Non Parametric Contributions to the Study of Copula Symmetries

Lorenzo Frattarolo

A thesis presented for the degree of
Doctor of Philosophy

Centre d'Économie de la Sorbonne
Université Paris 1
Panthéon-Sorbonne

Dipartimento di Economia
Università Ca' Foscari Venezia

September 2014

Contents

List of Tables

Abstract

The modelling of dependence of random variables, a central subject in probability and statistics, it is becoming, more and more, important in finance and economics, being the core background for the description and evaluation of systemic risk and the insurgence of financial instabilities. The aim of this thesis, is to contribute to the design of a new generation of non parametric statistics, that are able to investigate some characteristics of the dependence structure of a system of random variables and random processes, in order to act as a guidance for parametric, statistical and economic modelling.

Abstract

Modellare la dipendenza di variabili casuali, un tema centrale in probabilità e statistica, sta diventando sempre più importante in economia e finanza, essendo il necessario substrato per la descrizione e la valutazione del rischio sistemico e l'insorgenza di instabilità finanziarie. Lo scopo di questa tesi è di contribuire alla progettazione di una nuova generazione di statistiche non parametriche, che siano in grado di investigare alcune caratteristiche della struttura di dipendenza di sistemi di variabili casuali e processi stocastici, in modo da fare da guida per modelli parametrici di tipo sia statistico che economico.

Abstract

Modélisation de la dépendance des variables aléatoires, un thème central de la probabilité et de statistiques, il est de plus en plus important dans l'économie et de la finance, comme le substrat nécessaire pour la description et l'évaluation du risque systémique et l'apparition de l'instabilité financière. L'objectif de cette thèse est de contribuer à la conception d'une nouvelle génération de statistiques non paramétriques, qui sont en mesure d'enquêter sur certaines caractéristiques de la structure de dépendance des systèmes de variables aléatoires et processus stochastiques, afin d'agir comme un guide pour les modèles paramétriques taper les deux termes statistiques et économiques.

Introduction

The Subprime and Sovereign bond crisis, partly originated from an incomplete and unreliable description of the dependences among financial and economic variables and the resulting mispricing of risks associated to them. The prototypical example were the triple A rating of MBSs and CDOs , and the incorrect valuation of exposures of other asset classes to the real estate market, but also wrong evaluation of the risk of contagion among the sovereign bonds of different European countries. The modeling of dependence of random variables, a central subject in probability and statistics, it is becoming, more and more, important in finance and economics, being the core background for the description and evaluation of systemic risk and the insurgence of financial instabilities. The aim of this thesis, is to contribute to the design of a new generation of non parametric statistics, that are able to investigate some characteristics of the dependence structure of a system of random variables and random processes, in order to act as a guidance for parametric, statistical and economic modelling.

Three are the main ingredients used to understand dependence: copula functions, symmetries and empirical processes. In particular, we study the validity of some symmetry of the copula of the system (i.e. invariance of the copula under some transformation), using non parametric estimator of the copula and of the transformed copula. The asymptotic properties of the estimators, and of the tests derived from them, can be obtained using some recent advances in empirical process theory.With our focus on copula functions, we can have a proper description of dependence beyond the usual one based on second moments of the underlying random variables, the covariance. In fact,it is well known, that only linear dependence can be captured by the covariance and that it can characterize completely only the multivariate normal distribution. The concept of copula, due to Sklar [79], that will be formally introduced in the first chapter, allows to separate the effect of dependence from the effects of the marginal distributions. In this way a general and convenient description of dependence, among random variables, is possible. For those reasons in recent years copula functions became a central tool in many applied fields [35],[17].

In particular, in the present thesis, we are concerned with maps of copula functions that are still copulas (transformations). If the copula and the transformed copula are the same we say that the copula have a symmetry. The probabilistic interpretation of this symmetry can ,then, characterize some properties of the dependence among the random variables considered.Most of the literature on copula symmetries focus on permutation symmetry, exchangability [61],[46] and the related time reversibility [2], we instead will focus on conditional independence and reflection symmetry, but will introduce the concept using the an easy example.

The simplest possible example of this line of reasoning is the product symmetry, corresponding to independence. Consider the transformation Π, product of the marginals:

$$\Pi\left(C\left(u,v\right)\right) = C\left(u,1\right)C\left(1,v\right) = uv \tag{1}$$

If and only if, two random variables have a copula $C(u,v)$, such that $C(u,v) = \Pi(C(u,v)) = uv$, they are independent. Incidentally, testing for this symmetry, was one of the main motivations that did lead Deheuvels [26] to the introduction of the empirical copula (although the concept was already, independently, introduced in [71] in the context of rank statistics) and to the use of empirical process theory for showing the weak convergence to a Gaussian process (see [29] [80], [76] and [13] for a modern treatment). In chapter 2, we study the weak convergence of an empirical process related to the conditional version of this symmetry.

In addition, the symmetry introduced in (1) is a relevant case, also because, under it, the covariance function of empirical copula process is easy to compute, leading to an analytical knowledge of the distributions of test statistics based on it. For a general copula, instead, the determination of the covariance function of the empirical copula process, requires the knowledge of the true copula and its derivatives. This was the main historical motivation that made, for a long time, independence tests, the only non parametric inference procedure based on copula characteristics. The situation changed, when resampling procedures like bootstrap [29] and the multiplier method [66], were introduced in this context. In particular, after the simulation study in [10] for comparing different methods, that showed the superiority of the multiplier approach, even if it requires the estimation of copula derivatives, a large number of test for different copula properties, were proposed (see [56] and references therein). In addition the recent extension of the multiplier method to strongly mixing variables [12],[11] leads directly to applications in finance and economics. The test of rank reflection symmetry developed in the chapter 3 is an example in this direction.

Finally,the improvement of multiplier method, along a different direction, is the subject the last part of the thesis, in chapter 4. There, we propose alternative estimators of copula derivatives that, based on orthogonal polynomials, go beyond the simple finite differences estimators currently used in the literature.

The thesis is organized as follows: In Chapter 1 we introduce the mathematical background by stating some basic definitions and results, which will be repeatedly used throughout this thesis. The definition of copula, and most of its notable properties, will be discussed at first. Then, techniques and results for weak convergence in general metric spaces based on outer integrals are quickly introduced. The last part of the chapter is devoted to orthogonal polynomials.

In chapter 2 Conditional dependence is expressed as a projection map in the trivariate copula space. The projected copula, its sample counterpart and the related process are defined. The weak convergence of the projected copula process to a tight centered Gaussian Process is obtained under weak assumptions on copula derivatives.

In chapter 3 We propose a new non parametric test of rank reflection symmetry, also known as radial symmetry, of copula functions valid in any number of dimensions and for strongly mixing random variables. The possibility of applying the test jointly to a high number of weakly dependent random variables, allows applications to financial time series whose asymmetric dependence has already been documented and linked to financial contagion. The asymptotic theory for the test is based on new result in empirical processes theory that allows for sub exponentially strongly mixing data. Simulation based study of empirical size and power is conducted and example of applications are provided. A presentation, on preliminar results of this chapter, was given at the PhD lunch seminars of the CES at the MSE of Paris 1 university. This work has, also, been accepted and will be presented at the ERCIM 2014 conference, in December.

Chapter 4 is of technical nature and its aim is to show that we can use orthogonal polynomials in the estimation of the copula derivatives, needed for the multiplier method. This could increase the power of test procedures based on the empirical copula processes. Result, from this chapter, were presented at the IWFOS 2014 conference and an extended abstract is published in the peer-reviewed proceedings "Contributions in infinite-dimensional statistics and related topics".

Finally proofs and technical arguments for all the chapters are deferred to the appendix.

Chapter 1

Mathematical Background

1.1 Copula and Dependence

A Copula, can be, simply, introduced as a joint cumulative distribution function on the hypercube, with uniform marginals. The advantage of using Copulas, in modeling the distribution of a D dimensional random vector, resides in the separation between the univariate marginal distribution of the components and their stochastic dependence, carried by their copula. This property, comes, directly, from the representation of the joint cumulative distribution function given in the Sklar's Theorem.

Theorem 1 *Sklar's Theorem [79] Let F be a D-dimensional cumulative distribution function with margins F_d for $d = 1, \ldots, D$. Then there exist a copula C such that for all* $\mathbf{x} \in \mathbb{R}^D$

$$F(\mathbf{x}) = C\left(F_1(x_1), \ldots, F_D(x_D)\right) \tag{1.1}$$

If all marginals are continuous, then C is unique and given by

$$C(\mathbf{u}) = F\left(F_1^-(u_1), \ldots, F_D^-(u_D)\right)$$

otherwise, is uniquely determined on $\mathrm{ran}F_1 \times \ldots \times \mathrm{ran}F_d$. *Conversely, if C is a copula and $F_1, \ldots, F_D$ are distribution function , then the function F defined in 1.1 is a joint cumulative distribution function with marginals $F_1, \ldots, F_D$.*

The F_d^- denote the generalized inverse of F_d, that can be defined by:

$$F^-(u) \equiv \begin{cases} \inf\{x \in \mathbb{R} | F(x) \geq u\} & 0 < u < 1 \\ \sup\{x \in \mathbb{R} | F(x) = 0\} & u = 0 \end{cases} \tag{1.2}$$

and satisfy:

$$F_d\left(F_d^-(u)\right) = u \tag{1.3}$$

The most notable properties of multivariate copulas are summarized in the following proposition:

proposition 1 *Copula Properties*

- *Frechet-Hoffending Bounds*

$$W(\mathbf{u}) = \max\left(\sum_{d=1}^{D} u_d - D + 1, 0\right) \leq C(\mathbf{u}) \leq \min(u_1, \ldots, u_d) = M(\mathbf{u}) \qquad (1.4)$$

The lower bound is a copula in 2 dimensions only, the upper bound correspond to perfect positive dependence (all the variables can be expressed as a monotonic increasing function of one of them)

- *Independence if for each* $d \in \{1, \ldots, D\}$, F_d *is continuous* $X_1, \ldots, X_D$ *are independent if and only if* $C(\mathbf{u}) = \prod_{d=1}^{D} u_d = \Pi(\mathbf{u})$

- *Lipshitz continuity* C *is Lipschitz-continuos with respect to the* L^1*-norm on* $[0,1]^D$

$$|C(\mathbf{u}) - C(\mathbf{v})| \leq \sum_{d=1}^{D} |u_d - v_d| \qquad (1.5)$$

- *Invariance under increasing transformations*
Let $\alpha = (\alpha_1, \ldots, \alpha_D)^T$ *a vector of strictly increasing transformations. If the marginals are continuous, the copula of* $(X_1, \ldots, X_D)^T$ *is equal to the copula of* $(\alpha_1 \circ X_1, \ldots, \alpha \circ X_D)^T$

