity is that the copula completely characterizes the stochastic dependence between X and Y. Investigating the notion of copulas within the framework of multivariate extreme value theory leads to the so-called extreme value copulas.

Moreover, some extreme events, such as heavy precipitation or wind speed has also spatial characteristics and geostatisticians are striving to better understand the physical process in hands. In geostatistic, we often consider $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space, S a set of locations and $(E, \mathcal{E})$ a measurable state space. One can define on this probability space a stochastic process $X = \{X(s), s \in S\}$ with values on $(E, \mathcal{E})$. It is classical to define the following second-order statistic as the variogram (see [Gaetan and Guyon, 2008] Chapter 1.3 for definition and basic properties) :

$$2\gamma(h) = \mathbb{E}[|X(s+h) - X(s)|^2],$$

where $\{X(s), s \in S\}$ represents a spatial and stationary process with a well-defined covariance function. The function $\gamma(\cdot)$ is called the semi-variogram of X. With respect to extremes, this definition is not well adapted because a second order statistic is difficult to interpret inside the framework of extreme value theory or may not even be defined. To ensure to always work with finite moments quantities, the following type of first-order variogram is introduced by [Cooley et al., 2006]

$$\nu(h) = \frac{1}{2}\mathbb{E}\left[|F(X(s+h)) - F(X(s))|\right], \tag{1}$$

where $F(u) = \mathbb{P}(X(s) \leq u)$ is named as the FMadogram. As $F(X)$ is distributed accordingly to the uniform distribution under the segment $[0,1]$, the problem of moments' existence is no longer relevant. This quantity is alo link to the pairwise extremal dependence function (Section 4.3 of [Coles et al., 1999]) or the Pickands dependence function ([Pickands, 1981]) make it an interesting quantity to capture the dependence between the extrema of stochastic processes or random variables. Furthermore, this quantity may be seen as a dissimilarity measure among bivariate maxima to be used for clustering time-series as shown by [Bernard et al., 2013] or [Bador et al., 2015].

The main drawback of this quantity is that it only focus on the value of the diagonal section of the pairwise extremal dependence function. In the bivariate case, the FMadogram characterize solely the extremal coefficient for random variables X and Y (see Section 8.2.7 of [Beirlant et al., 2004]). To overpass this drawback, [Naveau et al., 2009] introduce the λ-FMadogram defined as,

$$\nu(h, \lambda) = \frac{1}{2}\mathbb{E}\left[\left|F^{\lambda}(X(s+h)) - F^{1-\lambda}(X(s))\right|\right], \tag{2}$$

for every $\lambda \in [0,1]$. This quantity characterizes the pairwise extremal dependence function outside the diagonal section but also the whole Pickands dependence function (see [Marcon et al., 2017]) and contribute to the vast literature of the estimation of the Pickands dependence function for bivariate extreme value copulas (see for example [Pickands, 1981], [Deheuvels, 1991], [Capéraà et al., 1997] or [Hall and Tajvidi, 2000]). Statisticians may estimate these quantities, but the classical results such as strong consistency or weak convergence may only apply if the data in hands are clean as possible. This induces that the process of data collection has not been corrupted such as the dataset is complete, each observation is independent from others and that the implicit law of the observations is still the same.

Unfortunately, as the volume of data expands, the problem of missing or contaminated data is present in many fields. It frequently happens that some individuals of a sample from a multivariate population are not observed. If a sample is represented in matrix form by allowing rows to represent the individuals and columns the variables, then we may have to deal with a sparse matrix. In dealing with fragmentary samples, it is important to have at hand techniques which will enable the statistician to extract as much information as possible from the data. A useful reference for general parametric statistical inferences with missing data was provided by [Little R.J.A., 1987].

Consider a sample from a random vector (X, Y) of incomplete data,

$$(X_t, Y_t, \delta_t), \quad t \in \{1, \dots, T\}, \tag{3}$$

where all the X_t's are observed and $\delta_t = 0$ if Y_t is missing, otherwise $\delta_t = 1$. The simple missing data pattern described by (3) is basically created by the double sampling or two phase sampling (see Chapter 12 of [Cochran, 2007]). Samples like (3) may arise in survival analysis : the study of the duration time preceding an event of interest is considered with a series of random censors, which might prevent the capture of the whole survival time. This is known as the censoring mechanism and it arises from restrictions depending of the nature of the study. Typically, they may occur in medicine, with studies of the survival times before the recovery / decease from a specific disease. Another important example is often realized in comparing treatment effects of two educational programs. Individuals with lower scores on a preliminary test are more likely to receive the experimental treatment (*i.e.*, a compensatory study program), whereas those with higher preliminary scores are more inclined to take the standard control. This phenomenon is well-known as the selection problem and we refer to Chapter 2 of [Angrist and Pischke, 2008] for more details. Beside of missing observations, the process of data might be disturb in a way that innerly deteriorate the quality of some data and one may ask that the estimation process of the Madograms in Equation (1) and (2) be robust. The topic of robustness in estimation has known an important research activity developed in the 60's and 70's resulting in a large number of publications. For a summary, the interested reader is referred to [Huber, 2011]. Robustness can be seen as an estimation procedure in which both stochastic and approximation errors are low (see Section 1.1 from [Baraud et al., 2016]). In other words, an estimator is robust if a not well specified model provides a reasonable approximation of the true one and derive an estimator which remains close to the true distribution. In this report, we mean by *robust* as *robust against outlier*, *e.g.* the ϵ-contamination model (see [Huber, 1964]), or *robust again heavy-tailed data* where only low-order moments are assumed to be finite for the data distribution. There is no simple relation between the two definitions and the first framework of robustness that we have depicted. It is a main goal of this report to develop estimates of the (λ)-FMadogram in these sticky situations. To achieve our goal, we make of use of the hybrid copula estimator ([Segers, 2014]) for the missing data framework and we leverage the idea of Median-of-meaNs (MoN, see [Nemirovsky and Yudin, 1983]) for the contaminated data scheme.

Definitions and Notation

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (X, Y) be a bivariate random vector with values in $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. This random vector has a joint distribution function H and marginal distribution function F and G. A function $C : [0,1]^2 \rightarrow [0,1]$ is called a bivariate copula if it is the restriction to $[0,1]^2$ of a bivariate distribution function whose margins are given by the uniform distribution on the interval $[0,1]$. Since the work of [Sklar, 1959], it is well known that every distribution function H can be decomposed as $H(x,y) = C(F(x), G(y))$, for all $(x,y) \in \mathbb{R}^2$.

Definition 1 ([Gudendorf and Segers, 2009]). *A bivariate copula C is an extreme-value copula if and only if it admits a representation of the form*

$$C(u,v) = (uv)^{A\left(\frac{log(v)}{log(uv)}\right)}, \quad (u,v) \in (0,1]^2 \setminus \{(1,1)\} \tag{4}$$

where $A(\cdot)$ is the Pickands dependence function, i.e., *$A : [0,1] \longrightarrow [1/2, 1]$ is convex and satisfies $t \vee (1-t) \leq A(t) \leq 1$, $\forall t \in [0,1]$.*

The upper and lower bound of A has special meanings : the upper bound $A(t) = 1$ corresponds to independence, whereas the lower bound $A(t) = t \vee (1-t)$ corresponds to the perfect dependence (comonotonicity). Notice that, on sections, the extreme value copula is of the form

$$C(u^{1-t}, u^t) = u^{A(t)}. \tag{5}$$

Let $(X_t, Y_t)_{t=1,\dots,T}$ be an *i.i.d.* sample of a bivariate random vector whose underlying copula is denoted by C and whose margins by F, G. For $x, y \in \mathbb{R}$, let $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$. Let $(b_{t,j})_{t \geq 1, j \in \{1,2\}}$ and $(a_{t,j})_{t \geq 1, j \in \{1,2\}}$ be respectively sequences of numbers and sequences of positive reals. We say that the sequence $(a_{T,1}^{-1}(\bigvee_{t=1}^T X_t - b_{T,1}), a_{T,2}^{-1}(\bigvee_{t=1}^T Y_t - b_{T,2}))$ belongs to the domain of attraction of H, if for all real values x, y (at which the limit is continuous and non-degenerate)

$$\mathbb{P}\left(\frac{\bigvee_{t=1}^T X_t - b_{T,1}}{a_{T,1}} \leq x, \frac{\bigvee_{t=1}^T Y_t - b_{T,2}}{a_{T,2}} \leq y \right) \underset{T \to \infty}{\to} H(x, y).$$

If this relationship hold, H is said to be a multivariate extreme value distribution. The FMadogram is one of the quantity used in dependence modeling to characterize the extremal coefficient, *i.e.* in the bivariate case $2A(2^{-1})$. The FMadogram is then defined by the following quantity

$$\nu = \frac{1}{2}\mathbb{E}\left[|F(X) - G(Y)|\right], \tag{6}$$

and the λ-FMadogram by

$$\nu(\lambda) = \frac{1}{2}\mathbb{E}\left[\left|F(X)^\lambda - G(Y)^{1-\lambda}\right|\right]. \tag{7}$$

Suppose that we observe sequentially a quadruple defined by

$$(I_t X_t, J_t Y_t, I_t, J_t), \quad t \in \{1, \dots, T\}, \tag{8}$$

where $I_t = 0$ (resp. $J_t = 0$) if X_t (resp. Y_t) is missing, otherwise $I_t = 1$ (resp. $J_t = 1$), *i.e.* at each $t \in \{1, \dots, T\}$, one or both entries may be missing. The probability of observing a realization partially or completely, is denoted by $p_X = \mathbb{P}(I_t = 1) > 0$, $p_Y = \mathbb{P}(J_t = 1) > 0$, $p_{XY} = \mathbb{P}(I_t = 1, J_t = 1) > 0$ and we note $\mathbf{p} = (p_X, p_Y, p_{XY})$. Let us now define the empirical cumulative distribution of X (resp. Y and (X, Y))

in case of missing data,

$$\hat{F}_T(u) = \frac{\sum_{t=1}^T 1_{\{X_t \leq u\}} I_t}{\sum_{t=1}^T I_t}, \; \hat{G}_T(v) = \frac{\sum_{t=1}^T 1_{\{Y_t \leq v\}} J_t}{\sum_{t=1}^T J_t}, \; \hat{H}_T(u,v) = \frac{\sum_{t=1}^T 1_{\{X_t \leq u, Y_t \leq v\}} I_t J_t}{\sum_{t=1}^T I_t J_t}.$$

Here, we weight the estimator by the number of observed data which is a natural estimator if divided by T of the probabilities of missing. We have all tools in hand to recall the definition of the *hybrid copula estimator* introduced by [Segers, 2014],

$$\hat{C}_T^{\mathcal{H}}(u,v) = \hat{H}_T(\hat{F}_T(u), \hat{G}_T(v)).$$

The normalized estimation error of the hybrid copula estimator is

$$\mathbb{C}_T^{\mathcal{H}}(u,v) = \sqrt{T}\left(\hat{C}_T^{\mathcal{H}}(u,v) - C(u,v)\right).$$

We write the generalized inverse function of F (respectively G) as $F^{\leftarrow}(u) = inf\{v \in \mathbb{R} | F(v) \geq u\}$ (respectively $G^{\leftarrow}(u) = inf\{v \in \mathbb{R} | G(v) \geq u\}$) where $0 < u, v < 1$. Given $\mathcal{X} \subset \mathbb{R}^2$, let $\ell^\infty(\mathcal{X})$ denote the spaces of bounded real-valued function on $\mathcal{X}$. For $f : \mathcal{X} \to \mathbb{R}$, let $||f||_\infty = sup_{x \in \mathcal{X}} |f(x)|$. Here, we use the abbreviation $Qf = \int f dQ$ for a given measurable function f and signed measure Q. The arrows $\overset{a.s.}{\to}$, $\overset{d}{\to}$ denote almost sure convergence and convergence in distribution of random vectors. Weak convergence of a sequence of maps will be understood in the sense of J.Hoffman-Jørgensen (see Part 1 in the monograph by [van der Vaart and Wellner, 1996]). Given that $t \in \mathbb{N}^*, X, X_t$ are maps from $(\Omega, \mathcal{A}, \mathbb{P})$ into a metric space $\mathcal{X}$ and that X is Borel measurable, $(X_t)_{t \geq 1}$ is said to converge weakly to X if $\mathbb{E}^* f(X_t) \to \mathbb{E} f(X)$ for every bounded continuous real-valued function f defined on $\mathcal{X}$, where $\mathbb{E}^*$ denotes outer expectation in the event that X_t may not be Borel measurable. In what follows, weak convergence is denoted by $X_t \rightsquigarrow X$.

This work is organized as follows. In Chapter 1 we state some results on the weak convergence of the estimator of the λ-FMadogram with missing and complete data. We also propose a closed formula for the asymptotic variance of the λ-FMadogram for a fixed $\lambda \in [0, 1]$. To propose a robust estimator of the FMadogram, we leverage the idea of Median-of-meaNs (MoN) and state a concentration inequality that this estimator does verify. Chapter 2 will present the performance of our estimator in a finite-sample framework. The asymptotic variance of the normalized estimation error of several models will be drawn with their empirical counterpart obtained through simulation. We also propose a reproduction of the experiment of the λ-FMadogram with a Smith's process as in [Naveau et al., 2009] and we will explain the augmentation of the Mean Squared Error while h is close to zero, *i.e.* where two locations are strongly dependent. This phenomenon will be also explained with simulations and a counterexample. All codes are available on Github[1]. For the ease of reading, we postponed all technical arguments and proofs in Chapter 3. Three appendices may be find in end. In Appendix A, we study properties of the Pickands dependence function and assert the weak convergence of our empirical copula process for extreme value copula. Appendix B recall some results of the litterature that are used in this report. Supplementary results are given in Appendix C.

[1] https://github.com/Aleboul/var_FMado

Chapter 1

Non parametric estimation of the Madogram with missing data

1.1 Definition of the estimator

Under the notation of the introduction, we assume that the copula C is of extreme value type as in Definition 1. To guarantee the weak convergence of our empirical copula while C, we make the following assumption as suggested in [Segers, 2012] in Example 3.5.

Assumption A.

(i) The bivariate distribution function H has continuous margins F and G.

(ii) The derivative of the Pickands dependence function $A'(t)$ exists and is continuous on $(0,1)$.

The Assumption A *(i)* guarantees that the representation $H(x,y) = C(F(x), G(y))$ is unique on the range of (F, G). Under the Assumption A *(ii)*, the first-order partial derivatives of C with respect to u (*resp.* with respect to v) exists and is continuous on the set $\{(u,v) \in [0,1]^2 : 0 < u < 1\}$ (*resp.* on the set $\{(u,v) \in [0,1]^2 : 0 < v < 1\}$). This point is discussed in Appendix A by using Lemma A.1 and Lemma A.2. We now define our estimator of Equation (7) in the general content (allowing missing data).

Definition 2. *Let $(I_1X_1, J_1Y_1, I_1, J_1), \dots, (I_TX_T, J_TY_T, I_T, J_T)$ be a sample given by Equation (8), we define the hybrid estimator of the λ-FMadogram by*

$$\hat{\nu}_T^{\mathcal{H}}(\lambda) = \frac{1}{2\sum_{t=1}^T I_t J_t} \sum_{t=1}^T \left| \hat{F}_T(X_t)^\lambda - \hat{G}_T(Y_t)^{1-\lambda} \right| I_t J_t. \tag{1.1}$$

One may verify that in the complete data framework, *i.e.* with $p_X = p_Y = p_{XY} = 1$ we retrieve the λ-FMadogram such as defined in [Naveau et al., 2009], namely

$$\hat{\nu}_T(\lambda) = \frac{1}{2T} \sum_{t=1}^T \left| \hat{F}_T^\lambda(X_t) - \hat{G}_T(Y_t)^{1-\lambda} \right|,$$

with $\hat{F}_T$ (resp. $\hat{G}_T$) the empirical cumulative distribution function of X (resp. Y).

Remark 1. *Our estimator defined in (1.1) does not verify $\hat{\nu}_T^{\mathcal{H}}(0) = \hat{\nu}_T^{\mathcal{H}}(1) = 0.25$ while $\nu(0) = \nu(1) = 0.25$. In addition, the variance at $\lambda = 0$ or $\lambda = 1$ does not equal 0. Indeed, suppose that we evaluate this statistic at $\lambda = 0$, we thus obtain the following quantity :*

$$\hat{\nu}_T^{\mathcal{H}}(0) = \frac{1}{2\sum_{t=1}^T I_t J_t} \sum_{t=1}^T \left(1 - \hat{G}_T(Y_t)\right) I_t J_t.$$

In this situation, the sample $(X_t)_{t=1}^T$ is taken into account through the indicators sequence $(I_t)_{t=1}^T$ and induce a supplementary variance when estimating.

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

Definition 3. *Under the notation of Definition 2, we define the hybrid corrected estimator of the λ-FMadogram by*

$$\begin{aligned}
\hat{\nu}_T^{\mathcal{H}*}(\lambda) = & \frac{1}{2\sum_{t=1}^T I_t J_t} \sum_{t=1}^T \left| \hat{F}_T(X_t)^\lambda - \hat{G}_T(Y_t)^{1-\lambda} \right| I_t J_t \\
& - \frac{\lambda}{2\sum_{t=1}^T I_t J_t} \sum_{t=1}^T \{1 - \hat{F}_T(X_t)^\lambda\} I_t J_t \\
& - \frac{1-\lambda}{2\sum_{t=1}^T I_t J_t} \sum_{t=1}^T \{1 - \hat{G}_T(Y_t)^{1-\lambda}\} I_t J_t + \frac{1-\lambda+\lambda^2}{2(2-\lambda)(1+\lambda)}
\end{aligned} \tag{1.2}$$

Remark 2. *In the missing data framework, the asymptotic behaviour of the process* $\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right)$ *is not the same as the process* $\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda)\right)$ *and they should be studied apart.*

Let us now introduce a condition on the missing mechanism :

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

Under this Assumption, we state the strong consistency of our hybrid estimator of the λ-FMadogram.

Proposition 1 (Strong consistency). *Let* $(I_1 X_1, J_1 Y_1, I_1, J_1), \dots, (I_T X_T, J_T Y_T, I_T, J_T)$ *a i.i.d sample given by Equation (8). We have, under Assumption B for a fixed* $\lambda \in [0,1]$*, as* $T \to \infty$

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

Details on the proof are given in Section 3.1.

1.2 Functional central limit theorem with possible missing data

In this section, we present with Theorem 1 our main result concerning the weak convergence of the following processes

$$\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right), \quad \sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda)\right). \tag{1.3}$$

Without missing data, the weak convergence of the normalized estimation error of the empirical copula process has been proved by [Fermanian et al., 2004] under a more restrictive condition than Assumption A. The difference being that C should be continuously differentiable on the closed cube. This statement make use of previous results on the Hadamard differentiability of the map $\phi : D([0,1]^2) \to \ell^\infty([0,1]^2)$ which transforms the cumulative distribution function H into its copula function C (see Lemma 3.9.28 from [van der Vaart and Wellner, 1996]). With the hybrid copula estimator, we need a following technical assumption in order to guarantee the weak convergence of the process $\mathbb{C}_T^{\mathcal{H}}$ (see [Segers, 2014]),

Assumption C. *In the space $\ell^\infty(\mathbb{R}^2) \otimes (\ell^\infty(\mathbb{R}), \ell^\infty(\mathbb{R}))$ equipped with the topology of uniform convergence, we have the joint weak convergence*

$$\left(\sqrt{T}(\hat{H}_T - H); \sqrt{T}(\hat{F}_T - F), \sqrt{T}(\hat{G}_T - G)\right) \rightsquigarrow (\alpha \circ (F, G), \beta_1 \circ F, \beta_2 \circ G) .$$

The stochastic processes α and $\beta_j, j \in \{1, 2\}$ take values in $l^\infty([0, 1]^2)$ and $l^\infty([0, 1])$ respectively, and are such that $\alpha \circ F$ and $\beta_j \circ F_j$ have continuous trajectories on $[-\infty, \infty]^2$ and $[-\infty, \infty]$ almost surely.

Under Assumptions A and C (see Theorem B.2 on page vi), the stochastic process $\mathbb{C}_T^{\mathcal{H}}$ converges weakly to the tight Gaussian process S_C defined by,

$$S_C(u, v) = \alpha(u, v) - \frac{\partial C(u, v)}{\partial u}\beta_1(u) - \frac{\partial C(u, v)}{\partial v}\beta_2(v), \quad \forall (u, v) \in [0, 1]^2.$$

Considering the same statistical framework and missing mechanism, [Segers, 2014] shows (in Example 3.5) that the processes α, β_1 and β_2 take the following closed form

$$\begin{aligned}
\beta_1(u) &= p_X^{-1}\mathbb{G}\left(\mathbb{1}_{X \le F^{\leftarrow}(u), I=1} - u\mathbb{1}_{I=1}\right), \\
\beta_2(v) &= p_Y^{-1}\mathbb{G}\left(\mathbb{1}_{Y \le G^{\leftarrow}(v), J=1} - v\mathbb{1}_{J=1}\right), \\
\alpha(u, v) &= p_{XY}^{-1}\mathbb{G}\left(\mathbb{1}_{X \le F^{\leftarrow}(u)}\mathbb{1}_{Y \le G^{\leftarrow}(v), I=1, J=1} - C(u, v)\mathbb{1}_{I=1, J=1}\right).
\end{aligned}$$

Furthermore, we are able to compute their covariance functions given in the following lemma.

Lemma 1. *The covariance function of the process $\beta_1(u)$, $\beta_2(v)$ and $\alpha(u, v)$ are,*

for $(u, u_1, u_2, v, v_1, v_2) \in [0, 1]^6$,

$$\begin{aligned}
cov\left(\beta_1(u_1), \beta_1(u_2)\right) &= p_X^{-1}\left(u_1 \wedge u_2 - u_1 u_2\right), \\
cov\left(\beta_2(v_1), \beta_2(v_2)\right) &= p_Y^{-1}\left(v_1 \wedge v_2 - v_1 v_2\right), \\
cov\left(\beta_1(u), \beta_2(v)\right) &= \frac{p_{XY}}{p_X p_Y}\left(C(u, v) - uv\right),
\end{aligned}$$

and

$$
\begin{aligned}
&cov\left(\alpha(u_1,v_1),\alpha(u_2,v_2)\right) = p_{XY}^{-1}\left(C(u_1\wedge u_2, v_1\wedge v_2) - C(u_1,v_1)C(u_2,v_2)\right),\\
&cov\left(\alpha(u_1,v),\beta_1(u_2)\right) = p_X^{-1}\left(C(u_1\wedge u_2, v) - C(u_1,v)u_2\right),\\
&cov\left(\alpha(u,v_1),\beta_2(v_2)\right) = p_Y^{-1}\left(C(u, v_1\wedge v_2) - C(u,v_1)v_2\right).
\end{aligned}
$$

Proof of Lemma 1 is postponed to Section 3.1.

We have all tools in hand to consider the weak convergence of the stochastic processes in Equation (1.3).

Theorem 1 (**Functional central limit theorem with missing data**). *Under Assumptions A, B, C we have the weak convergence in $\ell^\infty([0,1])$ for the hybrid estimator defined in (1.1) and (1.2), as $T\to\infty$,*

$$
\begin{aligned}
\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda)\right) \rightsquigarrow \Bigg(&\frac{1}{2}\int_{[0,1]} \alpha(x^{\frac{1}{\lambda}},1) - \beta_1(x^{\frac{1}{\lambda}})dx + \frac{1}{2}\int_{[0,1]} \alpha(1,x^{\frac{1}{1-\lambda}}) - \beta_2(x^{\frac{1}{1-\lambda}})dx \\
&- \int_{[0,1]} S_C(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx\Bigg)_{\lambda\in[0,1]},
\end{aligned}
$$

$$
\begin{aligned}
\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right) \rightsquigarrow \Bigg(&\frac{1-\lambda}{2}\int_{[0,1]} \alpha(x^{\frac{1}{\lambda}},1) - \beta_1(x^{\frac{1}{\lambda}})dx + \frac{\lambda}{2}\int_{[0,1]} \alpha(1,x^{\frac{1}{1-\lambda}}) - \beta_2(x^{\frac{1}{1-\lambda}})dx \\
&- \int_{[0,1]} S_C(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx\Bigg)_{\lambda\in[0,1]},
\end{aligned}
$$

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

For a given value of λ, it is possible to precise things. Indeed, as an integral of a tight Gaussian process, we know that the two normalized estimation errors follows a centered Gaussian variable for a given $\lambda\in[0,1]$. Furthermore, some computations are able to give a closed form of the variance of the limiting Gaussian law as an integral of the Pickands dependence function. This is summarized with the following result.

Proposition 2 (**Asymptotic variance closed formula**). *For $\lambda\in[0,1]$, let $A_1(\lambda) =$*

$A(\lambda)/\lambda$, $A_2(\lambda) = A(\lambda)/(1-\lambda)$. *Then, under the framework of Theorem 1, the random variables* $\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda)\right)$ *and* $\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right)$ *converges in distribution toward a centered Gaussian random variable with variance which has the following closed form*

$$\mathcal{S}^{\mathcal{H}}(\boldsymbol{p}) = \frac{1}{4}\sigma_1^2(p_X, p_{XY}) + \frac{1}{4}\sigma_2^2(p_Y, p_{XY}) + \sigma_3^2(\boldsymbol{p}) + \frac{1}{2}\sigma_{12}(\boldsymbol{p}) - \sigma_{13}(\boldsymbol{p}) - \sigma_{23}(\boldsymbol{p}), \tag{1.4}$$

$$\begin{aligned}\mathcal{S}^{\mathcal{H}*}(\boldsymbol{p}) =& \frac{(1-\lambda)^2}{4}\sigma_1^2(p_X, p_{XY}) + \frac{\lambda^2}{4}\sigma_2^2(p_Y, p_{XY}) + \sigma_3^2(\boldsymbol{p}) \\ &+ \lambda(1-\lambda)\frac{1}{2}\sigma_{12}(\boldsymbol{p}) - (1-\lambda)\sigma_{13}(\boldsymbol{p}) - \lambda\sigma_{23}(\boldsymbol{p}),\end{aligned} \tag{1.5}$$

where expressions of the functions $\sigma_1^2(p_X, p_{XY}), \sigma_2^2(p_Y, p_{XY}), \sigma_3^2(\boldsymbol{p})$ *and* $(\sigma_{ij}(\boldsymbol{p}))_{i,j\in\{1,2,3\}}$ *with* $i < j$ *are defined in 3.1 for the sake of readibility.*

Proof is postponed in Section 3.1. Note that there is no strict dominance between the variances of the limiting processes in Equation (1.4) and Equation (1.5) for a given λ. Considering the special case of independent copula, Corollary 1 below gives a closed form of the limit variance which no longer depend of the Pickands dependence function.

Corollary 1 (**Asymptotic variance in the independent copula case**). *In the framework of Theorem 1 and if* $C(u,v) = uv$, *then the functions* $\sigma_3^2(\boldsymbol{p})$ *and* $(\sigma_{ij}(\boldsymbol{p}))_{i,j\in\{1,2,3\}}$ *with* $i < j$ *has the following form, for* $\lambda \in [0,1]$

$$\begin{aligned}\sigma_3^2(\boldsymbol{p}) &= \left(\frac{\lambda(1-\lambda)}{1+\lambda(1-\lambda)}\right)^2 \left(\frac{p_{XY}^{-1}}{1+2\lambda(1-\lambda)}\right. \\ &\qquad \left. - \frac{p_X^{-1}(1-\lambda)}{2-(1-\lambda)+2\lambda(1-\lambda)} - \frac{p_Y^{-1}\lambda}{2-\lambda+2\lambda(1-\lambda)}\right), \\ \sigma_{13}(\boldsymbol{p}) &= \left(p_{XY}^{-1} - p_X^{-1}\right)\lambda^2(1-\lambda)\left(\frac{1}{(1+\lambda)(1-\lambda)+\lambda+\lambda(1-\lambda)}\right. \\ &\qquad \left. + \frac{1}{1+\lambda(1-\lambda)}\left[\frac{1-\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+\lambda}\right]\right), \\ \sigma_{23}(\boldsymbol{p}) &= \left(p_{XY}^{-1} - p_Y^{-1}\right)\lambda(1-\lambda)^2\left(\frac{1}{\lambda(1+1-\lambda)+1-\lambda+\lambda(1-\lambda)}\right. \\ &\qquad \left. + \frac{1}{1+\lambda(1-\lambda)}\left[\frac{\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+1-\lambda}\right]\right),\end{aligned}$$

and $\sigma_{12}(\boldsymbol{p}) = 0$.

Proof is postponed in Section 3.1.

1.3 Asymptotic law and variances' closed expression in complete data framework

In the complete data framework we get $p_X = p_Y = p_{XY} = 1$, (*i.e.* when $\mathbf{p} = \mathbf{1}$) and the hybrid copula estimator becomes the empirical copula process. We know (see Section B.1) that the limit Gaussian process $\mathbb{C}_T$ of the normalized error of the empirical copula process is given by

$$N_C(u,v) = B_C(u,v) - \frac{\partial C}{\partial u}(u,v)B_C(u,1) - \frac{\partial C}{\partial u}(u,v)B_C(1,v),$$

where B_C is a Brownian bridge in $[0,1]^2$ with covariance function

$$\mathbb{E}[B_C(u,v)B_C(u',v')] = C(u \wedge u', v \wedge v') - C(u,v)C(u',v').$$

In this setting, we defined the corrected estimator of the λ-FMadogram such as

$$\begin{aligned}\hat{\nu}_T^*(\lambda) = &\frac{1}{2T\sum_{t=1}^T} \left| \hat{F}_T(X_t)^\lambda - \hat{G}_T(Y_t)^{1-\lambda} \right| - \frac{\lambda}{2T}\sum_{t=1}^T \{1 - \hat{F}_T(X_t)^\lambda\} \\ &- \frac{1-\lambda}{2T}\sum_{t=1}^T \{1 - \hat{G}_T(Y_t)^{1-\lambda}\} + \frac{1-\lambda+\lambda^2}{2(2-\lambda)(1+\lambda)}.\end{aligned}$$

We show that the asymptotic behaviour of the processes in Equation (1.3) is the same. We refer the reader to Section 3.2 for details. We also find a functional central limit theorem in this setting as show by the following statement.

Proposition 3 (**Asymptotic behavior theorem with complete data**). *Under Assumption A and complete data framework we have the weak convergence in* $\ell^\infty([0,1])$ *for* $\sqrt{T}(\hat{\nu}_T(\lambda) - \nu(\lambda))$, *as* $T \to \infty$,

$$\sqrt{T}\left(\hat{\nu}_T(\lambda) - \nu(\lambda)\right) \rightsquigarrow \left(- \int_{[0,1]} N_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})du \right)_{\lambda \in [0,1]}.$$

This result comes down from the proof of Theorem 1 and the fact that, in this setup, we can show that almost surely

$$B_C(u,1) - \frac{\partial C}{\partial u}(u,v)B_C(u,1) = 0, \quad \forall u \in [0,1],$$
$$B_C(1,v) - \frac{\partial C}{\partial v}(u,v)B_C(1,v) = 0, \quad \forall v \in [0,1].$$

We retrieve the asymptotic limit in law of the normalized estimation error of the λ-FMadogram as studied in [Marcon et al., 2017] (up to a given parametrization). Furthermore, as the integral of a tight Gaussian process, we know that for a fixed $\lambda \in [0,1]$, the asymptotic law is a Gaussian random variable.

Proposition 4 (**Asymptotic variance closed formula**). *Let $\lambda \in [0,1]$. Under Assumptions A and complete data framework we have the convergence in distribution for $\sqrt{T}(\hat{\nu}_T(\lambda) - \nu(\lambda))$, as $T \to \infty$,*

$$\sqrt{T}\left(\hat{\nu}_T(\lambda) - \nu(\lambda)\right) \xrightarrow{d} \mathcal{N}(0, \sigma_3^2(\boldsymbol{1})),$$

where $\sigma_3^2(\boldsymbol{1})$ is defined in Equation (3.8) in Section 3.1.

Remark 3. *For a fixed $\lambda \in [0,1]$, [Naveau et al., 2009] proved that the asymptotic law of $\sqrt{T}(\hat{\nu}(\lambda) - \nu(\lambda))$ can be written as*

$$\int_{[0,1]^2} f(x,y)dJ(x,y) = \frac{1}{2}\int_{[0,1]} f(1, y^{\frac{1}{(1-\lambda)}})dy + \frac{1}{2}\int_{[0,1]} f(x^{\frac{1}{\lambda}}, 1)dx - \int_{[0,1]} f(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx, \tag{1.6}$$

for every measurable and bounded function $f : [0,1]^2 \to \mathbb{R}$. Some details explaining Equation (1.6) are given in Section C.1 in Appendix C. The special case $J(x,y) = 2^{-1}|x^\lambda - y^{1-\lambda}|$ satisfies the conditions, then some computations gives that :

$$\int_{[0,1]^2} N_C(u,v)dJ(u,v) = -\int_{[0,1]} N_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})du.$$

We hence retrieve the result of Proposition 4 where we give a closed expression of the variance.

We are able to infer the closed form without integral of the Pickands of the λ-Madogram's variance in the case of an independent copula, *i.e.* when $C(u,v) = uv$.

Corollary 2 (**Asymptotic variance closed formula in the independent copula case**). *Let λ be fixed in $[0,1]$. Under Assumption A and if $C(u,v) = uv$, then* *the asymptotic variance of* $\sqrt{T}(\hat{\nu}_T(\lambda) - \nu(\lambda))$ *has the following form*

$$\sigma_3^2(\boldsymbol{1}) = \left(\frac{\lambda(1-\lambda)}{1+\lambda(1-\lambda)}\right)^2 \left(\frac{1}{1+2\lambda(1-\lambda)} \right.$$
$$\left. - \frac{1-\lambda}{2-(1-\lambda)+2\lambda(1-\lambda)} - \frac{\lambda}{2-\lambda+2\lambda(1-\lambda)}\right).$$

1.4 FMadogram with outliers and complete data

In order to propose a robust estimator we assume that the sample is partitioned into K disjoint subsets $B_1, \dots, B_K$ of cardinalities $T_j := card(B_j)$ respectively, where the partitioning scheme is independent of the data. Let f be a measurable function from $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ to $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$, we define the following estimator of $\mathbb{E}[f(X,Y)]$ by

$$\mathbb{P}_{T_j} f = \frac{1}{T_j} \sum_{j \in B_j} f(X_j, Y_j).$$

The MoN estimator defines in [Nemirovsky and Yudin, 1983] is given by

$$\hat{f}_{MoN} \in \underset{z \in \mathbb{R}}{argmin} \sum_{j=1}^{K} \left|\mathbb{P}_{T_j} f - z\right|. \tag{1.7}$$

Note that the usual univariate median

$$med(\mathbb{P}_{T_1} f, \dots, \mathbb{P}_{T_K} f),$$

is a solution of Equation (1.7). We restrict our analysis to the FMadogram in Equation (6) to avoid technical difficulties, but the proof would be similar with a discussion according to the value of the concentration bound and the value λ using that $||x|^\lambda - |y|^\lambda| \leq |x-y|^\lambda$ in the proof.

Intuitively, we replace the linear operator of expectation with the median of averages taken over non-overlapping blocks of the data, in order to get a robust estimate thanks to the median step (see [Lerasle et al., 2019] for a similar idea applied to Kernel). The MoN is one of the mean estimators that achieve a sub-Gaussian behavior under mild conditions. Introduced during the 1980s [Nemirovsky and Yudin, 1983] for the estimation of the mean of real-valued random variables, that is easy to compute, while exhibiting attractive robustness properties by the median step.

Definition 4. *Let $B_1, \dots, B_K$ a partition of the set $\{1, \dots, T\}$. Denote by $\hat{F}_{T_j}$ (resp. $\hat{G}_{T_j}$) the empirical cumulative distribution fonction of X (resp. Y) computed within block B_j. The MoN-based FMadogram estimator is defined by*

$$\hat{\nu}_{MoN} \in \underset{z \in \mathbb{R}}{argmin} \sum_{j=1}^{K} \left| \hat{\nu}_{T_j} - z \right| ,$$

where $\hat{\nu}_{T_j} = \frac{1}{2T_j} \sum_{t \in B_j} |\hat{F}_{T_j}(X_t) - \hat{G}_{T_j}(Y_t)|$.

That is, in Equation (1.7), we take $f(x, y) = |x - y|$ and $\mathbb{P}_{T_j} = C_{T_j}$ the empirical copula constructed on the block B_j.

Assumption D. *The* i.i.d. *sample $((X_1, Y_1), \dots, (X_T, Y_T))$ contains $T - T_o$ observations drawn according to distribution H, and T_o outliers upon which no assumption is made.*

In the presence of outliers, the key point is to focus on sane blocks, *i.e* on blocks that does not contain a single outlier, since no inference can be made about blocks hit by an outlier. One way to ensure that sane blocks are in majority is to assume that we have twice more blocks than outliers. Indeed, in the worst case scenario, each outlier contaminates one block, but the sane blocks remain more numerous. Let K_s denote the total number of sane block containing no outliers. In other words, we suppose that there exists $\delta \in (0, 1/2]$ such that $K_s \geq K(1/2 + \delta)$. If the data are free from contaminations, then $K_s = K$ and $\delta = 1/2$. This is a problematic assumption in practice since we do not know the number of outliers. Thus this lead to a compromise between a large number of blocks (allowing to deal with a large number of outliers) and a sufficient number of data in each block (allowing to make reasonable estimations).