In the following proposition, we show, that the population versions of association measures can be expressed as a functional of the difference among a copula and the independence copula.

proposition 2 *Copulas and Association measures*

$$\begin{aligned}
\rho_{X,Y} &= 12 \int_{[0,1]^2} (C(u,v) - uv)\, dudv \\
\tau_{X,Y} &= 4 \left(\int_{[0,1]^2} C(u,v)\, dC(u,v) - \int_{[0,1]^2} uvdudv \right)
\end{aligned}$$

1.1.1 Differentiability, Darsow Products, conditional independence

The importance of copula derivatives was pioneered in a seminal paper by Darsow and coauthors [24], where they show the relationship among them, conditional independence and Markov processes. In the following proposition, we summarize some basic properties of copula derivatives:

proposition 3 *Copulas Derivatives*

- *First partial derivatives of a copula* $\dfrac{\partial C(\mathbf{u})}{\partial u_d}$ *exist almost everywhere in* $[0,1]^D$, *being derivatives of a monotonic function*

- *Partial derivatives are bounded*

$$0 \leq \frac{\partial C(\mathbf{u})}{\partial u_d} \leq 1$$

- *Partial derivatives are non decreasing function in non deriving arguments*

The relation among copula partial derivatives and conditional probabilities is given in the following theorem from [24].

Theorem 2 *Let ω be and event in the sample space of the random vector* $\mathbf{X}$. *If the random vector* $\mathbf{X}$ *has copula* C *then*

$$\begin{aligned} & \mathbb{E}\left(\mathbb{I}\left(X_d < x_d \middle| X_1, \ldots, X_{d-1}, X_{d+1}, \ldots X_D\right)\right)(\omega) \\ = & \frac{\partial}{\partial u_d} C\left(F_1\left(X_1(\omega)\right), \ldots, F_{d-1}\left(X_{d-1}(\omega)\right), F_d\left(x_d\right), F_{d+1}\left(X_{d+1}(\omega)\right), \ldots, F_D\left(X_D(\omega)\right)\right) \end{aligned}$$

This representation of conditional probabilities, in terms of copula derivatives and marginals, allows them to introduce a product operation, that can be used to characterize conditional independence and Markovianity.

Definition 1 *Darsow $\star$ product Let A be an M dimensional copula and B be an L dimensional copula. Define* $A \star B : [0,1]^{M+L-1} \mapsto [0,1]$

$$A \star B\left(u_1, \ldots, u_{M+L-1}\right) = \int_0^{x_M} \frac{\partial}{\partial \xi} A\left(u_1, \ldots, u_{M-1}, \xi\right) \frac{\partial}{\partial \xi} B\left(\xi, u_{M+1}, \ldots, u_{M*+L-1}\right) d\xi$$

It should be noted, that $A \star B$ is an $(M+L-1)$-copula and that the product is distributive over convex combinations, is associative (in the sense that $(A \star B) \star C = A \star (B \star C)$) and is continuous in each place.
The following theorem, links Darsow product and conditional independence and it is a corollary of the more general theorem 3.3 about Markovianity in their paper.

Theorem 3 *$\star$-product and conditional independence Let $C_{XYZ}\left(u_X, u_Y, u_W\right)$ a three dimensional copula of the three variables X, Y, Z and $C_{XY}\left(u_X, u_Z\right) = C_{XYZ}\left(u_X, 1, u_W\right)$ and $C_{YZ}\left(u_Y, u_Z\right) = C_{XYZ}\left(1, u_Y, u_Z\right)$. Then*

$$\begin{aligned} & \mathbb{E}\left(\mathbb{I}\left(X < x\right) \mathbb{I}\left(Z < z\right) \middle| Y\right) = \mathbb{E}\left(\mathbb{I}\left(X < x\right) \mathbb{I} \middle| Y\right) \mathbb{E}\left(\mathbb{I}\left(Z < z\right) \middle| Y\right) \\ \Leftrightarrow \quad & C_{XYZ} = C_{XY} \star C_{YZ} \end{aligned}$$

1.2 Weak Convergence and Empirical Processes

This section provides a quick introduction to convergence in distribution for random elements that are not measurable. This is necessary for the study of empirical processes, since they can be seen as taking values in non separable Banach spaces, and even in the most elementary cases are non Borel measurable. The use this theory allows to derive asymptotic results for the empirical copula process, one of the main object of interest of this thesis.
Let $(\mathbb{D}, d)$ a metric space and let $\left(\mathbb{P}_N\right)_{N \in \mathbb{N}}$ and $\mathbb{P}$ be probability measures on $(\mathbb{D}, \mathcal{D})$, where $\mathcal{D}$ is the Borel σ-field on $\mathbb{D}$. The definition of weak convergence, that is denoted by $\mathbb{P}_N \rightsquigarrow \mathbb{P}$, is defined by requiring:

$$\int_{\mathbb{D}} f d\mathbb{P}_N \to \int_{\mathbb{D}} f d\mathbb{P} \qquad \text{for all} f \in C_b(\mathbb{D}) \tag{1.6}$$

Where $C_b(\mathbb{D})$ is the space of all real-valued bounded and continuous function on $\mathbb{D}$ [5]. Considering $\mathbb{D}$-valued random variables $(X_N)_{N\in\mathbb{N}}$ and X, $\mathbb{X}_n \rightsquigarrow \mathbb{X}$ is equivalent to:

$$\mathbb{E}f(X_N) \to \mathbb{E}f(X) \qquad \text{for all} f \in C_b(\mathbb{D}) \tag{1.7}$$

In this formulation of weak convergence, all random variables must be Borel measurable and this condition, easily, fails to hold for non separable metric spaces. A classical example of relevant random variable on a non separable metric space, that is not measurable, is the empirical distribution function of iid random variables $X_1, \ldots, X_N$ on $[0,1]$ that can be seen as a random variable in $D[0,1]$, the space of all cadlag function on the unit interval which are right continuous and posses left limit. If we equip $D[0,1]$ with the sup norm, the space is not separable and the measurability fails to hold for the random variables empirical distribution function :

$$\hat{F}_N(x) = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(X_i \leq x), \qquad x \in [0,1] \tag{1.8}$$

and the associated empirical process [5]

$$\hat{\mathbb{F}}_N(x) = \sqrt{n}\left(\hat{F}_N(x) - F(x)\right), \qquad x \in [0,1] \tag{1.9}$$

Lots of solutions were proposed ,to overcome this difficulty. The most fruitful one, proposed in recent years by J. Hoffman-Jorgensen and applied in [82],[77],[55], is to avoid, completely, the requirement of measurability, by introducing the concept of outer integral, expectation and probability.

Definition 2 *Outer integral and outer probability Let T be an arbitrary map from a probability space $(\omega, \mathcal{A}, \mathbb{P})$ to the extended real line $\bar{\mathbb{R}}$. The outer integral of T with respect to $\mathbb{P}$ is defined as*

$$\mathbb{E}^*T = \inf\left\{\mathbb{E}U : U \geq T, U : \Omega \mapsto \bar{\mathbb{R}}\,\text{measurable and}\,\mathbb{E}U\,exist\right\}. \tag{1.10}$$

The outer probability of an arbitrary subset $B \subset \Omega$ is defined as

$$\mathbb{P}^*(B) = \inf\left\{\mathbb{P}(A) : A \supset B, A \in \mathcal{A}\right\} \tag{1.11}$$

The infima in the latter definition are always achieved by lemma 1.2.1 in [82]. The introduction of outer expectation, leads to the possibility of defining weak convergence, outer almost sure convergence and convergence in outer probability, for non measurable maps.

Definition 3 *Let $X_N : \Omega \mapsto \mathbb{D}$ be arbitrary maps defined on some probability spaces $(\Omega_N, \mathcal{A}_n, \mathbb{P}_N)$,$(\Omega, \mathcal{A}, \mathbb{P})$*

1. *If X is Borel measurable we say that X_N weakly converges to X, denoted by $X_n \rightsquigarrow X$, if and only if*

$$\mathbb{E}^*f(X_N) \to \mathbb{E}f(X) \qquad \text{for all} f \in C_b(\mathbb{D}) \tag{1.12}$$

2. *If X_N, X are defined on a common probability space we say that X_N converges outer almost surely to X if $d(X_N, X)^* \to 0$ almost surely for some version of $d(X_N, X)^* \to 0$. This is denoted by $X \stackrel{\text{as}*}{\to} X_N$.*

3. *If X_N, X are defined on a common probability space we say that X_N converges in outer probability to X if $\mathbb{P}^*(d(X_N, X) > \epsilon) \to 0$ for every $\epsilon > 0$ and is denoted by $X \stackrel{\mathbb{P}^*}{\to} X_N$.*

Now, we focus on the relevant example of space of bounded functions over some set T, $\ell^\infty(T)$, because the general empirical process is the most important example of a sequence of maps in a space of the latter form. Given a sample $X_1, \ldots, X_N$ of random variables with distribution P on an arbitrary sample space $\mathcal{X}$ we define the empirical measure as $\mathbb{P}_N = N^{-1} \sum_{i=1}^N \delta_{X_i}$ where δ_x is the Dirac measure at x. For $f \in \mathcal{F} = \{f : \mathcal{X} \to \mathbb{R}\}$ we set $\mathbb{P}_N f = N^{-1} \sum_{i=1}^N f(X_i)$ and define the empirical process as

$$\mathbb{G}_N f = \sqrt{n}(\mathbb{P}_N - P) f = \frac{1}{\sqrt{N}} \sum_{i=1}^N f(X_i) - Pf,$$

which can be seen as an element of $\ell^\infty(\mathcal{F})$ provided

$$\sup_{f \in \mathcal{F}} |f(x) - Pf| < \infty \qquad \text{for every } x$$

Then, since $D[0,1] \subset \ell^\infty[0,1]$ the example of the empirical process $\hat{\mathbb{F}}_N$ as defined in 1.9 is considered if we take $\mathcal{F} = \{\mathbb{I}(0 \le t \le x) : x \in [0,1]\}$.
We follow the functional central limit theorem of [64] section 10 , in giving conditions for weak convergence in $\ell^\infty(T)$

Theorem 4 *Theorem 10.2 in Pollard [64] Let (T, ρ) be a totally bounded pseudo metric space and let $\{X_N(\omega, t) : t \in T\}$ be a sequence of random processes on T. If and only if*

1. *Convergence of the marginals: for any finite subset $\{t_1, \ldots, t_k\}$ of T, $(X_N(\cdot, t_1), \ldots, X_N(\cdot, t_k))$ converge weakly to a limit $(X(\cdot, t_1), \ldots, X(\cdot, t_k))$*

2. *Asymptotic Equicontinuity: For any positive ϵ and η, there exists a positive δ such that*

$$\limsup_{n \to \infty} \mathbb{P}^* \left\{ \sup_{\substack{(s,t) \in T \times T \\ \rho(s,t) \le \delta}} |X_N(\omega, s) - X_N(\omega, t)| > \eta \right\} < \epsilon$$

then $X_N \rightsquigarrow X$

Results for the convergence of the empirical processes in the iid case can be found in chapter 2 of [82]. The hypothesis of iid random variables is quite strong and doesn't allow the application of those results in the time series context. For this reason we introduce strongly mixing random variables as in [70]. The strong mixing coefficient between two σ-fields $\mathcal{A}$ and $\mathcal{B}$ is defined by

$$\alpha(\mathcal{A},\mathcal{B}) = 2\sup\{\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B) : (A,B)\in\mathcal{A}\times\mathcal{B}\} \tag{1.13}$$

For a sequence $(X_i)_{i\in\mathbb{Z}}$ of random variables, in some polish space $\mathcal{X}$, let $\mathcal{F}_k = \sigma(X_i : i\leq k)$ and $\mathcal{G}_l = \sigma(X_i : i\geq l)$, the strong mixing coefficients $(\alpha)_{n\geq 0}$ of the sequence $(X_i)_{i\in\mathbb{Z}}$ are defined by:

$$\alpha_0 = 1/2 \quad \text{and} \quad \alpha_N = \sup\alpha(\mathcal{F}_k,\mathcal{G}_{k+n}) \quad \text{for any}\, n>0$$

The sequence $(X_i)_{i\in\mathbb{Z}}$ is said to be strongly mixing, in the sense of Rosenblatt, if $\lim_{n\uparrow\infty}\alpha_N = 0$. In the stationary case, this means that the σ-field $\mathcal{G}_n$ of the future after time n is asymptotically independent of $\mathcal{F}_0$, σ-field of the past before time 0. In chapter 7 of [68] under the condition

$$\alpha_n \leq cn^{-a}\, \text{for some real}\, a>1\, \text{and some constant}\, c\geq 1 \tag{1.14}$$

and stationarity, marginal convergence and asymptotic equicontinuity are demonstrated for the univariate and multivariate empirical process leading to a tight Gaussian limiting process $\mathbb{G}$ with covariance function:

$$\mathbb{C}\text{ov}(\mathbb{G}(x),\mathbb{G}(y)) = \sum_{t\in\mathbb{Z}}\mathbb{C}\text{ov}(\mathbb{I}(X_0\leq x),\mathbb{I}(X_t\leq y)) \tag{1.15}$$

The following theorems allow to obtain the weak convergence of map of processes.

Theorem 5 *Extended continuous mapping [82] Let $\mathbb{D}_n\subset\mathbb{D}$ and $g_N:\mathbb{D}_n\mapsto\mathbb{E}$ satisfy the following statements: if $x_N\to x$ with $x_N\in\mathbb{D}_N$, for every N and $x\in{}_0$, then $g_N(x_N)\to g(x)$, where $\mathbb{D}_0\subset\mathbb{D}$ and $g:\mathbb{D}_0\mapsto\mathbb{E}$. Let X_N be maps with values in $\mathbb{D}_N$, be Borel measurable and separable, and take values in $\mathbb{D}_0$. then*

1. *$X_N \rightsquigarrow X$ implies that $g_N(X_N)\rightsquigarrow g(X)$*
2. *$X_N \stackrel{\text{as}*}{\to} X$ implies that $g_n(X_n)\stackrel{\text{as}*}{\to} g(X)$*
3. *$X_N \stackrel{\mathbb{P}^*}{\to} X$ implies that $g_N(X_N)\stackrel{\mathbb{P}^*}{\to} g(X)$*

The next theorem requires the notion of Hadamard differentiability that we now define

Definition 4 *Hadamard Differentiability A map $\phi:\mathbb{D}_\phi\mapsto\mathbb{E}$ defined on a subset $\mathbb{D}_\phi$ of a normed space $\mathbb{D}$ that contains θ is called Hadamard differentiable at θ if there exist a continuous linear map $\phi'_\theta:\mathbb{D}\mapsto\mathbb{E}$ such that*

$$\left\|\frac{\phi(\theta+th_t)-\phi(\theta)}{t} - \phi'_\theta(h)\right\|_{\mathbb{E}}, \qquad as\, t\downarrow 0$$

for every $h_t\to h$ such that $\theta+th_t\in\mathbb{D}_\phi$ for all $t>0$.

and the following chain rule for Hadamard differentiability:

Theorem 6 *Chain Rule [82] Let* $\phi : \mathbb{D}_\phi \mapsto \mathbb{E}_\psi$ *and* $\mathbb{E}_\psi \mapsto \mathbb{F}$ *be maps on subsets* $\mathbb{D}_\phi$ *e* $\mathbb{E}_\psi$ *of normed spaces* $\mathbb{D}$ *e*$\mathbb{E}$ *let* ϕ *be Hadamard differentiable at* θ *tangentially to* $\mathbb{D}_0$ *and let* ψ *be Hadamard differentiable at* $\phi(\theta)$ *tangentially to* $\phi'_\theta(\mathbb{D}_0)$ *Then* $\psi \circ \phi : \mathbb{D}_\phi \mapsto \mathbb{F}$ *is Hadamard differentiable at* θ *tangentially to* $\mathbb{D}_0$ *with derivative* $\psi'_{\phi(\theta)} \circ \phi'_\theta$

Theorem 7 *Delta Method Let* $\mathbb{D}$ *and* $\mathbb{E}$ *be normed linear spaces. Let* $\phi : \mathbb{D}_\phi \subset \mathbb{D} \mapsto \mathbb{E}$ *be Hadamard differentiable at* θ *tangentially to* $\mathbb{D}_0$*. Let* $T_N : \Omega_n \mapsto \mathbb{D}_\phi$ *be maps such that* $r_N(T_N - \theta) \rightsquigarrow T$ *for numbers* $r_N \to \infty$ *and a random element* T *that takes values in* $\mathbb{D}_0$*. Then* $r_N(\phi(T_N) - \phi(\theta)) \rightsquigarrow \phi'_\theta(T)$ *. If* ϕ'_θ *is continuous on the whole space* $\mathbb{D}$*, then we also have* $r_N(\phi(T_N) - \phi(\theta)) - \phi'_\theta(r_N(T_N - \theta))$ *converges to zero in probability.*

We are now ready to introduce the object of interest of this thesis: the empirical copula

$$\begin{aligned} \hat{C}_N(u_1, \ldots, u_D) &= \frac{1}{N} \sum_{i=1}^{N} \prod_{d=1}^{D} \mathbb{I}\left(\hat{U}_{Nid} \leq u_i\right) \\ \hat{U}_{Nid} &= \hat{F}_{Ni}(X_{id}) \end{aligned}$$

that is a non parametric estimator of the copula, joining the random vector $\mathbf{X}$. We can define an empirical copula process by:

$$\hat{\mathbb{C}}_N(\mathbf{u}) = \sqrt{N}\left(\hat{C}_N(\mathbf{u}) - C(\mathbf{u})\right) \tag{1.16}$$

The weak convergence of the process, is easily obtained, from the weak convergence of the multivariate empirical process with uniform margins, using the delta method on the map [13]:

$$\Phi : \begin{cases} \mathbb{D}_\Phi \mapsto \ell^\infty [0,1]^D \\ H \mapsto H\left(H_1^-, \ldots, H_D^-\right) \end{cases} \tag{1.17}$$

where $\mathbb{D}_\Phi$ denotes the set of all distribution functions on $[0,1]^D$ whose marginal cdfs H_d satisfy $H_d(0) = 0$. In fact we have

$$\mathbb{C}_n = \sqrt{N}\left(\hat{C}_N - C\right) = \sqrt{N}\left(\Phi\left(\hat{G}_N\right) - \Phi(C)\right)$$

Where $\hat{G}_N$ is the empirical distribution function with uniform margins.
The representation

$$\hat{C}_N = \Phi\left(\hat{G}_N(\mathbf{u})\right) = \hat{G}_N\left(G_{N1}^-(u_1), \ldots, G_{ND}^-(u_D)\right)$$

for the empirical copula was introduced in [80]. The Hadamard differentiability of Φ for a variety of strictly stationary processes and under non restrictive assumptions for the copula derivatives is obtained in [13] In particular, they require, the empirical process, based on $\mathbf{X}$, to have the limit $\mathbb{B}_C$ and the copula derivatives, to satisfy:

A 1 *consider the strictly stationary sequence* $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ *. Let* F_d *the marginal distribution function of the sequence* $(X_{id})_{i\in\mathbb{Z}}$*. Define the uniform sequence on* $[0,1]^D$ $(\mathbf{U}_i)_{i\in\mathbb{Z}} =$

$(F_1(X_{iD}), \ldots, F_D(X_{iD}))$ *that has as empirical distribution function* G_N *and as distribution the copula* C. *the uniform multivariate empirical process* $\mathbb{G}_N = \sqrt{N}(G_N - C)$. $\hat{\mathbb{G}}_N$ *weakly converges in ell*$^\infty[0,1]^D$ *to a tight centered Gaussian field* $\mathbb{B}_C$ *concentrated on* $\mathbb{D}_0$*, where*

$$\mathbb{D}_0 = \left\{\alpha \in C[0,1]^D \,|\alpha(1,\ldots,1) = 0\, and\, \alpha(\mathbf{x}) = 0\, if\ some\ components\ of\,\mathbf{x}\, are\ equal\ to\, 0\right\} \tag{1.18}$$

A 2 *[76] For each* $j \in \{1,2,3\}$*, the jth first-order partial derivative* $\partial_j C$ *exists and is continuous on the set* $V_{D,j} := \left\{u \in [0,1]^D : 0 < u_j < 1\right\}$.