We present a concentration inequality that the MoN-based estimator of the FMadogram does verify. Suppose without loss of generality that $T_j = \lceil T/K \rceil$ for every $j \in \{1, \dots, K\}$. We can prove the following deviation bounds for our MoN-based estimator.

Theorem 2 (Consistency and outlier-robustness). *Under Assumption D, for* *any* $\eta \in]0,1[$ *such that* $K = \delta^{-1} log(1/\eta)$ *it holds that with probability* $1-\eta$,

$$|\hat{\nu}_{MoN} - \nu| \leq \frac{3}{\sqrt{2}} \frac{log\left(6e2^{\frac{1}{\delta}}\right)}{\delta} \sqrt{\frac{log\,(1/\eta)}{T}}.$$

Details of the proof are available in Section 3.3 in Chapter 3.

Let us discuss parameters of the concentration bound. Dependence on T guarantees that the MoN-based estimator is robust to outliers, providing consistient estimates with high probability even under arbitrary contamination (affecting less than half of the samples). A higher δ corresponds to less outliers in which ase the bounds above become tighter. A lower η gives a greater bound for which the inequality holds with a greater probability.

Chapter 2

Empirical study

2.1 Presentation of the different parametric models

We present several models that will be used in the simulation section in order to analyze the performance of our estimator in finite-sample settings.

1. The asymmetric logistic model [Tawn, 1988] defined by the following dependence function:

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

   with parameters $\theta \in [1, \infty[$, ψ_1, $\psi_2 \in [0,1]$. The special case $\psi_1 = \psi_2 = 1$ gives us the symmetric Gumbel model. In the Gumbel model, we retrieve the independent case when $\theta = 1$, the dependence between the two variables is stronger as θ goes to infinity.

2. The asymmetric negative logistic model [Joe, 1990], namely,

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

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

3. The asymmetric mixel model [Tawn, 1988] :

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

with parameters θ and κ 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 get the independent copula.

4. The model of Hüsler and Reiss [Hüsler and Reiss, 1989],

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

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

5. The t-EV model [Demarta and McNeil, 2005], in which

$$A(w) = wt_{\nu+1}(z_w) + (1-w)t_{\nu+1}(z_{1-w}),$$
$$\text{with } z_w = (1+\nu)^{1/2}[w/(1-w)^{\frac{1}{\nu}} - \theta](1-\theta^2)^{-1/2},$$

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.

2.2 Complete data framework

2.2.1 Extreme value copulas

A Monte Carlo study is implemented here to illustrate Theorem 1 of Chapter 1 in finite-sample settings with no missing data. In Figure 1, for each $\lambda \in [0, 1]$, 500 random samples of size $T = 256$ were generated from the Gumbel copula (see Model 1 in Section 2.1) with $\theta \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$. For each sample, the λ-FMadogram estimators were computed with unknow margins. For each estimator we estimate the empirical version of the normalized estimation error variance, namely

$$\mathcal{E}_T(\lambda) := \widehat{Var}\left(\sqrt{T}\left(\hat{\nu}_T^*(\lambda) - \nu(\lambda)\right)\right), \tag{2.1}$$

Then we compare $\mathcal{E}_T(\lambda)$ with its associated theoretical asymptotic variance using the form exhibits in Proposition 4 (see Equation 3.8 in Section 3.1). In all figures, y-axis denote the estimate and theoretical value of Equation (2.1) and is labelled as σ^2.

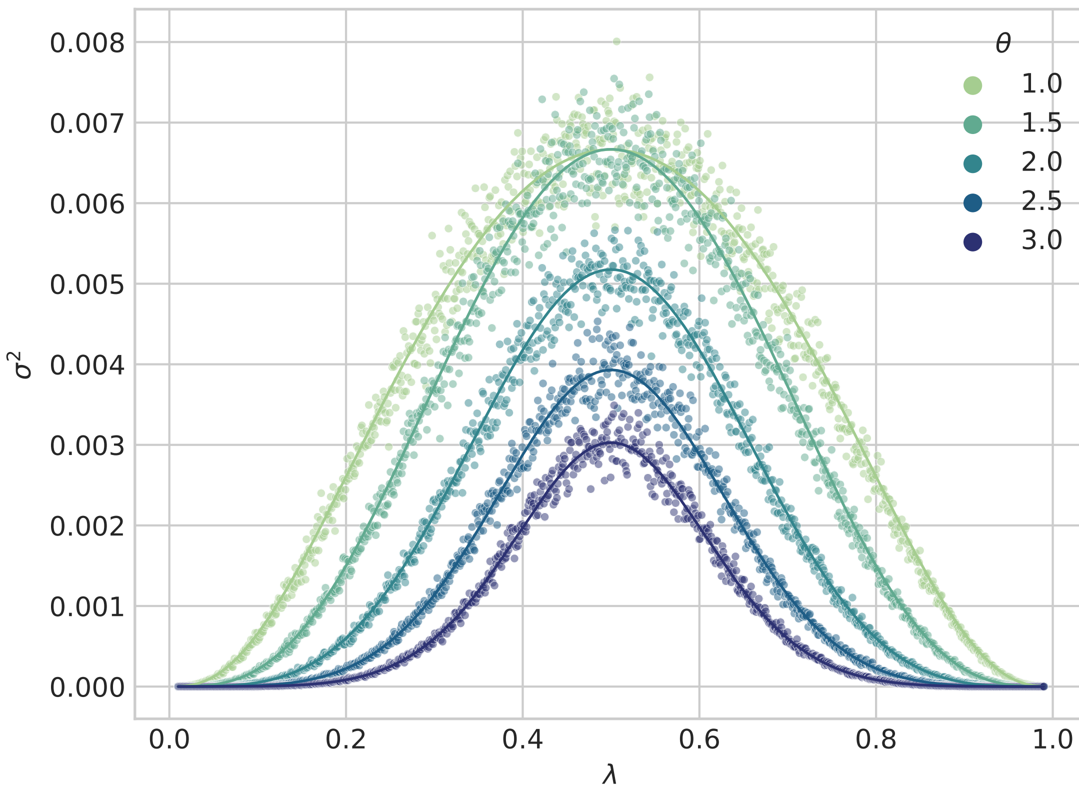


Figure 1: $\mathcal{E}_T(\lambda)$ in Equation (2.1), as a function of λ, based on 500 samples of size $T = 256$ from the Gumbel copula with $\theta = \{1.0, 1.5, 2.0, 2.5, 3.0\}$ chosen in such a way that $\lambda \in \{i/1000 : i = 10, \dots, 990\}$. Solid lines are obtained using Proposition 4.

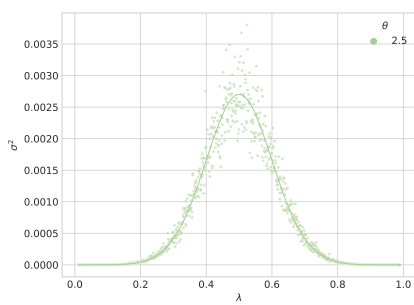
Similar numerical results were obtained with Figure 2 where Equation (2.1) and its theoretical value are drawn for many other extreme-value dependence models. We can notice the following :

Remark 4.

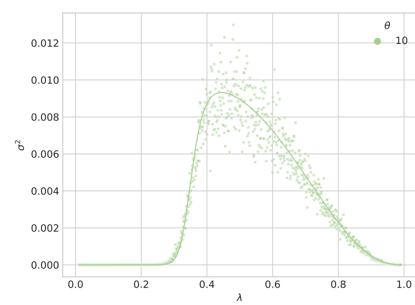
1. *When A is symmetric, one would expect the asymptotic variance of the estimator to reach its maximum at $\lambda = 1/2$. However, we can see in Figure 2f that this is not the case the t-EV model.*

2. *In the asymmetric negative logistic model (see Model 2 in Section 2.1), the asymptotic of the λ-FMadogram is close to zero for all $\lambda \in [0, 0.3]$. This is due to the fact that $A(\lambda) \approx 1 - t$ for this model.*

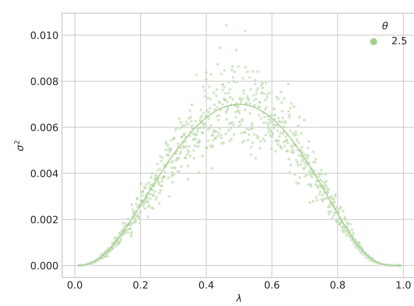
The two previous points are also observed in [Genest and Segers, 2009]. We propose in Figure 3 the theoretical asymptotic variance depending of θ and λ for the same models. The parameters of Figure 2 and Figure 3 are chosen accordingly to [Genest and Segers, 2009].



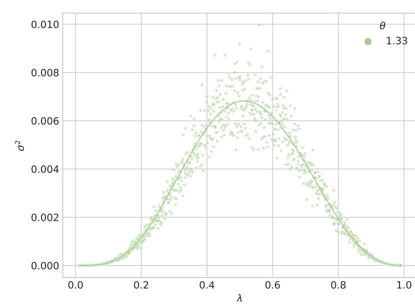
(a) **Galambos ($\theta = 2.5$)**



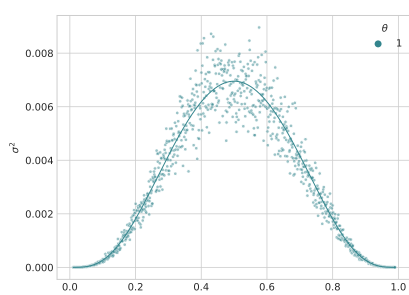
(b) **A. Neg. Log. ($\theta = 10$, $\psi_1 = 0.5$, $\psi_2 = 1.0$)**



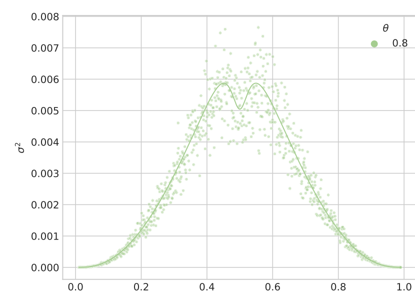
(c) **Asy. Log. ($\theta = \frac{5}{2}$, $\psi_1 = 0.1$, $\psi_2 = 1.0$)**



(d) **Asy. Mixed ($\theta = \frac{4}{3}$, $\kappa = -\frac{1}{3}$)**

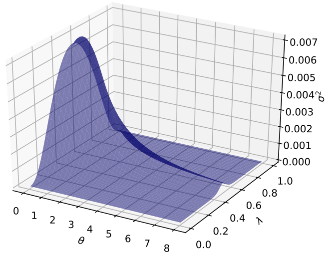


(e) **Hüsler-Reiss ($\theta = 1$)**

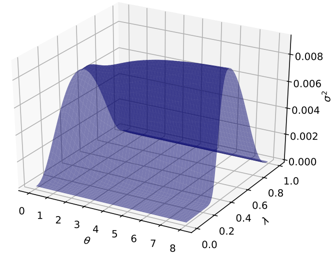


(f) **t-EV ($\theta = 0.8$, $\nu = 0.2$)**

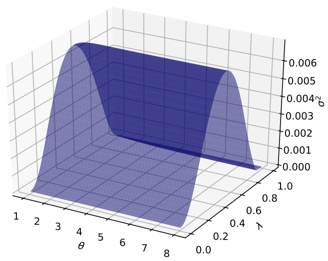
Figure 2: $\mathcal{E}_T(\lambda)$ in Equation (2.1), as a function of λ, based on 500 samples of size $T = 256$ of the λ-FMadogram. Solid lines are obtained using Proposition 4.



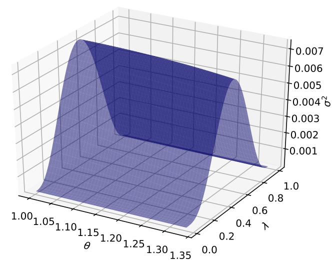
(a) **Galambos**



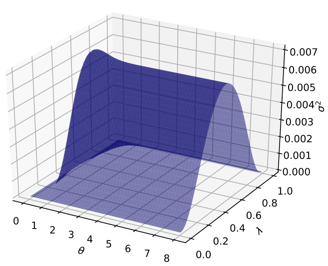
(b) **Asy. Neg. Log. ($\psi_1 = 0.5$, $\psi_2 = 1.0$)**



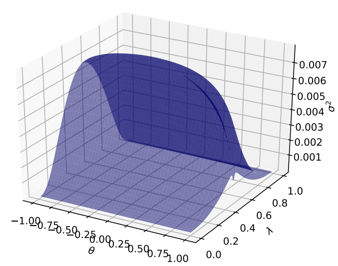
(c) **Asy. log. ($\psi_1 = 0.1, \psi_2 = 1.0$)**



(d) **Asy. Mixed ($\kappa = -\frac{1}{3}$)**



(e) **Hüsler-Reiss**



(f) **t-EV ($\nu = 0.2$)**

Figure 3: Asymptotic variance in Equation (3.8) in Section 3.1 depending on λ and the parameter θ of the chosen Pickands dependence function.

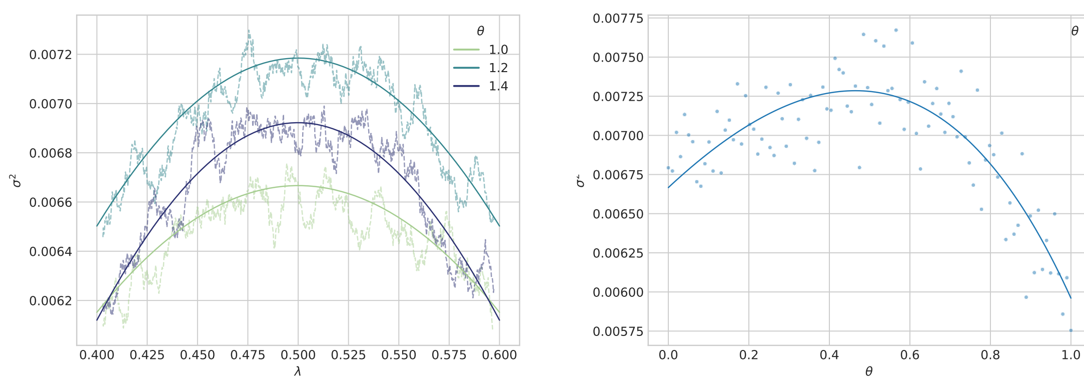
2.2.2 Non-monotonicity of the variance with respect to the dependence parameter

Looking at Figure 1, one may make another remark :

Remark 5. *There is no strict dominance between the asymptotic variance of the normalized estimator and the parametric extreme value copula. In other words, let A be a Pickands dependence function then we have the following*

$$1 \geq A(t), \quad \forall t \in [0,1] \not\Rightarrow \sigma^2_{3_\Pi} \geq \sigma^2_{3_A}, \quad \forall \lambda \in [0,1],$$

with $\sigma^2_{3_\Pi} := \sigma^2_3(\boldsymbol{1})$ in Proposition 4 (resp. $\sigma^2_{3_A}$) denotes the theoretical asymptotic variance of the normalized estimation error when $A(t) = 1$ (resp. for $1 \geq A(t)$), $\forall t \in [0,1]$.



(a) (b)

Figure 4: Panel (a) depicts $\mathcal{E}_T(\lambda)$ based on 500 samples of size $T = 256$ from the Gumbel copula with $\theta = \{1.0, 1.2, 1.4\}$ chosen in such a way that $\lambda \in \{i/1000 : i = 400, \dots, 600\}$. The dotted lines are moving averages made on the 1000 empirical estimates of the variance with a rolling window of 10. Panel (b) shows $\mathcal{E}_T(\lambda)$ based on 2000 sample of size $T = 512$ from the symmetric mixed model with $\lambda = 0.5$ chosen in such a way that $\theta \in \{i/100 : i = 0, \dots 100\}$. Solid lines are the theoretical asymptotic variance computed using Proposition 4, Equation (3.8)

To illustrate this phenomenon, we show in Figure 4a shows the same model with different values of θ and with a reduced scale for λ. The moving average is computed with a rolling window of 10 empirical variances for each θ. As the dependency parameter θ increases, we can find some λ for which the asymptotic variance is greater than the asymptotic variance in the case of independence. This figure supports our counterexample (see Section C.2 for details) that draws the same conclusion. Also,

Figure 4b depicts the asymptotic variance for a fixed $\lambda = 0.5$ for the symmetric mixed model (see Model 3 in Section 2.1) with $\theta \in [0, 1]$. When $\theta = 0$, we are turning back to the independent copula and its asymptotic variance is given by $1/150$ for this value of λ. When the random variables X and Y are becoming positively dependent, *i.e.* when θ increase in this model, the asymptotic variance for this given λ increase also, but after a certain threshold which depends on the chosen model, the variance starts to decrease.

2.2.3 Max-Stable processes

To determine the quality of the λ-FMadogram for estimating the pairwise dependence of maxima in space, [Naveau et al., 2009] computes on a particular class of simulated max-stable random fields. They focus on the Smith's max-stable process ([Smith, 2005]). Let us recall the bivariate distribution for the max-stable process model proposed by Smith is equal to :

$$\mathbb{P}\left(X(s) \leq u, X(s+h) \leq v\right) = exp\left[-\frac{1}{u}\Phi\left(\frac{a}{2} + \frac{1}{a}log\left(\frac{v}{u}\right)\right) - \frac{1}{v}\Phi\left(\frac{a}{2} + \frac{1}{a}log\left(\frac{v}{u}\right)\right)\right], \tag{2.2}$$

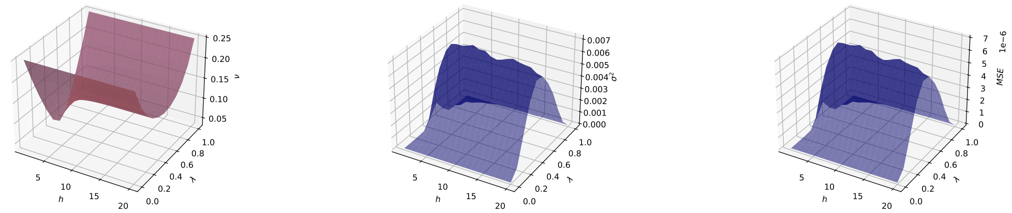
with $a^2 = (h^\top \Sigma^{-1} h)$ and Σ a covariance matrix. In case of isotropic field, we set $\Sigma = \sigma^2 I_2$. For this kind of process, the pairwise extremal dependence function $V(\cdot, \cdot)$ (see Section 4.3 of [Coles et al., 1999] for a definition) is given by :

$$V(u,v) = \frac{1}{u}\Phi\left(\frac{a}{2} + \frac{1}{a}log\left(\frac{v}{u}\right)\right) + \frac{1}{v}\Phi\left(\frac{a}{2} + \frac{1}{a}log\left(\frac{u}{v}\right)\right).$$

Furthermore, for a max-stable process, the theoretical value of the λ-FMadogram is given by

$$\nu(h, \lambda) = \frac{V(\lambda, 1-\lambda)}{1 + V(\lambda, 1-\lambda)} - \frac{3}{2(1+\lambda)(1+1-\lambda)}, \tag{2.3}$$

for any $\lambda \in [0, 1]$ (see Proposition 1 of [Naveau et al., 2009]). The λ-FMadogram was estimated independently for each simulated field with $T = 1024$. The xy-space $[0, 20] \times [0, 1]$, represent the distance h and parameter λ. In Smith's model, the pairwise dependence function between two locations s and $s+h$ decrease as the distance h between these two points increases. The red surface in Figure 5a provides the true value of the λ-FMadogram, the mean value of the 300 estimated λ-FMadogram in blue is juxtapose on it. Equation 2.1 is depicted on Figure 5b, its form is closed to the one that we have with the Hüsler-Reiss model in Figure 3e due to the similar expression of Equation 2.2.3 and the Pickands of Model 4 in Section 2.1. Figure



(a) **Estimated value** (b) **variance** (c) **Mean Squared Error**

Figure 5: Simulation results obtained by generating 300 independently and identically distributed Smith random fields. The dependence structure is characterized by (2.2) with $\Sigma = 25I_2$. Panel (5a) shows the estimated and the true λ-FMadogram. Panel (5b) represents $\mathcal{E}_T(\lambda)$ in Equation (2.1). Panel (5c) depicts the mean squared error between the true and estimated λ-FMadogram for all h and λ.