As we already said, this is true for strongly mixing data because in that case $\mathbb{B}_C$ is a tight, centered Gaussian process on $[0,1]^D$ with covariance function

$$\mathbb{C}\mathrm{ov}(\mathbb{B}_C(u), \mathbb{B}_C(v)) = \sum_{t\in\mathbb{Z}} \mathbb{C}\mathrm{ov}(\mathbb{I}(U_0 \leq u), \mathbb{I}(U_t \leq v)) = \sum_{t\in\mathbb{Z}} \{C_t((\mathbf{u},\mathbf{v})) - C_0((\mathbf{u},\mathbf{u}))\, C_0((\mathbf{v},\mathbf{v}))\}$$

With C_t being the 2D dimensional stationary autocopula linking random variables taken at times whose absolute difference is t. In particular, we have

$$C_0((\mathbf{u},\mathbf{v})) = C(\mathbf{u} \wedge \mathbf{v}) \tag{1.19}$$

$$C_0((\mathbf{u},\mathbf{u})) = C(\mathbf{u} \wedge \mathbf{u}) = C(\mathbf{u}) \tag{1.20}$$

and C is the D dimensional stationary copula. Under this conditions they derive

Theorem 8 *Hadamard Derivative of* Φ *Suppose condition* ***A 2*** *holds. Then* Φ *is Hadamard differentiable at* C *tangentially to* $\mathbb{D}_0$*. Its derivative at* C *in* $\alpha \in \mathbb{D}_0$ *is given by*

$$(\Phi'_c(\alpha))(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{d=1}^{D} \partial_d C(\mathbf{u})\, \alpha(1,\ldots u_d \ldots, 1).$$

where $\partial_d C$ *is defined as 0 on the set* $V_{D,j}$

Corollary 1 *Suppose* ***A 1*** *and* ***A2*** *hold. Then the empirical copula process* $\mathbb{C}_N = \sqrt{N}\left(\hat{C}_N - C\right)$ *weakly converges in* $\ell^\infty[0,1]^D$ *to a Gaussian field* $\mathbb{G}_C$*,*

$$\hat{\mathbb{C}}_N(\mathbf{u}) \rightsquigarrow \mathbb{C}(\mathbf{u}) = \mathbb{B}_C(\mathbf{u}) - \sum_{d=1}^{D} \partial_d C(\mathbf{u})\, \mathbb{B}_C(1,\ldots u_d \ldots, 1) \tag{1.21}$$

From the previous display, we can see, that the limiting process depend, in a complex way, on the true Copula. Aside the simple case of independence, a closed form evaluation of the limiting distribution is not available. For this reason bootstrap approximations of the limiting process are used. According to [10],the partial derivative multiplier method,introduced in the copula context by [73] is the best one, even if it requires the estimation of partial derivatives.
For strongly mixing random variables, it is based on the following dependent multiplier central limit theorem:

Definition 5 *Dependent multiplier sequence [11] Define a dependent multiplier sequence* $\{\xi_{i,n}\}_{i\in\mathbb{Z}}$ *i.e. a sequence that satisfies*

1. *The sequence* $\{\xi_{i,N}\}_{i\in\mathbb{Z}}$ *is strictly stationary with* $\mathbb{E}(\xi_{0,N}) = 0$, $\mathbb{E}\left(\xi_{0,N}^2\right) = 1$ *and* $\mathbb{E}(|\xi_{0,N}|^{\nu}) < \infty$ *for* $\nu > 2$ *and independent from the available sample*

2. *There exists a sequence* $\ell_n \to \infty$ *of strictly positive constants such that* $\ell_N = o(N)$ *and the sequence* $\{\xi_{i,N}\}_{i\in\mathbb{Z}}$ *is* ℓ_N*-dependent i.e.* $\xi_{i,N}$ *is independent from* $\xi_{i+h,N}$ *for all* $h > \ell_N$ *and* $i \in \mathbb{N}$*.*

3. *There exists a function* $\phi : \mathbb{R} \to [0,1]$*, symmetric around 0,continous at 0, satisfying* $\phi(0) = 1$ *and* $\phi(x) = 0$ *for all* $|x| > 1$ *such that* $\mathbb{E}(\xi_{0,N}\xi_{h,N}) = \phi(h/\ell_N)$ *for all* $h \in \mathbb{Z}$

Given M independent copies of the dependent multiplier sequence $\left\{\xi_{i,N}^{[1]}\right\}_{i\in\mathbb{Z}}, \ldots, \left\{\xi_{i,N}^{[M]}\right\}_{i\in\mathbb{Z}}$ we can define the new processes:

Definition 6 *Multiplier Empirical Process*

$$\tilde{\mathbb{B}}_N^{[m]}(\mathbf{u}) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \xi_{i,N}^{(m)} \left(\mathbb{I}\left(\hat{\mathbf{U}}_i \leq \mathbf{u}\right) - C(\mathbf{u})\right) \tag{1.22}$$

$$\tag{1.23}$$

Theorem 9 *Dependent Multiplier Central Limit Theorem [11]. Assume that* $\ell_n = O\left(n^{1/2-\epsilon}\right)$ *for some* $0 < \epsilon < 1/2$ *and that* $\mathbf{U}_1, \ldots, \mathbf{U}_n$ *is drawn from a strictly stationary sequence* $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ *whose strong mixing coefficients satisfy* $\alpha_r = O(r^{-\alpha}), a > 3 + 3d/2$*. Then,*

$$\left(\hat{\mathbb{G}}_N, \tilde{\mathbb{B}}_N^{[1]}, \ldots, \tilde{\mathbb{B}}_N^{[M]}\right) \rightsquigarrow \left(\mathbb{B}_C, \mathbb{B}_C^{(1)}, \ldots, \mathbb{B}_C^{(C)}\right)$$

in $\left\{\ell^{\infty}[0,1]^{D+1}\right\}^{M+1}$*, where* $\mathbb{B}_N^{(1)}, \ldots, \mathbb{B}_C^{(C)}$ *are independent copies of* $\mathbb{B}_C$*.*

1.3 Introduction to Orthogonal Polynomials

Discrete orthogonal polynomials, can be fruitfully, used to approximate function and their derivatives. Their use in estimation of derivatives on sampled data popularized by [72](see also [28])leads us to the idea, developed in chapter 4, of using them in the estimation of copula derivatives.

Given a positive Borel measure on $\mathbb{R}$,$\mu(x)$ such that $\int_{\mathbb{R}} |x|^n \, d\mu(x)$, with $n \geq 0$, orthogonal polynomials with respect to μ are defined as the set polynomials $p_i \equiv q_c = \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{i} c_k x^k$ of degree exactly i s.t.

$$\langle p_i, p_j \rangle = \int_{\mathbb{R}} p_i(x)\, p_j(x)\, d\mu(x) = 0 \qquad i \neq j \tag{1.24}$$

If the measure has finite discrete support $d\mu(x) = \sum_{k=1}^{K} \delta(x - x_k) w_k dx$, with $\delta(x - x_k)$ the point mass Dirac delta centered in x_k, we can use sums instead of integrals:

$$\langle p_i, p_j \rangle = \int_{\mathbb{R}} p_i(x) p_j(x) d\mu(x) = \sum_{k=1}^{K} p_i(x_k) p_j(x_k) w_k = 0 \qquad i \neq j \tag{1.25}$$

and, in this case, there are only K orthogonal polynomials.
There are different ways to normalize polynomials:

Definition 7 *Orthonormal Polynomials The polynomials p in a set of polynomials are orthonormal if they are mutually orthogonal and if* $\langle p_i, p_i \rangle = 1$.

Definition 8 *Monic Polynomials Polynomials in a set are said to be monic orthogonal polynomials if they are orthogonal, if the coefficient of the monomial of highest degree is 1, and their norms are strictly positive.*

One of the most important properties of orthogonal polynomials is the three terms recurrence relation that links the $k+1$ polynomial with the k and the $k-1$ polynomials

Theorem 10 *[41]For orthonormal polynomials, there exist a sequence of coefficients* α_k *and* β_k *such that*

$$\begin{aligned} \sqrt{\beta_{k+1}} p_{k+1} &= (x - \alpha_{k+1}) p_k(x) - \sqrt{\beta_k} p_{k-1}(x) \\ p_{-1}(x) \equiv 0, p_0(x) &\equiv \frac{1}{\sqrt{\beta_0}}, \beta_0 = \int_{\mathbb{R}} d\mu(x) \end{aligned} \tag{1.26}$$

where

$$\alpha_{k+1} = \frac{\langle x p_k, p_k \rangle}{\langle p_k, p_k \rangle} \tag{1.27}$$

and β_k *is computed such that* $\sqrt{\langle p_k, p_k \rangle} = 1$.

Theorem 11 *[41]For monic orthogonal polynomials, there exist a sequence of coefficients* α_k *and* β_k *such that*

$$\begin{aligned} p_{k+1} &= (x - \alpha_{k+1}) p_k(x) - \beta_k p_{k-1}(x) \\ p_{-1}(x) &\equiv 0, p_0(x) \equiv 1 \end{aligned} \tag{1.28}$$

where

$$\alpha_{k+1} = \frac{\langle x p_k, p_k \rangle}{\langle p_k, p_k \rangle} \tag{1.29}$$

$$\beta_k = \frac{\langle p_k, p_k \rangle}{\langle p_{k-1}, p_{k-1} \rangle} \tag{1.30}$$

Now let us focus on a measure with discrete finite support that put weight w_k on the node x_k. For that measure we have a relationship between weights nodes and the three terms recursions coefficients for orthonormal polynomials.

Theorem 12 *[33] There exists an orthogonal matrix Q such that*

$$Q^T \begin{pmatrix} 1 & \sqrt{w_1} & \sqrt{w_2} & \dots & \sqrt{w_N} \\ \sqrt{w_1} & x_1 & 0 & \dots & 0 \\ \sqrt{w_2} & 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{w_N} & 0 & 0 & \dots & x_n \end{pmatrix} Q \;=\; \begin{pmatrix} 1 & \sqrt{\beta_0} & 0 & \dots & 0 \\ \sqrt{\beta_0} & \alpha_0 & \sqrt{\beta_1} & \dots & 0 \\ 0 & \sqrt{\beta_1} & \alpha_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_n \end{pmatrix}$$

with $Q^T Q = I_N$.

Given this theorem, that is proved in [33] as a corollary of theorem 3.1, the knowledge of weights and nodes i.e. of the measure can be used to obtain the recursion coefficients with the use, for example, of the Lanczos algorithm for the tridiagonalization of matrices. If in addition we add a discretization procedure we have a strategy for numerically obtaining orthogonal polynomials for every measure.

Chapter 2

Empirical Projected Copula Process and Conditional Independence

2.1 Introduction

In this chapter, our purpose is to introduce the empirical projected copula process and explain its importance in non parametric conditional independence testing. The central role of conditional independence in statistical theory was first adressed in the paper of Dawid ([25]) in which he rephrases sufficency, ancillarity, exogeneity, identification, causal inference and other relevant statistical concepts in terms of conditional independence. In this paper, he also introduced the usual notation for conditional independence that we will use all along this chapter:

$$\begin{aligned} X_1 \perp\!\!\!\perp X_2 | X_3 \quad &\Leftrightarrow \\ \mathbb{P}\left(X_1 \leq x_1, X_2 \leq x_2 \,|X_3 = x_3\right) \quad &= \quad \mathbb{P}\left(X_1 \leq x_1 \,|X_3 = x_3\right) \mathbb{P}\left(X_2 \leq x_2 \,|X_3 = x_3\right). \end{aligned} \tag{2.1}$$

Conditional independence is a fundamental tool for probabilistic graphical models(PGM), [54] and a proper understanding of Granger causality [83]. To investigate Conditional Independence, both non parametric copulas estimators and empirical distributions have already been used independently in the literature: [6] consider empirical Bernstein copulas; [44] develop a strong consistent test based on empirical distribution function. The approach, that most resembles our methodology, is the use of non parametric pair copulas to test for conditional independence done in [45]. In this Chapter, using conditional copulas with three variables, we rephrase conditional independence for copula functions, [24]. The outcome is the definition of a projection map, the projected copula and the related empirical process. Then, considering recent advances in the study of empirical copula process [76], [13] we obtain the weak convergence of the projected empirical copula process to a tight centered Gaussian process, under weak assumptions for second derivatives in the conditioning argument.

The paper is organised as follows. We inroduce some notations and assumptions in Section 2.2, then, in Section 2.3, we develop the relationship (2.1) using copulas and introduce the projected copula, showing that it is the proper representation of conditional independence in the copula space, and we introduce and prove the weak convergence of the projected empirical copula process. Section 3.6 is devoted to the conclusion and possible extensions. The technical lemma 1 and application to finite difference derivatives are postponed to the appendix.

2.2 Notation and Assumptions

In this section we define some notations and introduce the assumptions needed for our main theorem. Thorough the paper, arguments in boldface are collection of single arguments, for example $\mathbf{x} \equiv \{x_i\}_{i=1}^k \equiv x_1, \ldots, x_k$, the maximum between a and b is $a \vee b$ and the minimum is $a \wedge b$. The n-th partial derivative in the arguments $\{x_{i_k}\}_{k=1}^n$ is written $\partial_{i_1 \ldots i_n}^{(n)} \equiv \frac{\partial^n}{\partial x_{i_1} \ldots \partial x_{i_n}}$ and when the arguments are all equal i abbreviated in $\partial_i^{(n)} \equiv \frac{\partial^n}{\partial x_i \ldots \partial x_i}$. For the first derivative the one is omitted. $\partial_i^{(1)} \equiv \partial_i$. The indicator function on the set A is $\mathbb{I}(A)$, the space of all bounded functions defined on A is $\ell^\infty(A)$ and the space of k-times differentiable functions on A is $C^k(A)$. Weak convergence and convergence in outer probability in Hoffman-Jorgesen sense [82] are denoted respectively by $\rightsquigarrow$ and $\xrightarrow{\mathbb{P}^*}$.We use o and O Landau symbols and their stochastic counterparts as defined in [81]. Given 3 random variables $\{X_i\}_{i=1}^3$, with marginals $\mathbb{P}(X_i \leq x_i) = F_i(x_i)$, $i = 1, 2, 3$ and joint cumulative distribution $\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3) = F(x_1, x_2, x_3)$ by Sklar' s theorem in three dimensions [59]
we know it exists a copula function $C : [0,1]^3 \mapsto [0,1]$ such that $F(x_1, x_2, x_3) = C(F_1(x_1), F_2(x_2), F_3(x_3))$. Any 3-variate copula is a 3-variate distribution function with uniform marginals. We denote the space of all such functions by $\mathfrak{C}_3$. Following [24] we define the conditional copulas of the first 2 variables given the third one $C_{U_1,U_2|U_3}(u_1, u_2 | u_3) = \partial_3 C(u_1, u_2, u_3)$. Let $(X_{11}, X_{21}, X_{31}) \ldots (X_{1N}, X_{2N}, X_{3N})$ be a random sample, distributed according to F, the empirical distribution function and its margins are $\hat{F}_N(x_1, x_2, x_3) = \frac{1}{N} \sum_{j=1}^N \prod_{i=1}^3 \mathbb{I}(X_{ij} \leq x_i)$ and $\hat{F}_{Ni}(x_i) = \frac{1}{N} \sum_{j=1}^N \mathbb{I}(X_{ij} \leq x_i)$.
The empirical copula is $\hat{C}_N(u_1, u_2, u_3) = \frac{1}{N} \sum_{j=1}^N \prod_{i=1}^3 \mathbb{I}\left(\hat{U}_{Nij} \leq u_i\right)$ with the pseudo-observations given by $\hat{U}_{Nij} = \hat{F}_{Ni}(X_{ij})$.
Several authors ([76],[13] and references therein) have studied - both in the iid and strongly mixing case - the weak convergence of the associated empirical process

$$\hat{\mathbb{C}}_N(u_1, u_2, u_3) = \sqrt{N}\left(\hat{C}_N(u_1, u_2, u_3) - C(u_1, u_2, u_3)\right) \tag{2.2}$$

and under the assumption:

A 3 *[76] For each $j \in \{1, 2, 3\}$, the jth first-order partial derivative $\partial_j C$ exists and is continuous on the set $V_{3,j} := \left\{u \in [0,1]^3 : 0 < u_j < 1\right\}$.*

they prove:

$$\hat{\mathbb{C}}_N(u_1, u_2, u_3) \rightsquigarrow \mathbb{C}(u_1, u_2, u_3) = \alpha_C(u_1, u_2, u_3) + \sum_{i=1}^3 \beta_{iC}(u_i)\, \partial_i C(u_1, u_2, u_3) \tag{2.3}$$

where $\alpha_C(u_1, u_2, u_3)$ is a C-Brownian Bridge on $[0,1]^3$ and $\beta_{iC}(u_i)$ are its margins.
To obtain our result, we need an additional assumption on the second derivative in the conditioning argument.

A 4 *The function* $u_3 \mapsto \partial_3^{(2)} C(u_1, u_2, u_3) \in C^0([0,1])$, $\forall u_1, u_2 \in [0,1]^2$

In addition be able to define the sample version of the projection, that we are going to introduce in the next section, we need a sequence of functional maps $\mathcal{D}_{N,i}^{(n)} : \ell^\infty\left([0,1]^3\right) \mapsto \ell^\infty\left([0,1]^3\right)$, satisfying several hypothesis in order to properly approximate the derivative $\partial_i^{(n)}$ and obtain weak convergence. In appendix A.1 the prototipical example of finite difference derivatives is discussed in detail.
All the limits are for $N \to \infty$. The first assumption guaranties that when $\mathcal{D}_{N,i}^{(n)}$ is applied to the subspace of functions for which the partial derivative exists: it is an uniform approximation of this derivative.

A 5 *For all* $G \in \ell^\infty\left([0,1]^3\right)$ *s.t.* $\exists \partial_i^{(n)} G$ *for* $i = 1, 2, 3$*:*

$$\sup_{\mathbf{u}\in[0,1]^3} \left|\mathcal{D}_{N,i}^{(n)} G(\mathbf{u}) - \partial_i^{(n)} G(\mathbf{u})\right| \leq R_N, \ \lim_{N\to\infty} R_N = 0$$

The next hypothesis guaranties that $\mathcal{D}_{N,i}^{(n)}$ when applied to the empirical copula is a consistent estimator of copula derivatives.

A 6 *For any copula:* $\sup_{\mathbf{u}\in[0,1]^3} \left|\mathcal{D}_{N,i}^{(n)} \hat{C}_N(\mathbf{u}) - \partial_i^{(n)} C(\mathbf{u})\right| \xrightarrow{\mathbb{P}^*} 0$ *for* $i = 1, 2, 3$

The following one allows the asymptotic integration by part in the conditioning argument. We need this assumption in order to avoid the derivation of a Gaussian process.

A 7 *Given* $f : u_3 \mapsto f(\mathbf{u}') = f(u_1', u_2', u_3)$, *s.t.* $f \in C^1([0,1])$, $u_1', u_2' \in [0,1]^2$

$$\sup_{\mathbf{u}\in[0,1]^3} \left|J\left(f, \hat{\mathbb{C}}_N\right)\right| \xrightarrow{\mathbb{P}^*} 0, \forall a \in [0,1]$$

$$J\left(f, \hat{\mathbb{C}}_N\right) = \int_0^a f(\mathbf{u}')\, \mathcal{D}_{N,3}^{(1)} \hat{\mathbb{C}}_N(\mathbf{u}) + \mathcal{D}_{N,3}^{(1)} f(\mathbf{u}')\, \hat{\mathbb{C}}_N(\mathbf{u})\, du_3 - f(\mathbf{u}')\, \hat{\mathbb{C}}_N(\mathbf{u})\Big|_{u_3=0}^{u_3=a}$$

The last one is a technical assumption on the rate of convergence of integrated difference between true derivative and its approximation, when we apply it to the true copula.