5c indicates the mean squared error between the estimated λ-FMadogram and the true one. As expected, the error is close to zero at the two boundary planes $\lambda = 0$ and $\lambda = 1$, by construction of the estimator. The largest mean squared errors are obtained where $\lambda = 0.5$, especially for very small distances, *i.e.* near $h = 0$. This behaviour is now well known from Section 2.2.2.

2.2.4 Block maxima model

In this section, we derive the behavior of the asymptotic variance of componentwise maxima of i.i.d random vectors having a t copula distribution. A bivariate t copula is defined as :

$$C_{\nu,\theta}(u,v) = \int_{]-\infty, t_\nu^{\leftarrow}(u)]} \int_{]-\infty, t_\nu^{\leftarrow}(v)]} \frac{1}{2\pi(1-\theta^2)^{1/2}} \left(1 + \frac{x^2 - 2\theta xy + y^2}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2} dydx,$$

where $\nu > 0$ is the number of degrees of freedom, $\theta \in [-1, 1]$ is the linear correlation coefficient, t_ν is the distribution function of a t-distribution with ν degrees of freedom. According to [Demarta and McNeil, 2005] the bivariate t copula $C_{\nu,\theta}$ is attracted to the t extreme value copula. Hence, we simulate $X_{1j}, \dots, X_{Mj}$, $j \in \{1, \dots, T\}$, a block of M variables from a t copula and we take the maximum in this block. This step is repeated several times in order to form a sample $(\bigvee_{i=1}^M X_{i1}, \dots, \bigvee_{i=1}^M X_{iT})$ of length T. The result depicts on Figure 6 is what we waited for. As we expected, when the number of observations in block maxima, the empirical variance converge towards the asymptotic variance for a t-EV copula. Furthermore, as the sample's length increases, better the empirical variance fits the theoretical one.

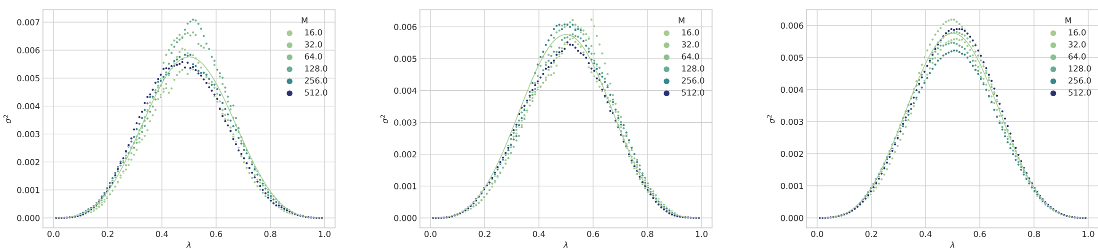


Figure 6: Simulation results obtained by generating $T \in \{128, 256, 512\}$ blocks maxima's of length $M \in \{16, 32, 64, 128, 256, 512\}$ from t-copula with parameters $\theta = 0.8$ and $\nu = 3$. Each T is associated, in increasing order, to the left, middle and right panel. For each $\lambda \in \{i/100, i = 1, \dots, 99\}$, we estimate $\mathcal{E}_T(\lambda)$ in Equation (2.1) on 100 estimator of λ-FMadogram. The solid line is the theoretical asymptotic variance of t-EV copula with $\theta = 0.8$ and $\nu = 3$ given by Proposition 4.

2.3 Missing data framework

In this section, we estimate the empirical variance on several Monte Carlo simulations of the normalized estimation error for both hybrid estimator. These errors are define respectively for the hybrid and the corrected estimators (see Equation (1.1) and (1.2)) by

$$\mathcal{E}_T^{\mathcal{H}}(\lambda) := \widehat{Var}\left(\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}}(\lambda) - \nu(\lambda)\right)\right), \tag{2.4}$$

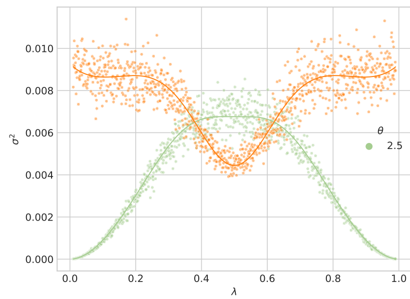
$$\mathcal{E}_T^{\mathcal{H}*}(\lambda) := \widehat{Var}\left(\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right)\right). \tag{2.5}$$

For $p_X \in]0,1]$ and $p_Y \in]0,1]$ the indicator of missing of variables X and Y is generated according to a Bernoulli distribution

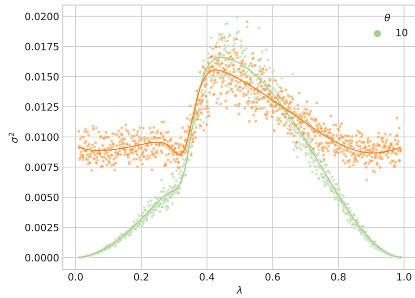
$$I \sim \mathcal{B}(p_X), \quad J \sim \mathcal{B}(p_Y).$$

We also set that $p_{XY} = p_X p_Y$, *i.e.* I and J are independent.

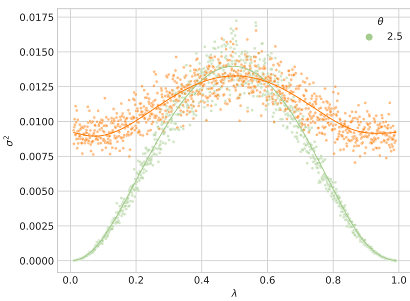
Figure 7 presents the results for the six models used in Section 2.2.1. For each model, the green (*resp.* red) line presents the asymptotic variance $\mathcal{E}_T^{\mathcal{H}}(\lambda)$ (*resp.* $\mathcal{E}_T^{\mathcal{H}*}(\lambda)$) for $\hat{\nu}_T^{\mathcal{H}}$ (*resp.* $\hat{\nu}_T^{\mathcal{H}*}$) given by Corollary 2. For each $\lambda \in \{\frac{i}{1000}, i = \frac{10}{1000}, \dots \frac{990}{1000}\}$, we estimate its empirical counterpart. Here, we take $p_X = p_Y = 0.75$. As waited, we directly see that both empirical and theoretical values of the variance of the normalized error of $\hat{\nu}_T^{\mathcal{H}}$ is different from zero for each extremity of λ. Furthermore, in some models, we also lose the "parabolic" shape of the curve (see Figure (7a)). The introduction of the corrected estimator may us to recover the same pattern as noticed in Figure 2 of Section 2.2.1. Notice that, in terms of variance, we do not have a strict dominance from one estimator to another as it was mentioned before.



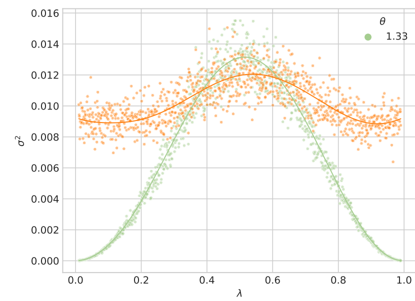
(a) **Galambos** ($\theta = 2.5$)



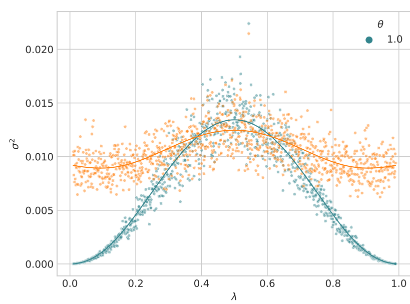
(b) **Asym. Neg. Log.** ($\theta = 10$, $\psi_1 = 0.5$, $\psi_2 = 1.0$)



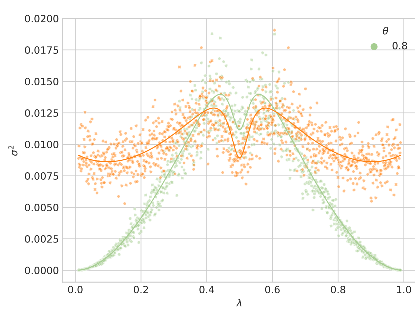
(c) **Asy. Log.** ($\theta = \frac{5}{2}$, $\psi_1 = 0.1$, $\psi_2 = 1.0$)



(d) **Asy. Mixed** ($\theta = \frac{4}{3}$, $\kappa = -\frac{1}{3}$)



(e) **Husler-Reiss** ($\theta = 1.0$)



(f) **t-EV** ($\theta = 0.8$, $\nu = 0.2$)

Figure 7: Equation (2.4) (in red) and Equation (2.5) (in green), as a function of λ, of the asymptotic variances of the estimators of the λ-FMadogram for six extreme-value copula models. The empirical variances are based on 500 samples of size $T = 256$. Solides lines are the theoretical value given by Proposition 2.

2.4 Robustness and outliers with complete data

In this section, we illustrate the results obtained in Section 1.4 concerning the robustness of the MoN-based estimator (see Equation 4). To do so, we set up two different types of outliers :

- *top-left*: the outliers are drawn *i.i.d* from a uniform distribution $\mathcal{U}([0, 0.05] \times [0.95, 1])$.

- *bottom-right*: the outliers are drawn *i.i.d* from a uniform distribution $\mathcal{U}([0.95, 1] \times [0, 0.05])$.

In each case, sane data are sampled from the desired copula model. Then all data are inverted by the quantile function of a standard Gaussian distribution. Moreover, we consider two models of contamination :

- *Huber* : $(X_t, Y_t)_{t=1}^T$ is drawn *i.i.d.* from the mixture $\tilde{\mathbb{P}} = (1-\epsilon)H + \epsilon\mathbb{P}_o$ where $\epsilon \in (0,1)$ and $\mathbb{P}_o$ is an arbitrary distribution.

- *Adversarial* : T *i.i.d.* copies are sample from H. Then an adversary is allowed to look at the samples and arbitrarily corrupt an ϵ-fraction of them. In this setup, outliers can be correlated to sane data.

In the adversarial contamination, our rule is to sample again points which are closer to the point $(0.5, 0.5)$ for the Gumbel model and the Asymmetric Logistic (see Model 1). For the Negative Asymmetric Logistic (see 2), we sample again points which are closer to $(1.0, 1.0)$. As the adversarial contamination breaks out the indepence of the sample in sane blocks, we are out of the framework of Theorem 2.

Figure 8 present an illustration of contaminated data for three selected copula models for all types of contaminations and outliers we consider. We report the squared bias of each estimator for all the models considered in Figure 9 for three extreme value copula, two types of outliers and contaminations. When there are no outliers, the MoN-based estimator and the FMadogram has the same bias, this is due that $\hat{\nu}_{MoN} = \hat{\nu}_T$ when $K = 1$. Otherwise, our MoN-based estimator is not always better than the FMadogram. This is the case when the Pickands dependence function is symmetric at $\lambda = 0.5$. When it is not, we can see that our MoN-based estimator become much more reliable when the Pickands is asymmetric. Indeed, when the asymmetry of the dependence model induces a concentration of the points on the bottom right, we see that the MoN-based estimator is better than the FMadogram for the *bottom-right* types of outliers. That is shown in Figure 9 for the asymmetric Negative Logistic Model for the second and the fourth rows. The same observation is drawn in the Asymmetric Logistic Model which performs comparatively for the Huber's contamination and better for the adversarial one. For these asymmetric models, the MoN-based estimator is more robust than the FMadogram because the *bottom-right* types of outliers (resp. *top-left*) upsets ranks of sane data for the asymmetric negative logistic model (resp. for the asymmetric logistic model).

(a) **Gumbel** (b) **Asy. Log.** (c) **Asy. Neg. Log.**

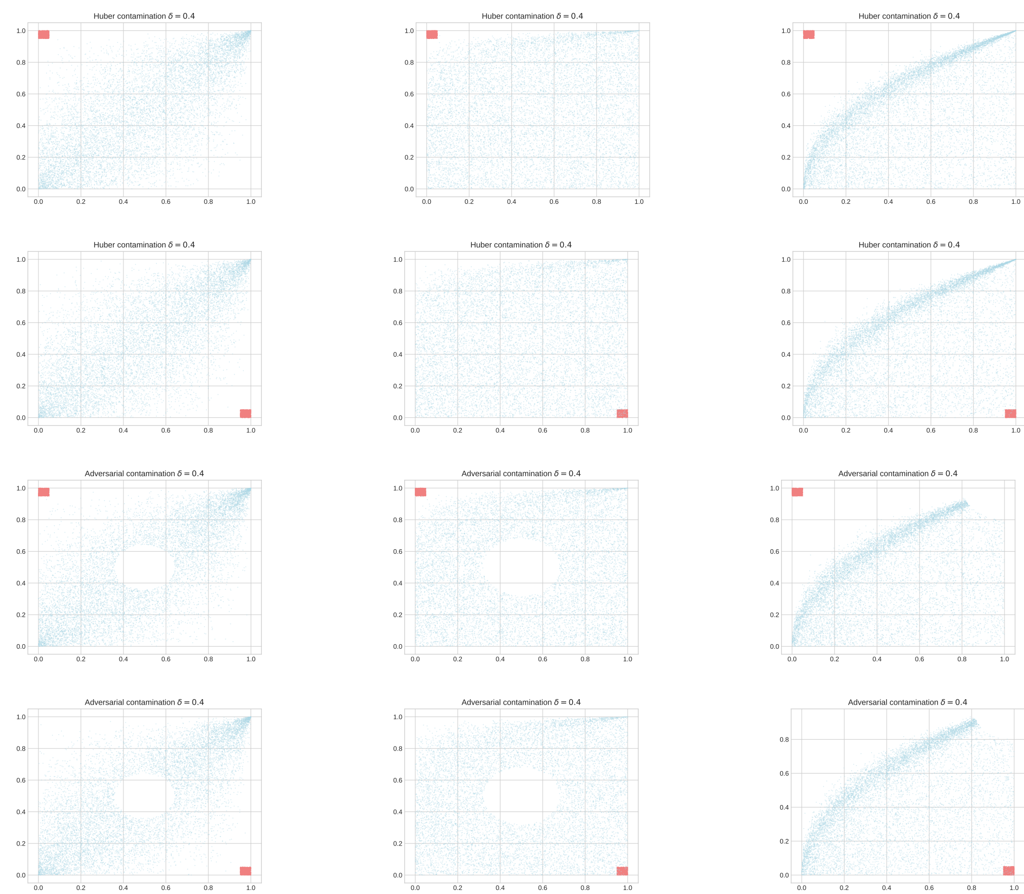


Figure 8: Sample of 10000 data with a fraction of 10% of outliers. Sane points are depicted in blue while contaminated ones are in red. The two first rows are respectively the *top-left* and *bottom-right* types of outliers for the Huber's contamination model while the two next are for the adversarial contamination. For Gumbel's model, we took $\theta = 1.5$, for the asymmetric logistic model, we consider $\theta = 2.5$, $\psi_1 = 0.1$, $\psi_2 = 1.0$ while the asymmetric negative logistic model is defined with $\theta = 10$, $\psi_1 = 0.5$, $\psi_2 = 1.0$.

(a) **Gumbel** (b) **Asy. Log.** (c) **Asy. Neg. Log.**

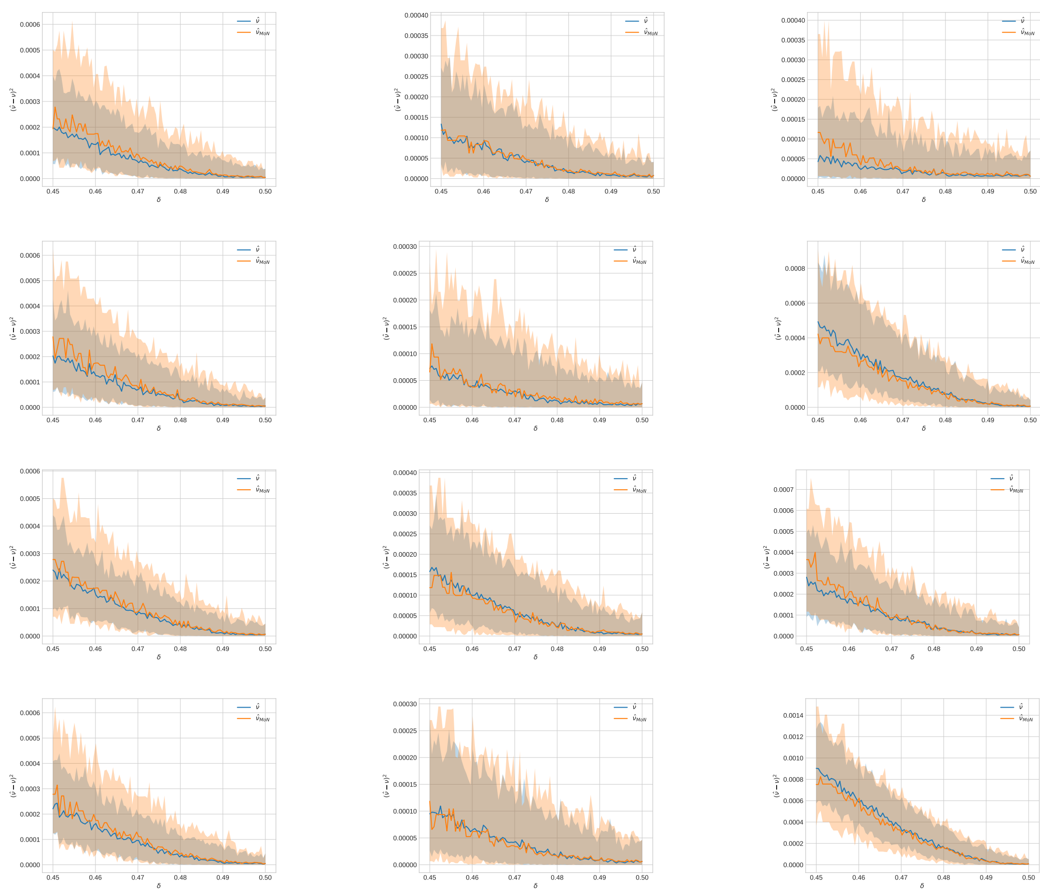


Figure 9: Squared bias of the MoN-based estimator (orange) and the FMadogram (blue) and the 90% confidence band computed on 100 estimators build on a sample of length $T = 1000$. The two first row is respectively the *top-left* and *bottom-right* types of outliers for the Huber's contamination model while the two next are for the adversarial contamination. For Gumbel's model, we took $\theta = 1.5$, for the asymmetric logistic model, we consider $\theta = 2.5$, $\psi_1 = 0.1$, $\psi_2 = 1.0$ while the asymmetric negative logistic model is defined with $\theta = 10$, $\psi_1 = 0.5$, $\psi_2 = 1.0$.

Chapter 3

Proofs

This chapter is devoted to prove of all the results of Chaper 1.

3.1 Proofs in the general case

Proposition 1 state the strong consistency of the hybrid and corrected estimator defined in Equation (1.1) and (1.2) (see page 9)

Proof of Proposition 1 We prove it for $\hat{\nu}_T^{\mathcal{H}}(\lambda)$ as the strong consistency for $\hat{\nu}_T^{\mathcal{H}*}(\lambda)$ use the same arguments. The estimator $\hat{\nu}_T(\lambda)$ is strongly consistent since it holds

$$\begin{aligned}
&\left|\frac{1}{2\sum_{t=1}^T I_tJ_t}\sum_{t=1}^T\left|\hat{F}_T(X_t)^\lambda-\hat{G}_T(Y_t)^{1-\lambda}\right|I_tJ_t-\frac{1}{2}\mathbb{E}\left|F(X)^\lambda-G(Y)^{1-\lambda}\right|\right|\\
&\leq\left|\frac{1}{2\sum_{t=1}^T I_tJ_t}\sum_{t=1}^T\left|\hat{F}_T(X_t)^\lambda-\hat{G}_T(Y_t)^{1-\lambda}\right|I_tJ_t-\frac{1}{2\sum_{t=1}^T I_tJ_t}\sum_{t=1}^T\left|F(X_t)^\lambda-G(Y_t)^{1-\lambda}\right|I_tJ_t\right|\\
&+\left|\frac{1}{2\sum_{t=1}^T I_tJ_t}\sum_{t=1}^T\left|F(X_t)^\lambda-G(Y_t)^{1-\lambda}\right|I_tJ_t-\frac{1}{2}\mathbb{E}\left|F(X)^\lambda-G(Y)^{1-\lambda}\right|\right|.
\end{aligned}$$

The second term converges almost surely to zero by the strong Law of Large Numbers and Assumption B. For the first term, we have

$$
\begin{aligned}
&\left|\frac{1}{2\sum_{t=1}^{T} I_t J_t}\sum_{t=1}^{T}\left|\hat{F}_T(X_t)^{\lambda}-\hat{G}_T(Y_t)^{1-\lambda}\right| I_t J_t-\frac{1}{2\sum_{t=1}^{T} I_t J_t}\sum_{t=1}^{T}\left|F(X_t)^{\lambda}-G(Y_t)^{1-\lambda}\right| I_t J_t\right|\\
&\leq \frac{1}{2\sum_{t=1}^{T} I_t J_t}\sum_{t=1}^{T}\left||\hat{F}_T(X_t)^{\lambda}-\hat{G}_T(Y_t)^{1-\lambda}|-\left|F(X_t)^{\lambda}-G(Y_t)^{1-\lambda}\right|\right| I_t J_t,\\
&\leq \frac{1}{2\sum_{t=1}^{T} I_t J_t}\sum_{t=1}^{T}\left|\hat{F}_T(X_t)^{\lambda}-F(X_t)^{\lambda}-\left(\hat{G}_T(Y_t)^{1-\lambda}-G(Y_t)^{1-\lambda}\right)\right| I_t J_t,\\
&\leq \frac{1}{2\sum_{t=1}^{T} I_t J_t}\sum_{t=1}^{T}\left|\hat{F}_T(X_t)^{\lambda}-F(X_t)^{\lambda}\right| I_t J_t+\left|\hat{G}_T(Y_t)^{1-\lambda}-G(Y_t)^{1-\lambda}\right| I_t J_t.\\
&\leq \frac{1}{2}\sup_{t\in\{1,\dots,T\}}\left|\hat{F}_T(X_t)^{\lambda}-F(X_t)^{\lambda}\right|+\frac{1}{2}\sup_{t\in\{1,\dots,T\}}\left|\hat{G}_T(Y_t)^{1-\lambda}-G(Y_t)^{1-\lambda}\right|
\end{aligned}
$$

which converges almost surely to zero, according to the Glivencko-Cantelli theorem.

□

Lemma 1 aims to assert that the hybrid margins $\hat{F}_T$, $\hat{G}_T$ and the hybrid cumulative distribution function $\hat{H}_T$ does verify Assumption C.

Proof of Lemma 1 Consider the following functions from $\{0,1\}^2 \times \mathbb{R}^2$ into $\mathbb{R}$: for $(x,y) \in \mathbb{R}^2$,

$$
\begin{aligned}
f_1(I,J,X,Y) = \mathbb{1}_{\{I=1\}}, &\quad g_{1,x} = \mathbb{1}_{\{X\leq x, I=1\}},\\
f_2(I,J,X,Y) = \mathbb{1}_{\{I=1\}}, &\quad g_{2,x} = \mathbb{1}_{\{X\leq x, I=1\}},\\
f_3 = f_1 f_2, &\quad g_{3,x,y} = g_{1,x} g_{2,y}.
\end{aligned}
$$

Let P denote the common distribution of the quadruples (I,J,X,Y). Consider the collection of functions

$$
\mathcal{F} = \{f_1, f_2, f_3\} \cup \{g_{1,x} : x \in \mathbb{R}\} \cup \{g_{2,y} : y \in \mathbb{R}\} \cup \{g_{3,x,y} : (x,y) \in \mathbb{R}^2\}.
$$

The empirical process $\mathbb{G}_T$ defined by

$$
\mathbb{G}_T(f) = \sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T} f(I_t, J_t, X_t, Y_t) - \mathbb{E}[f(I_t, J_t, X_t, Y_t)]\right), \quad f \in \mathcal{F},
$$

converge in $\ell^{\infty}(\mathcal{F})$ to a P-Brownian bridge $\mathbb{G}$ (see [Segers, 2014]). To establish such a statement, results on empirical processes based on the theory of Vapnik-Cervonenkis classes (VC-classes) of functions as formulated in [van der Vaart and Wellner, 1996] were used. We now add some lines of algebra to establish the weak convergence of the processes $\hat{F}_T(x)$, $\hat{G}_T(y)$ and $\hat{H}_T(x,y)$. These lines are made in the first process as the method is similar to the others. For $x \in \mathbb{R}$, we write

$$\hat{F}_T(x) = \frac{p_X F(x) + T^{-1/2}\mathbb{G}_T g_{1,x}}{p_X + T^{-1/2}\mathbb{G}_T f_1}.$$

We obtain :

$$\begin{aligned} p_X(\hat{F}_T(x) - F(x)) &= T^{-1/2}(\mathbb{G}_T(g_{1,x}) - \mathbb{G}_T(f_1)\hat{F}_T(x)), \\ &= T^{-1/2}(\mathbb{G}_T(g_{1,x} - F(x)f_1)) + T^{-1/2}\mathbb{G}_T(f_1)(F(x) - \hat{F}_T(x)). \end{aligned}$$

Multiplying by $\sqrt{T}$ and dividing by p_X gives :

$$\sqrt{T}(\hat{F}_T(x) - F(x)) = p_X^{-1}(\mathbb{G}_T(g_{1,x} - F(x)f_1)) + p_X^{-1}\mathbb{G}_T(f_1)(F(x) - \hat{F}_T(x)).$$

By the central limit theorem, we have that $\mathbb{G}_T(f_1) \xrightarrow{d} \mathcal{N}(0, \mathbb{P}(f_1 - \mathbb{P}f_1)^2)$, applying the law of the large number gives us that $(F(x) - \hat{F}_T(x)) = o_{\mathbb{P}}(1)$. With the help of Slutksy theorem, the second term in the right hand side is thus $o_{\mathbb{P}}(1)$, so

$$\sqrt{T}(\hat{F}_T(x) - F(x)) = p_X^{-1}(\mathbb{G}_T(g_{1,x} - F(x)f_1)) + o_{\mathbb{P}}(1).$$

As a consequence, we obtain the following limiting process of the Lemma :

$$\beta_1(u) = p_X^{-1}\mathbb{G}(1_{X \le F^{\leftarrow}(u), I=1} - u1_{I=1}).$$

We know that the covariance of a $\mathbb{P}$-Gaussian process is given by $\mathbb{E}[\mathbb{G}(f)\mathbb{G}(g)] = \mathbb{P}(fg) - \mathbb{P}(f)\mathbb{P}(g)$ where f, g are measurable functions. Now, using that, we have

$$\begin{aligned} cov[\beta_1(u_1), \beta_1(u_2)] &= p_X^{-2}\mathbb{E}\left[\mathbb{G}(1_{X \le F^{\leftarrow}(u_1), I=1} - u_1 1_{I=1})\mathbb{G}(1_{X \le F^{\leftarrow}(u_2), I=1} - u_2 1_{I=1})\right], \\ &= p_X^{-2}(\mathbb{P}\left[(1_{X \le F^{\leftarrow}(u_1), I=1} - u_1 1_{I=1})(1_{X \le F^{\leftarrow}(u_2), I=1} - u_2 1_{I=1})\right]), \\ &= p_X^{-2}(\mathbb{P}(I=1)\mathbb{P}(X \le F^{\leftarrow}(u_1), X \le F^{\leftarrow}(u_2)) - u_1 u_2 \mathbb{P}(I=1)), \\ &= p_X^{-1}(u_1 \wedge u_2 - u_1 u_2). \end{aligned}$$

□

We thus prove the functional central limit theorem of our processes defined in Equation (1.3).

Proof of Theorem 1 We do the proof only for the normalized error of $\hat{\nu}^{\mathcal{H}*}$ as the proof of $\hat{\nu}^{\mathcal{H}}$ is clearly similar. Using that $\mathbb{E}[F(X)^\alpha] = \frac{1}{1+\alpha}$ $(\alpha \neq 1)$, we can write $\nu(\lambda)$ as :

$$\begin{aligned}\nu(\lambda) &= \frac{1}{2}\mathbb{E}\left[|F(X)^\lambda - G(Y)^{1-\lambda}|\right] - \frac{\lambda}{2}\mathbb{E}\left[1 - F(X)^\lambda\right] - \frac{1-\lambda}{2}\mathbb{E}\left[1 - G(Y)^{1-\lambda}\right] \\ &+ \frac{1}{2}\frac{1-\lambda-\lambda^2}{(1+\lambda)(1+1-\lambda)}.\end{aligned}$$

Let us note, by g_λ the function defined as:

$$g_\lambda \colon [0,1]^2 \to [0,1], \quad (u,v) \mapsto u^\lambda \vee v^{1-\lambda} - \frac{1}{2}\left((1-\lambda)u^\lambda + \lambda v^{1-\lambda}\right).$$

We are able to write our estimator of the λ-FMadogram (resp. the λ-FMadogram) in missing data framework as an integral with respect to the hybrid copula estimator (resp. the copula function). We then have :

$$\begin{aligned}&\hat{\nu}_T^{\mathcal{H}*}(\lambda) = \frac{1}{\sum_{t=1}^T I_t J_t}\sum_{t=1}^T g_\lambda(\hat{F}_T(X_t), \hat{G}_T(Y_t)) I_T J_t + c_\lambda = \int_{[0,1]^2} g_\lambda(u,v) d\hat{C}_T^{\mathcal{H}}(u,v) + c_\lambda, \\ &\nu(\lambda) = \int_{[0,1]^2} g_\lambda(u,v) dC(u,v) + c_\lambda.\end{aligned}$$

Where c_λ a constant depending on λ. Using the same tools introduced to prove Lemma B.1, we are able to show that :

$$\begin{aligned}\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right) &= \frac{1}{2}\left((1-\lambda)\int_{[0,1]} \mathbb{C}_T^{\mathcal{H}}(x^{\frac{1}{\lambda}}, 1)dx + \lambda \int_{[0,1]} \mathbb{C}_T^{\mathcal{H}}(1, x^{\frac{1}{1-\lambda}})dx\right) \\ &- \int_{[0,1]} \mathbb{C}_T^{\mathcal{H}}(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx.\end{aligned}$$

Consider the function $\phi \colon \ell^\infty([0,1]^2) \to \ell^\infty([0,1]), f \mapsto \phi(f)$, defined by

$$(\phi(f))(\lambda) = \frac{1}{2}\left((1-\lambda)\int_{[0,1]} f(x^{\frac{1}{\lambda}}, 1)dx + \lambda \int_{[0,1]} f(1, x^{\frac{1}{1-\lambda}})dx\right) - \int_{[0,1]} f(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx.$$

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

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

in $\ell^\infty([0,1])$. We note that $S_C(u,1) = \alpha(u,1) - \beta_1(u)$ and $S_C(1,v) = \alpha(1,v) - \beta_2(v)$. Indeed, just remark that for the first one we have $\beta_2(1) = 0$ and $\partial C / \partial u(u,1) = 1$ a.s. We thus obtain our statement. $\square$

The following proof is really technical but gives the closed expression of the variance of the Gaussians limit law of our estimators defined in Equation (1.1) and (1.2).

Proof of Proposition 2 We are able to compute the variance for each process and they are given by the following expressions :

$$\sigma_1^2(p_X, p_{XY}) := Var\left(\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du\right),$$

$$\sigma_2^2(p_Y, p_{XY}) := Var\left(\int_{[0,1]} \alpha(1, u^{\frac{1}{1-\lambda}}) - \beta_2(u^{\frac{1}{1-\lambda}}) du\right),$$

$$\sigma_3^2(\mathbf{p}) := Var\left(\int_{[0,1]} S_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) du\right).$$

For $\sigma_1^2(p_X, p_{XY})$, we compute

$$\begin{aligned}
\sigma_1^2(p_X, p_{XY}) &= \mathbb{E}\left[\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du \int_{[0,1]} \alpha(v^{\frac{1}{\lambda}}, 1) - \beta_1(v^{\frac{1}{\lambda}}) dv\right], \\
&= \int_{[0,1]^2} \mathbb{E}\left[\alpha(u^{\frac{1}{\lambda}}, 1)\alpha(v^{\frac{1}{\lambda}}, 1) duv\right] - 2\int_{[0,1]^2} \mathbb{E}\left[\alpha(u^{\frac{1}{\lambda}}, 1)\beta_1(v^{\frac{1}{\lambda}}) duv\right] \\
&+ \int_{[0,1]^2} \mathbb{E}\left[\beta_1(u^{\frac{1}{\lambda}})\beta_1(v^{\frac{1}{\lambda}}) duv\right], \\
&= \left(p_{XY}^{-1} - p_X^{-1}\right) \int_{[0,1]^2} (u \wedge v)^{\frac{1}{\lambda}} - u^{\frac{1}{\lambda}} v^{\frac{1}{\lambda}} duv.
\end{aligned}$$

Using the same techniques, we have for $\sigma_2^2(p_Y, p_{XY})$

$$\sigma_2^2(p_Y, p_{XY}) := \left(p_{XY}^{-1} - p_Y^{-1}\right) \int_{[0,1]^2} (u \wedge v)^{\frac{1}{1-\lambda}} - u^{\frac{1}{1-\lambda}} v^{\frac{1}{1-\lambda}} duv.$$

We compute directly

$$\int_{[0,1]^2} (u \wedge v)^{\frac{1}{\lambda}} - u^{\frac{1}{\lambda}} v^{\frac{1}{\lambda}} duv = \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{1}{1+2\lambda},$$

and then

$$\sigma_1^2 = \left(p_{XY}^{-1} - p_X^{-1}\right) \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{1}{1+2\lambda}. \tag{3.1}$$

Similarly

$$\sigma_2^2 = \left(p_{XY}^{-1} - p_Y^{-1}\right) \left(\frac{1-\lambda}{1+1-\lambda}\right)^2 \frac{1}{1+2(1-\lambda)}. \tag{3.2}$$

For the covariances, we note

$$\begin{aligned}
\sigma_{12}(\mathbf{p}) &:= \mathbb{E}\left[\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du \int_{[0,1]} \alpha(1, v^{\frac{1}{1-\lambda}}) - \beta_2(v^{\frac{1}{1-\lambda}}) dv\right], \\
\sigma_{13}(\mathbf{p}) &:= \mathbb{E}\left[\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du \int_{[0,1]} S_C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) dv\right], \\
\sigma_{23}(\mathbf{p}) &:= \mathbb{E}\left[\int_{[0,1]} \alpha(1, u^{\frac{1}{1-\lambda}}) - \beta_2(u^{\frac{1}{1-\lambda}}) du \int_{[0,1]} S_C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) dv\right].
\end{aligned}$$

Some algebra gives for the first one

$$\sigma_{12}(\mathbf{p}) = \left(p_{XY}^{-1} - p_X^{-1} - p_Y^{-1} + \frac{p_{XY}}{p_X p_Y}\right) \int_{[0,1]^2} C(u^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) - u^{\frac{1}{\lambda}} v^{\frac{1}{1-\lambda}} duv. \tag{3.3}$$

For $\sigma_{13}(\mathbf{p})$, using the definition of the process $S_C(u,v)$ some lines gives

$$\begin{aligned}
\sigma_{13}(\mathbf{p}) = \mathbb{E}&\left[\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du \int_{[0,1]} \left(\alpha(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) dv - \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v} \beta_2(v^{\frac{1}{\lambda}})\right)\right], \\
&= \left(p_{XY}^{-1} - p_X^{-1}\right) \int_{[0,1]^2} C((u \wedge v)^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) - u^{\frac{1}{\lambda}} C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) duv \qquad (3.4) \\
&- \left(p_Y^{-1} - \frac{p_{XY}}{p_X p_Y}\right) \int_{[0,1]^2} \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v} \left(C(u^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) - u^{\frac{1}{\lambda}} v^{\frac{1}{1-\lambda}}\right) duv, \qquad (3.5)
\end{aligned}$$

where we use, in the first line that

$$\mathbb{E}\left[\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, 1) - \beta_1(u^{\frac{1}{\lambda}}) du \int_{[0,1]} \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} \beta_1(v^{\frac{1}{\lambda}}) dv\right] = 0.$$