A 8 *When* $N \to \infty$*:* $\sqrt{N} \int_0^{u_i} \left(\mathcal{D}_{N,i}^{(n)} C(\mathbf{u}) - \partial_i^{(n)} C(\mathbf{u})\right) du_i \to 0$

2.3 Projection and Weak Convergence

In this section we introduce the projection map in $\mathfrak{C}_3$, show that conditional independence is equivalent to invariance with respect to this map and obtain, in theorem 1, the weak convergence and asymptotic normality of the empirical projected copula process.
Using the notion of conditional copulas, relationship (2.1) can be rewritten:

$$C(F_1(x_1), F_2(x_2), F_3(x_3)) \;=\; \int_0^{F_3(x_3)} \partial_3 C(F_1(x_1), 1, v_3)\, \partial_3 C(1, F_2(x_2), v_3)\, dv_3$$

Thus $X_1 \perp\!\!\!\perp X_2|X_3$ is equivalent to invariance with respect to the map $\Pi_{|3} : \mathfrak{C}_3 \mapsto \mathfrak{C}_3$ where:

$$\Pi_{|3}\left(C\left(u_1,u_2,u_3\right)\right) = \int_0^{u_3} \partial_3 C\left(u_1,1,v_3\right)\partial_3 C\left(1,u_2,v_3\right)dv_3 \tag{2.4}$$

The map $\Pi_{|3}$ is the map projection onto X_3 and the right hand side of (2.4) is the projected copula. The map is not new as it can be rephrased using the $\star$-product of [24] from which follows the fact that the projected copula is always in $\mathfrak{C}_3$.Analogously,the empirical projected copula is defined as:

$$\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right) = \int_0^{u_3} \mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(1,u_2,v_3\right)dv_3,$$

From wich follows the definition of the empirical projected copula process:

$$\hat{\mathbb{C}}_{N|3} = \sqrt{N}\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right) - \Pi_{|3}\left(C\left(u_1,u_2,u_3\right)\right)\right) \tag{2.5}$$

We are now in the position to state our main theorem:

Theorem 1 *Under **A3-A8**,*

$$\begin{aligned}\hat{\mathbb{C}}_{N|3}\left(u_1,u_2,u_3\right) &\rightsquigarrow \mathbb{C}_{|3}\left(u_1,u_2,u_3\right) = \\ \partial_3 C\left(u_1,1,u_3\right)\mathbb{C}\left(1,u_2,u_3\right) &- \int_0^{u_3}\partial_3^{(2)}C\left(u_1,1,v_3\right)\mathbb{C}\left(1,u_2,v_3\right)dv_3 \\ +\partial_3 C\left(1,u_2,u_3\right)\mathbb{C}\left(u_1,1,u_3\right) &- \int_0^{u_3}\partial_3^{(2)}C\left(1,u_2,v_3\right)\mathbb{C}\left(u_1,1,v_3\right)dv_3 \end{aligned} \tag{2.6}$$

We remark that since $\mathbb{C}_{|3}$ is a linear combination of Gaussian processes it is also a Gaussian process.

2.3.1 Proof of theorem 1

In this section, we prove the theorem 1. For the proof the following technical Lemma 1 is needed:

Lemma 1 *let $\Psi_N : \ell^\infty[0,1]^3 \mapsto \ell^\infty[0,1]^3$ be the map:*

$$\Psi_N(f) = \partial_3 C\left(u_1,u_2,u_3\right)f\left(u_1,u_2,u_3\right) - \int_0^{u_3} dv_3 \mathcal{D}^{(1)}_{N,3}C\left(u_1,u_2,v_3\right)f\left(u_1,u_2,v_3\right)$$

*then, under **A5-A8**, we have:*

$$\hat{\mathbb{C}}_{N|3} = \Psi_N\left(\hat{\mathbb{C}}_N\left(u_1,1,v_3\right)\right) + \Psi_N\left(\hat{\mathbb{C}}_N\left(1,u_2,v_3\right)\right) + o_{\mathbb{P}^*}(1)$$

The empirical projected copula process (2.5) can be rewritten as difference between the empirical projection map applied to the empirical copula and the empirical projection map applied to the true copula, plus the difference between the empirical and asymptotic projection map applied to the true copula:

$$\begin{aligned}\hat{\mathbb{C}}_{N|3} &= \sqrt{N}\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right)-\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)\right) \\ &+ \sqrt{N}\left(\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)-\Pi_{|3}\left(C\left(u_1,u_2,u_3\right)\right)\right) \qquad (2.7)\end{aligned}$$

We develop first the second term of the right hand of (2.7) representing the difference between the empirical and asymptotic projection map applied to the true copula:

$$\sqrt{N}\left(\hat{\Pi}_{N|3}-\Pi_{|3}\right)\circ\left(C\left(u_1,u_2,u_3\right)\right)= \qquad (2.8)$$

$$\begin{aligned}&= \sqrt{N}\left(\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)-\Pi_{|3}\left(C\left(u_1,u_2,u_3\right)\right)\right) \\ &= \sqrt{N}\left(\int_0^{u_3}dv_3\left(\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)\mathcal{D}_{N,3}^{(1)}C\left(1,u_2,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\partial_3C\left(1,u_2,v_3\right)\right)\right)\end{aligned}$$

$$\begin{aligned}&= \sqrt{N}\left(\int_0^{u_3}dv_3\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)\left(\mathcal{D}_{N,3}^{(1)}C\left(1,u_2,v_3\right)-\partial_3C\left(1,u_2,v_3\right)\right)\right. \\ &\qquad \left.+\partial_3C\left(1,u_2,v_3\right)\left(\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\right)\right) \\ &= \sqrt{N}\left(\int_0^{u_3}dv_3\left(\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\right)\left(\mathcal{D}_{N,3}^{(1)}C\left(1,u_2,v_3\right)-\partial_3C\left(1,u_2,v_3\right)\right)\right. \\ &\qquad +\partial_3C\left(u_1,1,v_3\right)\left(\mathcal{D}_{N,3}^{(1)}C\left(1,u_2,v_3\right)-\partial_3C\left(1,u_2,v_3\right)\right) \\ &\qquad \left.+\partial_3C\left(1,u_2,v_3\right)\left(\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\right)\right)\end{aligned}$$

. Using the last expression,$|\partial_3C|\leq 1$ and assumption **A5** ,we can bound the absolute value of (2.8):

$$\begin{aligned}&\left|\sqrt{N}\left(\hat{\Pi}_{N|3}-\Pi_{|3}\right)\circ\left(C\left(u_1,u_2,u_3\right)\right)\right| \\ \leq\ & \sqrt{N}\left(1+R_N\right)\int_0^{u_3}\left|\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\right|dv_3 \\ +\ & \sqrt{N}\int_0^{u_3}dv_3\left|\mathcal{D}_{N,3}^{(1)}C\left(u_1,1,v_3\right)-\partial_3C\left(u_1,1,v_3\right)\right|dv_3\end{aligned}$$

By the dominated convergence theorem, the use assumption **A8** on this bound, implies that the limit of the absolute value of (2.8) is zero, i.e. that (2.8) is $o\left(1\right)$.
We consider now the first term of the right hand of the relationship (2.7) that represents the difference beetween the application of the empirical projection map to the empirical

copula and the application of the empirical projection map to the true copula :

$$
\begin{aligned}
& \sqrt{N}\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right)-\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)\right) \qquad (2.9)\\
= & \sqrt{N}\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(1,u_2,v_3\right)dv_3\\
- & \sqrt{N}\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}C\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}C\left(1,u_2,v_3\right)dv_3 =\\
= & \sqrt{N}\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(u_1,1,v_3\right)\left(\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(1,u_2,v_3\right)-\mathcal{D}^{(1)}_{N,3}C\left(1,u_2,v_3\right)\right)dv_3\\
+ & \sqrt{N}\int_0^{u_3}\left(\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(u_1,1,v_3\right)-\mathcal{D}^{(1)}_{N,3}C\left(u_1,1,v_3\right)\right)\mathcal{D}^{(1)}_{N,3}C\left(1,u_2,v_3\right)dv_3\\
= & \sqrt{N}\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{C}_N\left(u_1,1,v_3\right)\frac{1}{\sqrt{N}}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(1,u_2,v_3\right)dv_3\\
+ & \sqrt{N}\int_0^{u_3}\frac{1}{\sqrt{N}}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}C\left(1,u_2,v_3\right)dv_3\\
= & \int_0^{u_3}\partial_3C\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(1,u_2,v_3\right)dv_3+o_{\mathbb{P}^*}\left(1\right)\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(1,u_2,v_3\right)dv_3\\
+ & \int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(u_1,1,v_3\right)\partial_3C\left(1,u_2,v_3\right)dv_3+o\left(R_N\right)\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(u_1,1,v_3\right)dv_3
\end{aligned}
$$

Where the last equality follows from **A5** and **A6**.
Now, under **A7** and **A5**, we have for any $\mathbf{u}\in[0,1]^3$

$$
\begin{aligned}
\int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(u_1,u_2,v_3\right)dv_3 &= \hat{\mathbb{C}}_N\left(u_1,u_2,u_3\right)-\int_0^{u_3}\left(\mathcal{D}^{(1)}_{N,3}1\right)\hat{\mathbb{C}}_N\left(u_1,u_2,v_3\right)dv_3\\
&= \hat{\mathbb{C}}_N\left(u_1,u_2,u_3\right)+o\left(R_N\right)\int_0^{u_3}\hat{\mathbb{C}}_N\left(u_1,u_2,v_3\right)dv_3\\
&= \hat{\mathbb{C}}_N\left(u_1,u_2,u_3\right)+o_{\mathbb{P}^*}\left(1\right)
\end{aligned}
$$

The last expression implies that the second and the fourth term in the last inequality of (2.9) are $o_{\mathbb{P}^*}(1)$.
Summarizing, we have shown that:

$$
\begin{aligned}
\hat{\mathbb{C}}_{N|3} &= \sqrt{N}\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right)-\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)\right)\\
&+ \sqrt{N}\left(\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)-\Pi_{|3}\left(C\left(u_1,u_2,u_3\right)\right)\right)\\
&= \sqrt{N}\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\left(u_1,u_2,u_3\right)\right)-\hat{\Pi}_{N|3}\left(C\left(u_1,u_2,u_3\right)\right)\right)+o\left(1\right)\\
&= \int_0^{u_3}\partial_3C\left(u_1,1,v_3\right)\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(1,u_2,v_3\right)dv_3\\
&+ \int_0^{u_3}\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N\left(u_1,1,v_3\right)\partial_3C\left(1,u_2,v_3\right)dv_3+o_{\mathbb{P}^*}\left(1\right) \qquad (2.10)
\end{aligned}
$$

Under **A7**,to obtain the result we can again "integrate by part". In particular, we can rewrite the last equality of (2.10) in the following way:

$$\begin{aligned}
\hat{\mathbb{C}}_{N|3} &= \int_0^{u_3} \partial_3 C(u_1,1,v_3)\,\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N(1,u_2,v_3)\,dv_3 \\
&+ \int_0^{u_3} \mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N(u_1,1,v_3)\,\partial_3 C(1,u_2,v_3)\,dv_3 + o_{\mathbb{P}^*}(1) \\
&= \Psi_N\left(\hat{\mathbb{C}}_N(1,u_2,v_3)\right) + J\left(\partial_3 C(u_1,1,v_3), \hat{\mathbb{C}}_N(1,u_2,v_3)\right) \\
&+ \Psi_N\left(\hat{\mathbb{C}}_N(u_1,1,v_3)\right) + J\left(\partial_3 C(1,u_2,v_3), \hat{\mathbb{C}}_N(u_1,1,v_3)\right) + o_{\mathbb{P}^*}(1) \\
&= \Psi_N\left(\hat{\mathbb{C}}_N(1,u_2,v_3)\right) + \Psi_N\left(\hat{\mathbb{C}}_N(u_1,1,v_3)\right) + o_{\mathbb{P}^*}(1)
\end{aligned}$$