Similarly,

$$\sigma_{23}(\mathbf{p}) = \left(p_{XY}^{-1} - p_Y^{-1}\right) \int_{[0,1]^2} C(v^{\frac{1}{\lambda}}, (u \wedge v)^{\frac{1}{1-\lambda}}) - u^{\frac{1}{1-\lambda}} C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) duv \tag{3.6}$$

$$- \left(p_X^{-1} - \frac{p_{XY}}{p_X p_Y}\right) \int_{[0,1]^2} \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} \left(C(v^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) - v^{\frac{1}{\lambda}} u^{\frac{1}{1-\lambda}}\right) duv. \tag{3.7}$$

We go back to the quantity σ_3^2. But before, we introduce some notations for convenience. Let $\lambda \in [0,1]$, using the property exhibited in Equation (5). We find a similar pattern for partial derivatives,

$$\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u} = \frac{C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{u^{\frac{1}{\lambda}}} \left(A(\lambda) - A'(\lambda)\lambda\right),$$
$$\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v} = \frac{C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{v^{\frac{1}{1-\lambda}}} \left(A(\lambda) + A'(\lambda)(1-\lambda)\right).$$

Furthermore, the integral $\int_{[0,1]^2} C(u,v)duv$ does not admit, in general, a closed form. But we are able to express it with respect to a simple integral of the Pickands dependence function. We note, for notational convenience the following functional

$$f \colon [0,1] \times \mathcal{A} \to [0,1], \quad (\lambda, A) \mapsto \left(\frac{\lambda(1-\lambda)}{A(\lambda) + \lambda(1-\lambda)}\right)^2.$$

Now, we write σ_3^2 in the following way

$$\sigma_3^2(\mathbf{p}) = \left(p_{XY}^{-1}\gamma_1^2 + p_X^{-1}\gamma_2^2 + p_Y^{-1}\gamma_3^2\right) - 2p_X^{-1}\gamma_{12} - 2p_Y^{-1}\gamma_{13} + 2\frac{p_{XY}}{p_X p_Y}\gamma_{23}, \tag{3.8}$$

with

$$
\begin{aligned}
&Var\left(\int_{[0,1]} \alpha(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})du\right) \\
&= p_{XY}^{-1} \int_{[0,1]^2} \left(C((u \wedge v)^{\frac{1}{\lambda}}, (u \wedge v)^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})\right)duv \\
&= p_{XY}^{-1}\gamma_1^2, \\
&Var\left(\int_{[0,1]} \frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\beta_1(u^{\frac{1}{\lambda}})du\right) \\
&= p_X^{-1} \int_{[0,1]^2} \left((u \wedge v)^{\frac{1}{\lambda}} - u^{\frac{1}{\lambda}}v^{\frac{1}{\lambda}}\right)\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u}duv \\
&= p_X^{-1}\gamma_2^2, \\
&Var\left(\int_{[0,1]} \frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial v}\beta_2(u^{\frac{1}{1-\lambda}})du\right) \\
&= p_Y^{-1} \int_{[0,1]^2} \left((u \wedge v)^{\frac{1}{1-\lambda}} - u^{\frac{1}{1-\lambda}}v^{\frac{1}{1-\lambda}}\right)\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial v}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v}duv \\
&= p_Y^{-1}\gamma_3^2.
\end{aligned}
$$

These integrals are tractable and we compute

$$
\gamma_1^2 = f(\lambda, A)\left(\frac{A(\lambda)}{A(\lambda) + 2\lambda(1-\lambda)}\right),
$$

$$
\gamma_2^2 = f(\lambda, A)\left(\frac{\kappa^2(\lambda, A)(1-\lambda)}{2A(\lambda) - (1-\lambda) + 2\lambda(1-\lambda)}\right),
$$

$$
\gamma_3^2 = f(\lambda, A)\left(\frac{\zeta^2(\lambda, A)\lambda}{2A(\lambda) - \lambda + 2\lambda(1-\lambda)}\right).
$$

We now compute the covariance :

$$
\begin{aligned}
p_X^{-1}\gamma_{12} &:= cov\left(\int_{[0,1]} S_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})du, \int_{[0,1]} \beta_1(v^{\frac{1}{\lambda}})\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u}dv\right) \\
&= \int_{[0,1]}\int_{[0,1]} \mathbb{E}[S_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})\beta_1(v^{\frac{1}{\lambda}})]\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u}dudv, \\
&= p_X^{-1}\int_{[0,1]^2} \left(C((u \wedge v)^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})v^{\frac{1}{\lambda}}\right)\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u}dudv.
\end{aligned}
$$

Let us decompose the integrals into two parts, one under the segment $[0,1] \times [0,v]$, the other under $[0,1] \times [v,1]$. The first one gives

$$\int_{[0,1]} \int_{[0,v]} (C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) v^{\frac{1}{\lambda}}) \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dudv =$$
$$\frac{\kappa(\lambda, A)}{2} f(\lambda, A) \left(\frac{1-\lambda}{2A(\lambda) + (2\lambda - 1)(1-\lambda)} \right).$$

For the second part, using Fubini, we have :

$$\int_{[0,1]} \int_{[0,u]} (C(v^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) v^{\frac{1}{\lambda}}) \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dvdu.$$

For the right hand side of the minus sign, we compute :

$$\int_{[0,1]} \int_{[0,u]} C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) v^{\frac{1}{\lambda}} \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dvdu = \frac{\kappa(\lambda, A)}{2} f(\lambda, A).$$

For the last one, some substitutions be considered.

$$\int_{[0,1]} \int_{[0,u]} C(v^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dvdu. \tag{3.9}$$

Following the proof of Proposition 3.3 from [Genest and Segers, 2009], the substitution $v^{\frac{1}{\lambda}} = x$ and $u^{\frac{1}{1-\lambda}} = y$ yields

$$\begin{aligned}
&\int_{[0,1]} \int_{[0,u]} C(v^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) \frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dvdu \\
&= \lambda(1-\lambda) \int_{[0,1]} \int_{[0, y^{\frac{1-\lambda}{\lambda}}]} C(x,y) \frac{\partial C(x, x^{\frac{\lambda}{1-\lambda}})}{\partial u} x^{\lambda - 1} y^{-\lambda} dxdy \\
&= \lambda(1-\lambda)\kappa(\lambda, A) \int_{[0,1]} \int_{[0, y^{\frac{1-\lambda}{\lambda}}]} C(x,y) x^{\frac{A(\lambda)}{1-\lambda} - (1-\lambda) - 1} y^{-\lambda} dxdy.
\end{aligned}$$

Next, use the substitution $x = w^{1-s}$ and $y = w^s$. Note that $w = xy \in [0,1]$, $s = log(y)/log(xy) \in [0,1]$, $C(x,y) = w^{A(s)}$ and the Jacobian of the transformation is $-log(w)$. As the constraint $x < y^{-1+1/\lambda}$ reduces to $s < \lambda$, the integral becomes:

$$
\begin{aligned}
&-\lambda(1-\lambda)\kappa(\lambda, A)\int_{[0,\lambda]}\int_{[0,1]} w^{A(s)+(1-s)(A_2(\lambda)-(1-\lambda)-1)-s\lambda} log(w)dwds \\
&= \lambda(1-\lambda)\kappa(\lambda, A)\int_{[0,\lambda]} \left[A(s)+(1-s)(A_2(\lambda)-1-(1-\lambda))-s\lambda+1\right]^{-2} ds.
\end{aligned}
$$

and we thus obtain γ_{12} by doing the sum. Next we compute the following integral :

$$
\begin{aligned}
\frac{p_{XY}}{p_X p_Y}\gamma_{23} &:= cov\left(\int_{[0,1]}\beta_1(u^{\frac{1}{\lambda}})\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}du, \int_{[0,1]}\beta_2(v^{\frac{1}{1-\lambda}})\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v}dv\right) \\
&= \frac{p_{XY}}{p_X p_Y}\mathbb{E}\left[\int_{[0,1]}\int_{[0,1]}\beta_1(u^{\frac{1}{\lambda}})\beta_2(v^{\frac{1}{1-\lambda}})\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v}dudv\right] \\
&= \frac{p_{XY}}{p_X p_Y}\int_{[0,1]}\int_{[0,1]}\left(C(u^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) - u^{\frac{1}{\lambda}}v^{\frac{1}{1-\lambda}}\right)\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v}dudv.
\end{aligned}
$$

The second term can be easily handled and its value is given by :

$$
\int_{[0,1]}\int_{[0,1]} u^{\frac{1}{\lambda}}v^{\frac{1}{1-\lambda}}\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v}dudv = f(\lambda, A)\kappa(\lambda, A)\zeta(\lambda, A).
$$

For the first, use the substitutions $u^{\frac{1}{\lambda}} = x$ and $v^{\frac{1}{1-\lambda}} = y$. This yields :

$$
\lambda(1-\lambda)\int_{[0,1]}\int_{[0,1]} C(x,y)\frac{\partial C(x, x^{\frac{\lambda}{1-\lambda}})}{\partial u}\frac{\partial C(y^{\frac{1-\lambda}{\lambda}}, y)}{\partial v}x^{\lambda-1}y^{-\lambda}dxdy.
$$

Then, make the substitutions $x = w^{1-s}$, $y = w^s$ that were used for the preceding integral gives :

$$
\begin{aligned}
&-\lambda(1-\lambda)\kappa(\lambda, A)\zeta(\lambda, A)\int_{[0,1]}\int_{[0,1]} w^{A(s)+(1-s)(A_2(\lambda)-(1-\lambda)-1)+s(A_1(\lambda)-\lambda-1)} log(w)dwds \\
&= \lambda(1-\lambda)\kappa(\lambda, A)\zeta(\lambda, A)\int_{[0,1]} \left[A(s)+(1-s)(A_2(\lambda)-(1-\lambda)-1)+s(A_1(\lambda)-\lambda-1)+1\right]^{-2} ds.
\end{aligned}
$$

Similarly, the last covariance requires the same tools as used before, it is left to the reader to find that

$$
\zeta(\lambda, A)\left(f(\lambda, A)\left(\frac{\lambda^2 - A(\lambda)}{2A(\lambda) - \lambda + 2\lambda(1-\lambda)}\right)\right),
$$

under the segment $[0,1] \times [0,v]$ and,

$$\lambda(1-\lambda)\zeta(\lambda, A) \int_{[\lambda,1]} [A(s) + s(A_1(\lambda) - 1 - \lambda) - (1-s)(1-\lambda) + 1]^{-2} \, ds,$$

on the segment $[0,1] \times [1,v]$. Using all the tools depicted, we compute, in the same manner, the integral in Equation (3.3), (3.4), (3.5), (3.6), (3.7) and their values are given by

$$(3.3) = \lambda(1-\lambda) \left(\int_{[0,1]} [A(s) - (1-s)(1-\lambda) - s\lambda + 1]^{-2} \, ds - \frac{1}{(1+\lambda)(1+1-\lambda)} \right),$$

$$\begin{aligned} (3.4) = \lambda(1-\lambda) \Bigg(& \int_{[0,\lambda]} [A(s) - (1-s)(1-\lambda) - s\lambda + 1]^{-2} \, ds \\ & + \frac{\lambda}{A(\lambda) + \lambda(1-\lambda)} \left[\frac{1-\lambda}{A(\lambda) + 2\lambda(1-\lambda)} - \frac{1}{1+\lambda} \right] \Bigg), \end{aligned}$$

$$\begin{aligned} (3.5) = \zeta(\lambda, A)\lambda(1-\lambda) \Bigg[& \int_{[0,1]} [A(s) + s(A_1(\lambda) - \lambda - 1) - (1-s)(1-\lambda) + 1]^{-2} ds \\ & - \frac{\lambda}{(1+\lambda)(A(\lambda) + \lambda(1-\lambda))} \Bigg], \end{aligned}$$

$$\begin{aligned} (3.6) = \lambda(1-\lambda) \Bigg(& \int_{[\lambda,1]} [A(s) - (1-s)(1-\lambda) - s\lambda + 1]^{-2} \, ds \\ & + \frac{(1-\lambda)}{A(\lambda) + \lambda(1-\lambda)} \left[\frac{\lambda}{A(\lambda) + 2\lambda(1-\lambda)} - \frac{1}{1+1-\lambda} \right] \Bigg), \end{aligned}$$

$$\begin{aligned} (3.7) = \kappa(\lambda, A)\lambda(1-\lambda) \Bigg[& \int_{[0,1]} [A(s) + (1-s)(A_2(\lambda) - \lambda - 1) - s(1-\lambda) + 1]^{-2} ds \\ & - \frac{1-\lambda}{(1+1-\lambda)(A(\lambda) + \lambda(1-\lambda))} \Bigg]. \end{aligned}$$

And that's the end of the proof. □

Corollary 1 aims to give a closed form of the limit variance without the integral of the Pickands dependence function.

Proof of Corollary 1 We use the same notations as in Section 3.1. As $C(u,v) = uv$, we have that

$$(3.3) = (3.5) = (3.7) = 0.$$

And the integral in Equation (3.4) become :

$$\int_{[0,1]}\int_{[0,v]} u^{\frac{1}{\lambda}} v^{\frac{1}{1-\lambda}} dudv + \frac{\lambda^2(1-\lambda)}{1+\lambda(1-\lambda)}\left[\frac{1-\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+\lambda}\right] =$$
$$\lambda^2(1-\lambda)\left(\frac{1}{(1+\lambda)(1-\lambda)+\lambda+\lambda(1-\lambda)} + \frac{1}{1+\lambda(1-\lambda)}\left[\frac{1-\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+\lambda}\right]\right).$$

The multiplication by $(p_{XY}^{-1} - p_X^{-1})$ gives σ_{13}. Same is for Equation (3.6)

$$\int_{[0,1]}\int_{[0,u]} v^{\frac{1}{\lambda}} u^{\frac{1}{1-\lambda}} dvdu + \frac{\lambda(1-\lambda)^2}{1+\lambda(1-\lambda)}\left[\frac{\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+1-\lambda}\right] =$$
$$\lambda(1-\lambda)^2\left(\frac{1}{\lambda(1+1-\lambda)+1-\lambda+\lambda(1-\lambda)} + \frac{1}{1+\lambda(1-\lambda)}\left[\frac{\lambda}{1+2\lambda(1-\lambda)} - \frac{1}{1+1-\lambda}\right]\right).$$

And we obtain σ_{23} with multiplying by $(p_{XY}^{-1} - p_Y^{-1})$. In the independent case $A(t) = 1$ for every $t \in [0,1]$ implies that that $\kappa(\lambda, A) = \zeta(\lambda, A) = 1$ for every $\lambda \in [0,1]$. Then for σ_3^2, Equation (3.9) equals :

$$f(\lambda, 1)\left(\frac{1+\lambda(1-\lambda)}{2-(1-\lambda)+2\lambda(1-\lambda)}\right).$$

Summing all the elements gives

$$\int_{[0,1]^2}\left(C((u\wedge v)^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})v^{\frac{1}{\lambda}}\right)\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial u} dudv = f(\lambda,1)\left(\frac{1-\lambda}{2-(1-\lambda)+2\lambda(1-\lambda)}\right).$$

Same computations gives

$$\int_{[0,1]^2}\left(C(u^{\frac{1}{\lambda}}, (u\wedge v)^{\frac{1}{1-\lambda}}) - C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})v^{\frac{1}{1-\lambda}}\right)\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v} dudv = f(\lambda,1)\left(\frac{\lambda}{2-\lambda+2\lambda(1-\lambda)}\right).$$

In independent case, we have the following equality :

$$\int_{[0,1]^2}\left(C(u^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}}) - u^{\frac{1}{\lambda}} v^{\frac{1}{1-\lambda}}\right)\frac{\partial C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})}{\partial u}\frac{\partial C(v^{\frac{1}{\lambda}}, v^{\frac{1}{1-\lambda}})}{\partial v} dudv = 0.$$

We then obtain for σ_3^2 when we sum all the elements :

$$\sigma_3^2(\mathbf{p}) = \left(\frac{\lambda(1-\lambda)}{1+\lambda(1-\lambda)}\right)^2 \left(\frac{p_{XY}^{-1}}{1+2\lambda(1-\lambda)} - \frac{p_X^{-1}(1-\lambda)}{2-(1-\lambda)+2\lambda(1-\lambda)} - \frac{p_Y^{-1}\lambda}{2-\lambda+2\lambda(1-\lambda)}\right)$$

□

3.2 Proofs in complete data framework

This Lemma aims to show that, under complete data framework, the asymptotic behavior of the processes $\sqrt{T}(\hat{\nu}_T(\lambda) - \nu(\lambda))$ and $\sqrt{T}(\hat{\nu}_T^*(\lambda) - \nu(\lambda))$ is the same.

Proof of the asymptotic behavior of the corrected estimator It is readily verified that

$$\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right) =$$
$$\frac{1}{2}\left((1-\lambda)\int_{[0,1]} \mathbb{C}_T(x^{\frac{1}{\lambda}}, 1)dx + \lambda \int_{[0,1]} \mathbb{C}_T(1, x^{\frac{1}{1-\lambda}})dx\right) - \int_{[0,1]} \mathbb{C}_T(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx.$$

Using the same argument as in the proof of Theorem 1, we can show, with complete data, that $(1-\lambda)\int_{[0,1]} \mathbb{C}_T(x^{\frac{1}{\lambda}}, 1)dx \rightsquigarrow \delta_{\{0\}}$ and $\lambda \int_{[0,1]} \mathbb{C}_T(1, x^{\frac{1}{1-\lambda}})dx \rightsquigarrow \delta_{\{0\}}$, where $\delta_{\{0\}}$ refers to the Dirac measure at 0. We thus obtain that, by extended Slutsky's lemma (example 1.4.7 of [van der Vaart and Wellner, 1996])

$$\sqrt{T}\left(\hat{\nu}_T^{\mathcal{H}*}(\lambda) - \nu(\lambda)\right) \rightsquigarrow -\int_{[0,1]} N_C(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx.$$

That's what we wanted to prove. □

3.3 Concentration inequality with contaminated data

We will denote by S the index set of sane blocks. For the rest of the section, λ is a fixed constant between 0 and 1.

Lemma 2. *For every positive ϵ, it holds that*

$$\mathbb{P}\left\{ |\hat{\nu}_{MoN} - \nu| > \epsilon \right\} \leq \mathbb{P}\left\{ \left|\hat{\nu}_{T_j} - \nu\right| > \epsilon \right\}^{K\delta} 2^K, \quad j \in S.$$

Proof If at least $K/2$ sane blocks have an empirical estimate that is ϵ close to the expectation, then so is the MoN. Reversing the implication gives that event

$$\{|\hat{\nu}_{MoN} - \nu| > \epsilon\},$$

implies that at least $K/2$ of $\hat{\nu}_{T_j}$ for $j \in S$ has to be outside distance ϵ of ν. Namely,

$$\{|\hat{\nu}_{MoN} - \nu| > \epsilon\} \subset \left\{ \left|\sum_{j\in S} \mathbb{1}_{\{|\hat{\nu}_{T_j} - \nu| \geq \epsilon\}}\right| \geq \frac{K}{2} \right\}.$$

Then for the first inequality, we have

$$\begin{aligned}
\mathbb{P}\Big\{ |\hat{\nu}_{MoN} - \nu| > \epsilon \Big\} &\leq \mathbb{P}\left\{ \left|\sum_{j\in S} \mathbb{1}_{\{|\hat{\nu}_{T_j} - \nu| > \epsilon\}}\right| \geq \frac{K}{2} \right\}, \\
&\leq \mathbb{P}\left\{ \left|\sum_{j\in S} \mathbb{1}_{\{|\hat{\nu}_{T_j} - \nu| > \epsilon\}}\right| \geq K_s - \frac{K}{2} \right\}, \\
&\leq \mathbb{P}\left\{ \left|\sum_{j\in S} \mathbb{1}_{\{|\hat{\nu}_{T_j} - \nu| > \epsilon\}}\right| \geq K_s \left(1 - \frac{1}{2}(\frac{1}{2} + \delta)^{-1}\right) \right\}.
\end{aligned}$$

All these inequalities results from $K \geq K_s \geq K(2^{-1} + \delta)$ and that $K_s + K_o = K$. Notice that the random variable $\sum_{j\in S} \mathbb{1}_{\{|\hat{\nu}_{T_j} - \nu|\}}$ is distributed according to a binomial random variable with K_s trials and probability p_ϵ with

$$p_\epsilon = \mathbb{P}\left\{ \left|\hat{\nu}_{T_j} - \nu\right| > \epsilon \right\}.$$

It can thus be upper bounded by

$$\begin{aligned}
\sum_{n=\lceil K_s(1-\frac{1}{2}(\frac{1}{2}+\delta)^{-1})\rceil}^{K_s} \binom{K_s}{n} p_\epsilon^n \left(1 - p_\epsilon\right)^{n-K_s} &\leq p_\epsilon^{K_s\left(1-\frac{1}{2}(\frac{1}{2}+\delta)^{-1}\right)} \sum_{n=1}^{K_s} \binom{K_s}{n}, \\
&\leq p_\epsilon^{K_s\left(1-\frac{1}{2}(\frac{1}{2}+\delta)^{-1}\right)} 2^{K_s}, \\
&\leq p_\epsilon^{K\delta} 2^K.
\end{aligned}$$

When we use that $K_s(1 - 2^{-1}(2^{-1} + \delta)^{-1}) \geq K\delta$ and $K_s \leq K$. That is our statement. $\square$

Lemma 3. *For every $j \in S$ and $\epsilon > 0$, we have*

$$p_\epsilon \leq \mathbb{P}\left\{\left|\frac{1}{2T_j}\sum_{t\in B_j}|F(X_t)-G(Y_t)| - \frac{1}{2}\mathbb{E}\,|F(X)-G(Y)|\right| > \frac{\epsilon}{3}\right\} \tag{3.10}$$

$$+\,\mathbb{P}\left\{\underset{t\in B_j}{sup}\left|\hat{F}_{T_j}(X_t)-F(X_t)\right| > \frac{2\epsilon}{3}\right\} + \mathbb{P}\left\{\underset{t\in B_j}{sup}\left|\hat{G}_{T_j}(Y_t)-G(Y_t)\right| > \frac{2\epsilon}{3}\right\}. \tag{3.11}$$

Proof First, notice that we can obtain the following upper bound

$$\begin{aligned}
\left|\hat{\nu}_{T_j}-\nu\right| &= \left|\frac{1}{2T_j}\sum_{t\in B_j}|\hat{F}_{T_j}(X_t)-\hat{G}_{T_j}(Y_t)| - \frac{1}{2}\mathbb{E}|F(X)-G(Y)|\right|,\\
&\leq \left|\frac{1}{2T_j}\sum_{t\in B_j}\Big(|\hat{F}_{T_j}(X_t)-\hat{G}_{T_j}(Y_t)| - |F(X_t)-G(Y_t)|\Big)\right|\\
&+ \left|\frac{1}{2T_j}\sum_{t\in B_j}|F(X_t)-G(Y_t)| - \frac{1}{2}\mathbb{E}|F(X)-G(Y)|\right|.
\end{aligned}$$

The first expression can be bounded by

$$\begin{aligned}
&\left|\frac{1}{2T_j}\sum_{t\in B_j}\Big(|\hat{F}_{T_j}(X_t)-\hat{G}_{T_j}(Y_t)| - |F(X_t)-G(Y_t)|\Big)\right|,\\
&\overset{(a)}{\leq} \frac{1}{2T_j}\sum_{t\in B_j}\left||\hat{F}_{T_j}(X_t)-\hat{G}_{T_j}(Y_t)| - |F(X_t)-G(Y_t)|\right|,\\
&\overset{(b)}{\leq} \frac{1}{2T_j}\sum_{t\in B_j}\left|\hat{F}_{T_j}(X_t)-F(X_t) - \Big(\hat{G}_{T_j}(Y_t)-G(Y_t)\Big)\,|\right|,\\
&\overset{(c)}{\leq} \frac{1}{2}\underset{t\in B_j}{sup}\left|\hat{F}_{T_j}(X_t)-F(X_t)\right| + \frac{1}{2}\underset{t\in B_j}{sup}\left|\hat{G}_{T_j}(Y_t)-G(Y_t)\right|.
\end{aligned}$$

We used triangle inequality in (a), $||x| - |y|| \leq |x - y|$ in (b) and both triangle inequality and that $\sum_{t=1}^{T} x_t \leq T \underset{t \in \{1,\dots,T\}}{sup} x_t$ in (c). Since :

$$\left\{ \left| \hat{\nu}_{T_j} - \nu \right| \leq \epsilon \right\} \supseteq \left\{ \frac{1}{2} \underset{t \in B_j}{sup} \left| \hat{F}_{T_j}(X_t) - F(X_t) \right| \leq \frac{\epsilon}{3} \right\} \cap \left\{ \frac{1}{2} \underset{t \in B_j}{sup} \left| \hat{G}_{T_j}(Y_t) - G(Y_t) \right| \leq \frac{\epsilon}{3} \right\}$$
$$\cap \left\{ \left| \frac{1}{2T_j} \sum_{t \in B_j} |F(X_t) - G(Y_t)| - \frac{1}{2} \mathbb{E}|F(X) - G(Y)| \right| \leq \frac{\epsilon}{3} \right\}.$$

We thus obtain our lemma using union bound. □

Now we got to the proof of Theorem 2.

Proof of the concentration bound Let $j \in S$ be fixed. We can write Equation (2) in Lemma 2 such as

$$\mathbb{P}\left\{ |\hat{\nu}_{MoN} - \nu| > \epsilon \right\} \leq exp\left(K\delta log\left(p_\epsilon 2^{\frac{1}{\delta}} \right) \right). \tag{3.12}$$

The DKW inequality (see page 384 in [Boucheron et al., 2013], [Massart, 1990] or in the proof of Theorem 1 in [Alquier et al., 2020] for a similar application) gives us an upper bound for Equation (3.11) in the following form :

$$4exp\left(-\frac{8}{9} T_j \epsilon^2 \right) \leq 4exp\left(-\frac{2}{9} T_j \epsilon^2 \right).$$

Clearly, we have that

$$\mathbb{E}\left[\frac{1}{2T_j} \sum_{t \in B_j} |F(X_t) - G(Y_t)| \right] = \frac{1}{2} \mathbb{E}|F(X) - G(Y)|,$$

and, for every $t \in B_j$

$$\frac{1}{2T_j} |F(X_t) - G(Y_t)| \leq \frac{1}{T_j}.$$

Applying Hoeffding's inequality permits us to bound Equation (3.10) by

$$2exp\left(-\frac{2}{9} T_j \epsilon^2 \right).$$

Summing all these components and the use of Lemma 3 yields to

$$p_\epsilon \leq 6exp\left(-\frac{2}{9}T_j\epsilon^2\right).$$

Plugging this inequality in Equation (3.12) leads to

$$\mathbb{P}\left\{|\hat{\nu}_{MoN} - \nu| > \epsilon\right\} \leq exp\left(K\delta log\left(6e^{-\frac{2\epsilon^2 T_j}{9}}2^{\frac{1}{\delta}}\right)\right).$$

It can be set to η by choosing $K = log(1/\eta)\delta^{-1}$ and ϵ such that $6e^{-\frac{2\epsilon^2 T_j}{9}}2^{\frac{1}{\delta}} = 1/e$, or again

$$\epsilon = \frac{3}{\sqrt{T_j}}log\left(6e2^{\frac{1}{\delta}}\right) = \frac{3}{\sqrt{2}}\frac{log\left(6e2^{\frac{1}{\delta}}\right)}{\delta}\sqrt{\frac{log\left(1/\eta\right)}{T}}.$$

Hence the result. $\square$

Conclusion

This report dealt with estimation of Madograms with missing data and outliers setting. Two setups were addressed. The first one consists in estimating the Madogram under complete observations while considering all information that can be drawn from the observations of the margins. The second one concern how to estimate the Madogram in presence of outliers.

In the first framework, we suppose that the missing mechanism is independent of the observations' law. Strong consistency and weak convergence of the constructed estimator with missing data and its corrected version were asserted. Also, for a given $\lambda \in [0,1]$, we detailed the closed form of the variance for both estimators of their Gaussians limit law. The same considerations in the complete data framework were also investigated and follows from the general case. Numerical results made from several extreme value copula models support our findings.

In the second framework, we leverage the idea of Median-of-meaNs (MoN) to propose a robust estimator thanks to the median step. We proved a concentration inequality that this estimator does verify and assert its robustness against outliers. Nevertheless, the procedure does not allow that the proportion of outliers be unknown, which is restrictive. Numerical results allow us to identify illustrative cases under which the MoN-based estimator is better in sense of the mean squared error than the FMadogram but also cases which is not.

To go beyond, one may leverage the notion of the depth of a point in a multivariate dataset [Tukey, 1975]. A robust estimator of the bivariate copula could be

$$\hat{C}_T(u,v) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}_{\{\hat{F}_T(X_t)\leq u, \hat{G}_T(Y_t) \leq v \leq y\}} \mathbb{1}_{\{D_T(\hat{F}_T(X_t), \hat{G}_T(Y_t))\}}, \quad \forall (u,v) \in [0,1]^2,$$

where D_T is the empirical Tukey depth, that is, the Tukey depth with respect to $\hat{H}_T$. The objective of further work is to develop clustering methods for maxima. An interesting proposal is to see the λ-FMadogram as a dissimilarity measure.

For clustering with respect to maxima, one may consider the following dissimilarity measure between two random variables X and Y with law F and G by

$$\lambda(F,G) := \int_{[0,1]} \nu(u)du + log(2).$$

Using Proposition A.1(i), the following inequalities hold

$$\int_{[0,1]} \frac{1}{1+u(1-u)}du \geq \lambda(F,G) \geq 2log(\frac{2}{3}).$$

Furthermore, $\lambda(\cdot,\cdot)$ is symmetric, verify triangular inequality and its lower bound is unique and reached when $A(t) = t \vee 1-t$ which is suitable for clustering maxima. One drawback of this quantity is the integral step which add a further estimation.

Another quantity could be of interest in modeling pairwise maxima and follows the seminal idea of [Pickands, 1981] to estimate the Pickands dependence function. Let $S = -log(F(X))$ and $T = -log(G(Y))$, we define

$$\gamma(\lambda) = \frac{1}{2}\mathbb{E}\left[\left|\frac{S}{1-\lambda} - \frac{T}{\lambda}\right|\right],$$

with $\lambda \in (0,1)$. It follows after some details that we omit

$$\gamma(\lambda) = \frac{1}{2}\left(\int_{[0,1]} C(u^{1-\lambda},1)\frac{du}{u} + \int_{[0,1]} C(1,u^{\lambda})\frac{du}{u}\right) - \int_{[0,1]} C(u^{1-\lambda},u^{\lambda})\frac{du}{u}.$$

This form might permit us to prove a functional central limit theorem as we done in this report for the hybrid λ-FMadogram estimator.

Bibliography

[Alquier et al., 2020] Alquier, P., Chérief-Abdellatif, B.-E., Derumigny, A., and Fermanian, J.-D. (2020). Estimation of copulas via maximum mean discrepancy.

[Angrist and Pischke, 2008] Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

[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.

[Baraud et al., 2016] Baraud, Y., Birgé, L., and Sart, M. (2016). A new method for estimation and model selection: ρ-estimation. *Inventiones Mathematicae*.

[Beirlant et al., 2004] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley. Pagination: 522.

[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.

[Boucheron et al., 2013] Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration inequalities. A nonasymptotic theory of independence*. Oxford University Press.

[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.

[Cochran, 2007] Cochran, W. (2007). *Sampling Techniques, 3Rd Edition*. A Wiley publication in applied statistics. Wiley India Pvt. Limited.

[Coles et al., 1999] Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2:339 – 365.

[Cooley et al., 2006] Cooley, D., Naveau, P., and Poncet, P. (2006). *Variograms for spatial max-stable random fields*, pages 373–390. Springer New York, New York, NY.

[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., Radulovic, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847 – 860.

[Gaetan and Guyon, 2008] Gaetan, C. and Guyon, X. (2008). *Modélisation et statistique spatiales*. Mathématiques & applications. Springer, Berlin Heidelberg New York.

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

[Gudendorf and Segers, 2009] Gudendorf, G. and Segers, J. (2009). Extreme-value copulas.

[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.

[Huber, 1964] Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1):73 – 101.

[Huber, 2011] Huber, P. J. (2011). *Robust Statistics*, pages 1248–1251. Springer

Berlin Heidelberg, Berlin, Heidelberg.

[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.

[Lerasle et al., 2019] Lerasle, M., Szabó, Z., Mathieu, T., and Lecué, G. (2019). MONK – Outlier-Robust Mean Embedding Estimation by Median-of-Means. In *ICML 2019 - 36th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Long Beach, United States.

[Little R.J.A., 1987] Little R.J.A., R. D. (1987). *Statistical analysis with missing data*. Wiley Series in Probability and Statistics.

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

[Massart, 1990] Massart, P. (1990). The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality. *The Annals of Probability*, 18(3):1269 – 1283.

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

[Nemirovsky and Yudin, 1983] Nemirovsky, A. and Yudin, D. (1983). *Problem Complexity and Method Efficiency in Optimization.* John Wiley, New York.

[Oliveira and Galambos, 1977] Oliveira, J. D. T. and Galambos, J. (1977). The asymptotic theory of extreme order statistics. *International Statistical Review*, 47:230.

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

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

[Segers, 2014] Segers, J. (2014). Hybrid copula estimators.

[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, 2005] Smith, R. L. (2005). Max-stable processes and spatial extremes. In *None*.

[Tawn, 1988] Tawn, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika*, 75(3):397–415.

[Tukey, 1975] Tukey, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 523–531.

[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.

Appendix A

Study of the Pickands dependence function

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and (X, Y) be a bivariate random vector with values in $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. This random vector has a joint distribution function H and marginal distribution function F and G. We suppose the copula function of H is an extreme-value copula type, *i.e.* if and only if it admits a representation of the form

$$C(u, v) = (uv)^{A(log(v)/(log(uv)))}, \tag{A.1}$$

for all $u, v \in [0, 1]$ and where $A(\cdot)$ is the Pickands dependence function, *i.e.*, $A : [0, 1] \longrightarrow [1/2, 1]$ is convex and satisfies $t \vee (1 - t) \leq A(t) \leq 1$, $\forall t \in [0, 1]$.

We will call by λ-FMadogram the following quantity

$$\nu(\lambda) = \frac{1}{2}\mathbb{E}\left[\left|F(X)^{\lambda} - G(Y)^{1-\lambda}\right|\right]. \tag{A.2}$$

In the following proposition, we establish some properties of the λ-FMadogram.

Proposition A.1. *Let (X, Y) a $\mathbb{R}^2$-valued random vector of distribution H. We have, for each $\lambda \in [0, 1]$,* *(i)* $\nu(0) = \nu(1) = 0.25$, *and if* $\lambda \in (0,1)$,

$$\nu(\lambda) = \frac{A(\lambda)}{A(\lambda) + \lambda(1-\lambda)} - \frac{1}{2}\left(\frac{1}{1+\lambda} + \frac{1}{1+1-\lambda}\right). \tag{A.3}$$

(ii) $\frac{\lambda \vee (1-\lambda)}{\lambda \vee (1-\lambda) + \lambda(1-\lambda)} - \frac{1}{2}\left(\frac{1}{1+\lambda} + \frac{1}{1+1-\lambda}\right) \leq \nu(\lambda) \leq \frac{1}{1+\lambda(1-\lambda)} - \frac{1}{2}\left(\frac{1}{1+\lambda} + \frac{1}{1+1-\lambda}\right)$.

Proof To show *(i)* we define the following function,

$$\nu_\lambda \colon [0,1]^2 \to [0,1], \quad (u,v) \mapsto u^\lambda \vee v^{1-\lambda} - \frac{1}{2}(u^\lambda + v^{1-\lambda}).$$

Using Fubini-Tonelli and the equality $|u^\lambda - v^{1-\lambda}| = u^\lambda \vee v^{1-\lambda} - 2^{-1}(u^\lambda + v^{1-\lambda})$ gives,

$$\begin{aligned}
\nu(\lambda) &= \frac{1}{2}\left(\int_{[0,1]} C(x^{\frac{1}{\lambda}}, 1)dx + \int_{[0,1]} C(1, x^{\frac{1}{1-\lambda}})dx\right) - \int_{[0,1]} C(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx, \\
&= \frac{1}{2}\left(\int_{[0,1]} x^{\frac{1}{\lambda}}dx + \int_{[0,1]} x^{\frac{1}{1-\lambda}}dx\right) - \int_{[0,1]} x^{\frac{A(\lambda)}{\lambda(1-\lambda)}}dx, \\
&= \frac{1}{2}\left(\frac{\lambda}{1+\lambda} + \frac{1-\lambda}{1+1-\lambda}\right) - \frac{\lambda(1-\lambda)}{A(\lambda)+\lambda(1-\lambda)}, \\
&= \frac{A(\lambda)}{A(\lambda)+\lambda(1-\lambda)} - \frac{1}{2}\left(\frac{1}{1+\lambda} + \frac{1}{1+1-\lambda}\right)
\end{aligned}$$

Hence the result. The second statement is directly implied by the first one. □

Remark 6. *The upper bound (resp. the lower bound) in (i) is exactly the value of the* λ*-FMadogram when* X *and* Y *are independent (resp. perfectly positive dependent), i.e. when* $A(t) = 1$ *(resp.* $A(t) = t \vee (1-t)$*).*

Let $\mathcal{A}$ be the space of Pickands dependence functions. We will denote by $\kappa(\lambda, A)$ and $\zeta(\lambda, A)$ the two functional such as :

$$\begin{aligned}
\kappa &\colon [0,1] \times \mathcal{A} \to [0,1], \quad (t, A) \mapsto A(t) - A'(t)t, \\
\zeta &\colon [0,1] \times \mathcal{A} \to [0,1], \quad (t, A) \mapsto A(t) + A'(t)(1-t).
\end{aligned}$$

We are able to prove the two following lemmas.