Consider, now, the map Ψ_N, introduced the previous lemma toghether with the map : $\Psi : \ell^\infty[0,1]^3 \mapsto \ell^\infty[0,1]^3$

$$\Psi(f) = \partial_3 C(u_1,u_2,u_3)\,f(u_1,u_2,u_3) - \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,u_2,v_3)\,f(u_1,u_2,v_3).$$

Under **A5-A8**,using lemma 1, we have:

$$\hat{\mathbb{C}}_{N|3} = \Psi_N\left(\hat{\mathbb{C}}_N(u_1,1,v_3)\right) + \Psi_N\left(\hat{\mathbb{C}}_N(1,u_2,v_3)\right) + o_{\mathbb{P}^*}(1)$$

Because $\Psi_N(f_N) \to \Psi(f)$ whenever $f_N \to f$, under **A3**, we have (4.9) and under **A4**, $\Psi(f)$ is continous, the hypotesis of the extended continous mapping theorem (1.11.1) in [82] pg. 67 are satisfied and the result follows by the application of the theorem to $\Psi_N\left(\hat{\mathbb{C}}_N(u_1,1,v_3)\right) + \Psi_N\left(\hat{\mathbb{C}}_N(1,u_2,v_3)\right)$.

2.4 Limit Process under Conditional Independence

In the hypothesis $\mathbb{H}_0 : X_1 \perp\!\!\!\perp X_2|X_3$ and **A4** and the continuity of the second mixed derivatives of the copula, the limit could be further simplified as in the following lemma:

Lemma 2 *if $\mathbb{H}_0$ and **A4** and if the second mixed derivatives of the copula are continous then:*

$$\begin{aligned}
&\mathbb{C}_{|3}(u_1,u_2,u_3) = \mathbb{C}_{1\perp 2|3}(u_1,u_2,u_3) \equiv \\
&= \partial_3 C(u_1,1,u_3)\,\alpha_C(1,u_2,u_3) - \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,1,v_3)\,\alpha_C(1,u_2,v_3) \\
&+ \partial_3 C(1,u_2,u_3)\,\alpha_C(u_1,1,u_3) - \int_0^{u_3} dv_3 \partial_3^{(2)} C(1,u_2,v_3)\,\alpha_C(u_1,1,v_3) \\
&+ \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,u_2,v_3)\,\beta_{3C}(v_3) - \beta_{3C}(u_3)\,\partial_3 C(u_1,u_2,u_3) \\
&+ \mathbb{C}(u_1,u_2,u_3) - \alpha_C(u_1,u_2,u_3)
\end{aligned} \tag{2.11}$$

2.4.1 Proof of lemma 2

Under **A4** the integrand in (2.4) is derivable and we can obtain the derivative of the integral using the fundamental theorem of calulus. Then, in the hypothesis of conditional independence, we have:

$$\partial_3 C(u_1, u_2, u_3) = \partial_3 C(u_1, 1, u_3)\, \partial_3 C(1, u_2, u_3) \tag{2.12}$$

This is the property tested, for example, in [6] and [45].
From wich we can derive the following, using the continuty of mixed partial derivatives and Schwarz's theorem:

$$\partial_3\partial_2 C(u_1, u_2, u_3) = \partial_3\partial_2 C(1, u_2, u_3)\, \partial_3 C(u_1, 1, u_3) \tag{2.13}$$

$$\partial_3\partial_1 C(u_1, u_2, u_3) = \partial_3\partial_1 C(u_1, 1, u_3)\, \partial_3 C(1, u_2, u_3) \tag{2.14}$$

We are ready to develop the expression for $\mathbb{C}_{|3}$, using the definition of $\mathbb{C}$:

$$\begin{aligned}
& \mathbb{C}_{|3}(u_1, u_2, u_3) \\
= \; & \partial_3 C(u_1, 1, u_3) \left[\alpha_C(1, u_2, u_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, u_3) - \beta_{3C}(u_3)\, \partial_3 C(1, u_2, u_3)\right] \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1, 1, v_3) \left[\alpha_C(1, u_2, v_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, v_3) - \beta_{3C}(u_3)\, \partial_3 C(1, u_2, v_3)\right] \\
+ \; & \partial_3 C(1, u_2, u_3) \left[\alpha_C(u_1, 1, u_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, u_3) - \beta_{3C}(u_3)\, \partial_3 C(u_1, 1, u_3)\right] \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(1, u_2, v_3) \left[\alpha_C(u_1, 1, v_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, v_3) - \beta_{3C}(v_3)\, \partial_3 C(u_1, 1, v_3)\right]
\end{aligned}$$

Using (2.12) and the rule for the derivation of a product on the terms that multiply β_{3C}, we can obtain:

$$\begin{aligned}
= \; & \partial_3 C(u_1, 1, u_3) \left[\alpha_C(1, u_2, u_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, u_3)\right] - \beta_{3C}(u_3)\, \partial_3 C(u_1, u_2, u_3) \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1, 1, v_3) \left[\alpha_C(1, u_2, v_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, v_3)\right] \\
+ \; & \int_0^{u_3} dv_3 \left[\partial_3 \left(\partial_3 C(u_1, 1, v_3)\, \partial_3 C(1, u_2, v_3)\right) - \partial_3^{(2)} C(1, u_2, v_3)\, \partial_3 C(u_1, 1, v_3)\right] \beta_{3C}(v_3) \\
+ \; & \partial_3 C(1, u_2, u_3) \left[\alpha_C(u_1, 1, u_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, u_3)\right] - \beta_{3C}(u_3)\, \partial_3 C(u_1, u_2, u_3) \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(1, u_2, v_3) \left[\alpha_C(u_1, 1, v_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, v_3) - \beta_{3C}(v_3)\, \partial_3 C(u_1, 1, v_3)\right]
\end{aligned}$$

Then using again (2.12) and simplyfing terms that are equal:

$$\begin{aligned}
= \; & \partial_3 C(u_1, 1, u_3) \left[\alpha_C(1, u_2, u_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, u_3)\right] \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1, 1, v_3) \left[\alpha_C(1, u_2, v_3) - \beta_{2C}(u_2)\, \partial_2 C(1, u_2, v_3)\right] \\
+ \; & \partial_3 C(1, u_2, u_3) \left[\alpha_C(u_1, 1, u_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, u_3)\right] \\
- \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(1, u_2, v_3) \left[\alpha_C(u_1, 1, v_3) - \beta_{1C}(u_1)\, \partial_1 C(u_1, 1, v_3)\right] \\
+ \; & \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1, u_2, v_3)\, \beta_{3C}(v_3) - 2\beta_{3C}(u_3)\, \partial_3 C(u_1, u_2, u_3)
\end{aligned}$$

Collecting the terms in β_{2C} and β_{1C} and using again the rule for the derivation of a product, we have:

$$\begin{aligned}
&= \partial_3 C(u_1,1,u_3)\left[\alpha_C(1,u_2,u_3) - \beta_{2C}(u_2)\,\partial_2 C(1,u_2,u_3)\right] \\
&- \int_0^{u_3} dv_3 \partial_3\partial_3 C(u_1,1,v_3)\,\alpha_C(1,u_2,v_3) \\
&+ \beta_{2C}(u_2)\int_0^{u_3} dv_3\left[\partial_3\left(\partial_3 C(u_1,1,v_3)\,\partial_2 C(1,u_2,v_3)\right) - \partial_3 C(u_1,1,v_3)\,\partial_3\partial_2 C(1,u_2,v_3)\right] \\
&+ \partial_3 C(1,u_2,u_3)\left[\alpha_C(u_1,1,u_3) - \beta_{1C}(u_1)\,\partial_1 C(u_1,1,u_3)\right] \\
&- \int_0^{u_3} dv_3 \partial_3^{(2)} C(1,u_2,v_3)\,\alpha_C(u_1,1,v_3) \\
&+ \beta_{1C}(u_1)\int_0^{u_3} dv_3\left[\partial_3\left(\partial_3 C(1,u_2,v_3)\,\partial_1 C(u_1,1,v_3)\right) - \partial_3 C(1,u_2,v_3)\,\partial_3\partial_1 C(u_1,1,v_3)\right] \\
&+ \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,u_2,v_3)\,\beta_{3C}(v_3) - 2\beta_{3C}(u_3)\,\partial_3 C(u_1,u_2,u_3)
\end{aligned}$$

Up to this point everything is valid under the assumptions of theorem 1, but now we are going to use (2.13) and (2.14) that rely onthe continuity of mixed partial derivatives

$$\begin{aligned}
&= \partial_3 C(u_1,1,u_3)\left[\alpha_C(1,u_2,u_3) - \beta_{2C}(u_2)\,\partial_2 C(1,u_2,u_3)\right] \\
&- \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,1,v_3)\,\alpha_C(1,u_2,v_3) \\
&+ \beta_{2C}(u_2)\int_0^{u_3} dv_3\left[\partial_3\left(\partial_3 C(u_1,1,v_3)\,\partial_2 C(1,u_2,v_3) - \partial_2 C(u_1,u_2,v_3)\right)\right] \\
&+ \partial_3 C(1,u_2,u_3)\left[\alpha_C(u_1,1,u_3) - \beta_{1C}(u_1)\,\partial_1 C(u_1,1,u_3)\right] \\
&- \int_0^{u_3} dv_3 \partial_3^{(2)} C(1,u_2,v_3)\,\alpha_C(u_1,1,v_3) \\
&+ \beta_{1C}(u_1)\int_0^{u_3} dv_3\left[\partial_3\left(\partial_3 C(1,u_2,v_3)\,\partial_1 C(u_1,1,v_3) - \partial_1 C(u_1,u_2,v_3)\right)\right] \\
&+ \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,u_2,v_3)\,\beta(v_3) - 2\beta_{3C}(u_3)\,\partial_3 C(u_1,u_2,u_3)
\end{aligned}$$