Lemma A.1. *Using the properties of the Pickands dependence function, we have* *that*

$$0 \leq \kappa(t, A) \leq 1, \quad 0 \leq \zeta(t, A) \leq 1, \quad 0 < t < 1.$$

Furthermore, if A admits a second derivative on $]0,1[$, $\kappa(\cdot, A)$ (resp. $\zeta(\cdot, A)$) is a decreasing function (resp. an increasing function).

Proof First, using that the graph of a (differentiable) convex function lies above all of its tangents and using that $A(t) \geq t$ for $0 < t < 1$, we get

$$A'(t) \leq \frac{A(1) - A(t)}{1 - t} = \frac{1 - A(t)}{1 - t} \leq 1.$$

Same reasoning using $A(t) \geq 1 - t$ leads to

$$A'(t) \geq \frac{A(t) - A(0)}{t - 0} = \frac{A(t) - 1}{t} \geq -1.$$

If we suppose that A admits a second derivative, the derivative of κ (resp ζ) with respect to λ gives:

$$\kappa'(t, A) = -tA''(t) < 0, \quad \zeta'(t, A) = (1 - t)A''(t) > 0, \quad \forall t \in]0, 1[.$$

Using $\kappa(0) = 1$, $\kappa(1) = 1 - A'(1) \geq 0$ gives $0 \leq \kappa(\lambda, A) \leq 1$. As $\zeta(0) = 1 + A'(0) \geq 0$ and $\zeta(1) = 1$, we have $0 \leq \zeta(\lambda, A) \leq 1$.

Now, we can obtain the same result while removing the hypothesis of A admits a second derivative. As A is a convex function, for $x, y \in [0, 1]$, we have the following inequality:

$$A(x) \geq A(y) + A'(y)(x - y).$$

Take $x = 0$ and $y = t$ gives $1 \geq A(t) - tA'(t) = \kappa(t)$. Now, using that $-tA'(t) \geq -t$, clearly $A(t) - tA'(t) \geq A(t) - t \geq 0$. As $A(t) \geq max(t, 1 - t)$. We thus obtain the result. $\square$

Lemma A.2. *If A admits a derivative, then $\lim_{t \to 0^+} A'(t)$ and $\lim_{t \to 1^-} A'(t)$ exists and are finite.*

Proof As A is convex and derivable, it follows that $A'(\cdot)$ is increasing. Furthermore, in the proof of Lemma A.1, we show that $-1 \leq A'(t) \leq 1$ for every $t \in (0, 1)$ and therefore bounded. Then the two limits exist and are finite. $\square$

The partial derivatives of the extreme value copula are given by

$$\frac{\partial C(u,v)}{\partial u} = \begin{cases} \frac{C(u,v)}{u}\kappa\left(\frac{log(v)}{log(uv)}, A\right), & \text{if}\, 0 < u, v \leq 1, \\ 0, & \text{if}\, v = 0, \quad 0 \leq u \leq 1, \end{cases}$$

$$\frac{\partial C(u,v)}{\partial v} = \begin{cases} \frac{C(u,v)}{v}\zeta\left(\frac{log(v)}{log(uv)}, A\right), & \text{if}\, 0 < u, v \leq 1, \\ 0, & \text{if}\, u = 0, \quad 0 \leq v \leq 1. \end{cases}$$

The properties of A imply $0 \leq A(t) - tA'(t) \leq 1$ and $0 \leq A(t) + (1-t)A'(t) \leq 1$ where $t = log(v)/log(uv)$ (see Lemma A.1). Therefore, if $v \searrow 0$, then $\partial C(u,v)/\partial u \to 0$ as required. We also need that the functionals $t \mapsto \kappa(t, A)$ and $t \mapsto \zeta(t, A)$ be defined on 0 and 1 in order that the previous discussion holds. Such is always the case as stated by Lemma A.2. We set with extend by continuity

$$A'(0) = \lim_{t \to 0^-} A'(t), \quad A'(1) = \lim_{t \to 1^+} A'(t).$$

Thus, both functionals are defined for every $t \in [0,1]$. An extreme value copula verify smoothness Condition 2.1 of [Segers, 2012]. This implies the weak convergence of the empirical copula process if C is on extreme value.

Appendix B

Auxiliary results

In this appendix, we present some results of the literature that is cited is this report for self-consistency.

Theorem B.1 (Theorem 3 of [Fermanian et al., 2004])**.** *Suppose that H has continuous marginal distribution functions and that the copula function $C(x, y)$ has continuous partial derivatives. Then the empirical copula process $\{\mathbb{C}_T(u, v), 0 \leq u, v \leq 1\}$ converges weakly to a Gaussian process $\{N_C(u, v), 0 \leq u, v \leq 1\}$ in $\ell^\infty([0, 1]^2)$.*

Under the framework of Theorem B.1, the following proposition holds.

Proposition B.1 (Proposition 3 of [Naveau et al., 2009])**.** *Suppose that Assumption A holds and let J be a function of bounded variation, continuous from above and with discontinuities of the first kind. Then:*

$$T^{-1/2} \sum_{t=1}^{T} \left(J(\hat{F}_T(X_t), \hat{G}_T(Y_t)) - \mathbb{E}\left[J(F(X), G(Y))\right] \right),$$

converges in distribution to $\int_{[0,1]^2} N_C(u, v) dJ(u, v)$ where $N_C(u, v)$ and the integral is well defined as a Lebesgue-Stieltjes integral. The special case, $J(x, y) = 2^{-1}|x^\lambda - y^{1-\lambda}|$ provide the weak of convergence of the λ-FMadogram estimator. It holds that

$$T^{1/2} \left\{ \hat{\nu}_T(\lambda) - \frac{1}{2}\mathbb{E}\left[|F(X)^\lambda - G^{1-\lambda}(Y)|\right] \right\},$$

converges in distribution to $\int_{[0,1]^2} N_C(u,v)dJ(u,v)$ *where the latter integral satisfies :*

$$\int_{[0,1]^2} f(x,y)dJ(x,y) = \frac{1}{2}\int_{[0,1]} f(0,y^{1/(1-\lambda)})dy + \frac{1}{2}\int_{[0,1]} f(x^{1/\lambda},0)dx - \int_{[0,1]} f(x^{1/\lambda},x^{1/(1-\lambda)})dx, \tag{B.1}$$

for all bounded-measurable function $f:[0,1]^2 \mapsto \mathbb{R}$.

Lemma B.1. *(Lemma A.1 of [Marcon et al., 2017]) For* $\lambda \in [0,1]$, *let* H *be any distribution function in* $[0,1]^2$, *let* ν_λ *be the function defined by*

$$\nu_\lambda\colon [0,1]^2 \to [0,1], \quad (u,v) \mapsto u^\lambda \vee v^{1-\lambda} - \frac{1}{2}(u^\lambda + v^{1-\lambda}),$$

Then

$$\int_{[0,1]^2} \nu_\lambda(u,v)dH(u,v) = \frac{1}{2}\left(\int_{[0,1]} H(x^{\frac{1}{\lambda}},1)dx + \int_{[0,1]} H(1,x^{\frac{1}{1-\lambda}})dx\right) - \int_{[0,1]} H(x^{\frac{1}{\lambda}},x^{\frac{1}{1-\lambda}})dx. \tag{B.2}$$

Theorem B.2 (Theorem 2.3 in [Segers, 2014])**.** *If Assumptions A and C hold, then uniformly in* $u \in [0,1]^2$,

$$\begin{aligned}\sqrt{T}\left\{\hat{C}_T(u,v) - C(u,v)\right\} &= \sqrt{T}\left\{\hat{H}_T((F,G)^{\leftarrow}(u,v) - C(u,v)\right\} \\ &- \frac{\partial C(u,v)}{\partial u}\sqrt{T}\left\{\hat{F}_T(F^{\leftarrow}(u)) - u\right\}1_{(0,1)}(u) \\ &- \frac{\partial C(u,v)}{\partial v}\sqrt{T}\left\{\hat{G}_T(G^{\leftarrow}(v)) - v\right\}1_{(0,1)}(v) + \circ_{\mathbb{P}}(1),\end{aligned}$$

as $T \to \infty$. *Hence in* $l^\infty([0,1]^2)$ *equipped with the supremum norm, as* $T \to \infty$,

$$\left(\sqrt{T}\left\{\hat{C}_T(u,v) - C(u,v)\right\}\right)_{u,v\in[0,1]^2} \rightsquigarrow \left(\alpha(u,v) - \frac{\partial C(u,v)}{\partial u}\beta_1(u) - \frac{\partial C(u,v)}{\partial v}\beta_2(v)\right)_{u,v\in[0,1]^2}. \tag{B.3}$$

The processes α, β_1 *and* β_2 *have continuous trajectories almost surely. The right-hand side in (B.3) is well-defined because* $\beta_j(0) = \beta_j(1) = 1$ *almost surely with* $j \in \{1,2\}$.

Appendix C

Supplementary results

Some out of blue results are depicted here.

C.1 A Lemma for Equation (1.6)

Lemma C.1. *For all bounded-measurable function* $f : [0,1]^2 \mapsto \mathbb{R}$*, if* $J(s,t) = 2^{-1}|s^\lambda - t^{1-\lambda}|$*, then the following integral satisfies:*

$$\int_{[0,1]^2} f(x,y)dJ(x,y) = \frac{1}{2}\int_{[0,1]} f(1, y^{1/(1-\lambda)})dy + \frac{1}{2}\int_{[0,1]} f(x^{1/\lambda}, 1)dx - \int_{[0,1]} f(x^{1/\lambda}, x^{1/(1-\lambda)})dx.$$

Proof Let $A = [0,s] \times [0,t]$, a closed pavement of $[0,1]^2$, where $s,t \in [0,1]$. Thus, $A \in \mathcal{B}([0,1])^2$. Let us introduce the following indicator function :

$$f_{s,t}(x,y) = \mathbb{1}_{\{(x,y)\in[0,1]^2, 0\le x\le s, 0\le y\le t\}}.$$

Then, for this function, we have in one hand :

$$\int_{[0,1]^2} f_{s,t}(x,y)dJ(x,y) = J(s,t) - J(0,0) = \frac{1}{2}|s^\lambda - t^{1-\lambda}|,$$

in other hand, using the equality $2^{-1}|x-y| = 2^{-1}(x+y) - x \wedge y$, one has to show

$$\begin{aligned}\frac{1}{2}\left|s^{\lambda} - t^{1-\lambda}\right| &= \frac{s^{\lambda}}{2} + \frac{t^{1-\lambda}}{2} - s^{\lambda} \wedge t^{1-\lambda} \\ &= \frac{1}{2}\int_{[0,1]} f_{s,t}(x^{\frac{1}{\lambda}}, 1)dx + \frac{1}{2}\int_{[0,1]} f_{s,t}(1, y^{\frac{1}{1-\lambda}})dy - \int_{[0,1]} f_{s,t}(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx.\end{aligned}$$

Notice that the class

$$\begin{aligned}\mathcal{E} = \Big\{A \in \mathcal{B}([0,1]^2) : \int_{[0,1]^2} \mathbb{1}_A(x,y)dJ(x,y) = \int_{[0,1]} \mathbb{1}_A(x^{\frac{1}{\lambda}}, 1)dx \\ + \int_{[0,1]} \mathbb{1}_A(1, y^{\frac{1}{1-\lambda}})dy - \int_{[0,1]} \mathbb{1}_A(x^{\frac{1}{\lambda}}, x^{\frac{1}{1-\lambda}})dx\Big\},\end{aligned}$$

contains the class $\mathcal{P}$ of all closed pavements of $[0,1]^2$. It is otherwise a monotone class (or λ-system). Hence as the class $\mathcal{P}$ of closed pavement is a π-system, the class monotone theorem ensure that $\mathcal{E}$ contains the sigma-field generated by $\mathcal{P}$, that is $\mathcal{B}([0,1]^2)$.

This result holds for simple function $f(x,y) = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}$ where $\lambda_i \in \mathbb{R}$ and $A_i \in \mathcal{B}([0,1]^2)$ for all $i \in \{1,\dots,n\}$. We then can prove the identity for positive measurable function by approximation with an increasing sequence of simple function using Beppo-Levy theorem. This identity is then extended to measurable function $f : [0,1]^2 \mapsto \mathbb{R}$ considering $f = f_+ - f_-$ with $f_+ = max(f,0)$ and $f_- = min(-f,0)$. We take the function bounded-measurable in order that the left hand size of the equality is well defined as a Lebesgue-Stieljes integral. □

C.2 A counter example against variance's monotony with respect to an increasing positive dependence

First, notive that, under dependency condition, the variance of the λ-FMadogram evaluated in $\lambda = 0.5$ is equal to $1/150$.

Lemma C.2. *Let us consider $A(t) = 1 - \theta t + \theta t^2$ where $\theta \in [0,1]$. If we take $\lambda = 0.5$, there exist $\theta \in (0,1)$ such that*

$$Var\left(\int_{[0,1]} N_C(u^{\frac{1}{\lambda}}, u^{\frac{1}{1-\lambda}})du\right) > \frac{1}{150}. \tag{C.1}$$

Proof For this dependence function, we have immediately :

$$\kappa(\lambda, A) = 1 - \theta\lambda^2, \quad \zeta(\lambda, A) = 1 - \theta(1-\lambda)^2.$$

For $\lambda = 0.5$, we notice that $\kappa(0.5, A) = \zeta(0.5, A)$. By a simple change of variable, we notice that :

$$\int_0^{0.5} [A(s) + (1-s)(2A(0.5) - 0.5 - 1) - 0.5s + 1]^{-2} ds =$$
$$\int_{0.5}^{1} [A(s) + s(2A(0.5) - 0.5 - 1) - 0.5(1-s) + 1]^{-2} ds.$$

By substitution, we have for the chosen copula that,

$$\int_{0.5}^{1} [A(s)+s(2A(0.5)-0.5-1)-0.5(1-s)+1]^{-2} ds = \int_{0.5}^{1} [\frac{3}{2} - s(\theta+1-2A(0.5)) + s^2\theta] ds.$$

Let us take $\theta = 2A(0.5) - 1$, which implies by direct computation that $\theta = 2/3 > 0$. Let us make of use of this lemma :

Lemma C.3. *Let a, b be two reals. Note $I_n = \int_{\mathbb{R}} (ax^2 + b)^n dx$, then :*

$$I_n = \frac{2n-3}{2b(n-1)} I_{n-1} + \frac{x}{2b(n-1)(ax^2+b)}.$$

Proof An integration by parts gives and some algebra gives:

$$\begin{aligned} I_{n-1} &= \frac{x}{(ax^2+b)^{n-1}} + 2(n-1) \int_{\mathbb{R}} \frac{ax^2}{(ax^2+b)^n} dx, \\ &= \frac{x}{(ax^2+b)^{n-1}} + 2(n-1) I_{n-1} - 2b(n-1) I_n. \end{aligned}$$

Solving the equation for I_n gives the result. $\square$

We want to compute the following quantity :

$$\int_{0.5}^{1} [\frac{3}{2} + s^2 \frac{2}{3}]^{-2} ds.$$

The lemma gives :

$$\int_{0.5}^{1}[\frac{3}{2}+s^2\frac{2}{3}]^{-2}ds = 36\int_{0.5}^{1}[4s^2+9]^{-2}ds$$
$$= 2\left(\frac{7}{20}+\int_{0.5}^{1}(4s^2+9)^{-1}ds\right),$$
$$= 2\left(\frac{7}{20}+\frac{1}{6}\int_{1/3}^{2/3}\frac{1}{u^2+1}du\right).$$

Where we have made the substitution $u = 2s/3$ in the third line. Then :

$$\int_{0.5}^{1}[\frac{3}{2}+s^2\frac{2}{3}]^{-2}ds = 2\left[\frac{7}{20}+\frac{1}{6}(atan(2/3)-atan(1/3))\right] \approx 0.142596.$$

For the last integral, we have, by substitution for $\lambda = 0.5$ and $\theta = 2/3$:

$$\int_{0}^{1}[A(s)+(1-s)(2A(0.5)-0.5-1)+s(2A(0.5)-0.5-1)+1]^{-2}ds = \int_{0}^{1}[\frac{13}{6}-\frac{2}{3}s+\frac{2}{3}s^2]^{-2}ds.$$

Then, we are able to compute :

$$\int_{0}^{1}[\frac{13}{6}-\frac{2}{3}s+\frac{2}{3}s^2]^{-2}ds = 36\int_{0}^{1}(13-4s+4s^2)ds \overset{u=(2s-1)}{=} 36\int_{0}^{1}((2s-1)^2+12)^2ds,$$
$$= 18\int_{-1}^{1}(u^2+12)^{-2}du \overset{\text{Lemma}}{=} \frac{3}{4}\left(\frac{2}{13}+\int_{-1}^{1}\frac{1}{u^2+12}du\right),$$
$$\overset{v=u/(2\sqrt{3})}{=} \frac{6}{52}+\frac{3}{8\sqrt{3}}\int_{\frac{-1}{2\sqrt{3}}}^{\frac{1}{2\sqrt{3}}}\frac{1}{v^2+1}dv.$$

$$\int_{0}^{1}[\frac{13}{6}-\frac{2}{3}s+\frac{2}{3}s^2]^{-2}ds = \frac{\sqrt{3}}{8}\left(atan(\frac{1}{2\sqrt{3}})-atan(-\frac{1}{2\sqrt{3}})\right)+\frac{6}{52} \approx 0.23707.$$

Summing all the components of the variance gives $Var\left(\int_{[0,1]}N_C(u^{\frac{1}{\lambda}},u^{\frac{1}{1-\lambda}})du\right) \approx 0.00713 > 1/150$, which gives our counterexample. $\square$