Doing the integrals that multiply β_{2C} and β_{1C} and simplyfing, we obtain:

$$\begin{aligned}
&= \partial_3 C(u_1,1,u_3)\,\alpha_C(1,u_2,u_3) - \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,1,v_3)\,\alpha_C(1,u_2,v_3) \\
&+ \partial_3 C(1,u_2,u_3)\,\alpha_C(u_1,1,u_3) - \int_0^{u_3} dv_3 \partial_3^{(2)} C(1,u_2,v_3)\,\alpha_C(u_1,1,v_3) \\
&+ \int_0^{u_3} dv_3 \partial_3^{(2)} C(u_1,u_2,v_3)\,\beta(v_3) - \beta_{3C}(u_3)\,\partial_3 C(u_1,u_2,u_3) \\
&- \beta_{3C}(u_3)\,\partial_3 C(u_1,u_2,u_3) - \beta_{2C}(u_2)\,\partial_2 C(u_1,u_2,u_3) - \beta_{1C}(u_1)\,\partial_1 C(u_1,u_2,u_3)
\end{aligned}$$

The final expression comes from the definition of $\mathbb{C}$.

$$
\begin{aligned}
&= \partial_3 C\left(u_1,1,u_3\right)\alpha_C\left(1,u_2,u_3\right) - \int_0^{u_3} dv_3 \partial_3^{(2)} C\left(u_1,1,v_3\right)\alpha_C\left(1,u_2,v_3\right)\\
&+ \partial_3 C\left(1,u_2,u_3\right)\alpha_C\left(u_1,1,u_3\right) - \int_0^{u_3} dv_3 \partial_3^{(2)} C\left(1,u_2,v_3\right)\alpha_C\left(u_1,1,v_3\right)\\
&+ \int_0^{u_3} dv_3 \partial_3^{(2)} C\left(u_1,u_2,v_3\right)\beta\left(v_3\right) - \beta_{3C}\left(u_3\right)\partial_3 C\left(u_1,u_2,u_3\right)\\
&+ \mathbb{C}\left(u_1,u_2,u_3\right) - \alpha_C\left(u_1,u_2,u_3\right)
\end{aligned}
$$

2.5 Conclusion

The objective of this Chapter is to lay the theoretical foundation for a new, non parametric, test of conditional independence based on a symmetry principle with respect to a projection map. The novelty of our approach is, in defining a sample estimator, for the projection, using empirical copula processes. This makes our results valid under very general assumption and widens the range of applications for our findings. For what concern our hypothesis, a closer look to the example section of [76] reveals that the discontinuity at the boundaries of the most common copula first derivatives occurs only when two or more arguments are involved in the limit so our hypothesis of continuity of the second partial derivative only in the conditioning argument are verified for most of the examples. We are only more restictive in considering twice continously differentiable Archimedean copula generators and dependence function of extreme value copula twice continously differentiable in each argument. For what concerns derivative approximations, in the appendix is shown that finite difference approximation as in [36] are copula consistent approximation. With a second order copula consistent derivative approximation, using the multiplier central limit theorem [82] as in [10] it is possible to evaluate through simulation the limit process distribution. This would be done in a different paper. Concerning the simplyfied limiting process introduced in lemma 2, instead, the hypothesis are more restrictive, and only a detailed simulation study, could ascertain, if the gain in empirical power, that we expect from test based on the simplyfied process with respect tests based on $\mathbb{C}_{|3}$, could justify its use, even if the theoretical applicability of the test is more restricted. This is left for future research.

Chapter 3

Multivariate Normalized Rank Reflection Symmetry of Copula Functions

3.1 Introduction

Statistical description of multivariate data, has suffered from serious limitations, due to the mathematical difficulties associated multivariate probability distributions. The vast majority of model used, were based on the multivariate Gaussian distribution. The situation has slowly changed when Sklar [79] introduced copula functions, allowing to model, separately, the dependence structure and the marginals. Recently, copula functions have become essential ingredients, in many applied fields concerning the modeling of multivariate data, such as actuarial sciences, finance, hydrology and survival analysis [35] [17]. Nevertheless, in most high dimensional applications of copula methods, in particular when temporal dependence cannot be disregarded, the Gaussian and t copulas remain the preferred copula models. A test of the symmetry proposed in this paper, among other things, allows to justify or reject the use of those models. In addition, testing this property could be evidence of asymmetric dependence that was empirically documented by different statistical means in [57], [1] and [49] for financial time series. The asymmetry of dependence and, in particular, an increase of dependence during downturns is usually regarded as the occurence of financial contagion and can be also related to systemic risk being a sympthom of a negative collective response of the financial system. With respect to many of previously cited methods for contagion detection, our copula framework allows to have a measure that include the effect of more than two variables and investigate, directly, a change in dependence, different from a simultaneus change in the characteristics of marginal distributions as, for example, a sudden shift in all volatilities due, for example, to a shift in the volatility of a common factor. Those are important advantages with respect to correlation and conditional correlation based measures. Rejection of parametric models and detection of contagion, motivate the choice of datasets for our applications of the test, but consequences of this symmetry and the related radial symmetry of the joint distribution are relevant in other financial and econometric areas: for portfolio management([62]) for quantile regressions ([7]), for up and down barrier symmetry in multivariate barrier options ([20]) and put-call parity for multivariate assets([18]).The Asymptotic test, being based on empirical copula, survival empirical copula and dependent multiplier bootstrap, is completely non parametric and is valid under the vast majority of hypothesis underling most common strictly stationary parametric models. The difference between our test and already introduced tests of this type [8], [69],[27],[37]and [56], is the use of the survival empirical copula process, that allows to avoid the use of the combicatorially complex multivariate generalization of the relation between the copula and the survival copula (see for example [40],[18]), making the application of the test to dimensions greater than 2 much easier. We remark that this difficulty arises also in the parametric context. For example, using the fact that from every copula function, being symmetric or not, we may build a symmetric copula by simply drawing from a mixture of the copula function itself and its survival counterpart obtained by a mixture of the copula function itself and its survival counterpart [52], a test for the difference beetween an asymmetric copula and its symmetrized version could be conceived. This kind of test is hard to implement with a number of variables greater than two due to the complexity of obtaining a manegeable expression for the survival copula. In this regards an extensive simulation study concerning the empirical power of the test is conducted with a particular emphasis on temporal dependence and high number of variables. The paper is structured as follows: In section 3.2 we introduce the rank reflection symmetry and motivate its importance for the right modellization of extreme events and asymmetric dependence. Then in section 3.3 we discuss the asymptotics of the test and introduce the dependent multiplier bootstrap. In section 3.4 there is our extensive simulation study where we discuss the empirical size and power of the test. Section 3.5 reports simple application of the methodology to the Cook Jonhson database and financial data. In the end, in section 3.6 we summarize our findings discussing their implications and extensions.

3.2 Copulas and Rank Reflection

The aim of this paper is the introduction of a new non parametric test for the study of the null hypothesis

$$H_0: \quad C(u_1, \ldots, u_D) = \bar{C}(u_1, \ldots, u_D) \tag{3.1}$$

where C is a copula and $\bar{C}$ is the corresponding survival copula, to be properly introduced in the following. This relationship, was first introduced, in the bivariate copula context, by [58] under the name of radial symmetry of the copula function. The reason, motivating the use of this name, was that this invariance property is linked to the so called radial symmetry of the bivariate distribution. We Instead follow [69], in calling it reflection symmetry because, in our opinion, the name radial symmetry is misleading in two ways. The first one is that the transformation under which the probability of the random variables is invariant is a reflection in a point(also called central inversion [23]), and this is true both for multivariate and copula symmetry. The second, most compelling, reason is that (3.1) has important probabilistic implications, that we try to point out, even without assuming symmetric marginals, needed together with (3.1) for the reflection symmetry of the joint distribution. Here, we will call a copula satisfying (3.1), normalized rank reflection symmetric,(rank reflection in brief), because under this property a multivariate

random variable and the multivariate random variable whose normalized ranks are the reflection trough the center of the Hypercube of the normalized rank of the original variable, have the same probability (see also [69]). In this section, after introducing copulas and survival copulas, we discuss normalized rank reflection symmetry, its probabilistic implication and which copulas and time series models manifest the symmetry.
We begin introducing the well known relationship among the cumulative marginal distribution function F_i of the i-th variable X_i and the marginal survival function $\bar{F}_i$ of the same variable:

$$\mathbb{P}\left(X_i > x_i\right) = \bar{F}_i\left(x_i\right) = 1 - \mathbb{P}\left(X_i \leq x_i\right) = 1 - F_i\left(x_i\right) \tag{3.2}$$

Applying the probability integral transforms $F_i\left(X_i\right) = U_i$ and $\bar{F}_i\left(X_i\right) = \bar{U}_i$ we can translate it to a relationship among uniform random variables

$$\bar{U}_i = 1 - U_i \tag{3.3}$$

.
The sample version of U_i represents the univariate ranks of the sample from X_i divided by the number of observations in the sample and, using (3.3), $\bar{U}_i$ can be interpreted as the one dimensional reflection around the center of the unitary interval of U_i. In this way the usual concept of symmetry of the distribution :

$$\begin{cases} F_i\left(c_i - x_i\right) & = & 1 - F_i\left(c_i + x_i\right) = \bar{F}_i\left(c_i + x_i\right) \\ F_i\left(c_i\right) & = & 1 - F_i\left(c_i\right) = \bar{F}_i\left(c_i\right) = \dfrac{1}{2} \end{cases} \tag{3.4}$$

can be interpreted(see the appendix) as a rank reflection symmetry in probability:

$$\mathbb{P}\left(U_i \leq u_i\right) = \mathbb{P}\left(\bar{U}_i \leq u_i\right) \tag{3.5}$$

whose extension to the multivariate setting can be translated in a relationship between the copula and its corresponding survival. Then, to introduce this extension, we need some results from copula theory.
According to the Sklar Theorem [79], a multivariate cumulative distribution could be expressed using the univariate marginal cumulative distributions and a copula function

$$\mathbb{P}\left(X_1 \leq x_1, \ldots, X_D \leq x_D\right) = F\left(x_1, \ldots, x_D\right) = C\left(F_1\left(x_1\right), \ldots, F_D\left(x_D\right)\right) \tag{3.6}$$

With the application of the probability integral transform to the original random variable we produce an uniform random vector $\mathbf{U} = \left(U_1, \ldots, U_D\right)^T$ with $U_i = F_i\left(X_i\right)$. Then, given the joint distribution and the marginals we can define a copula by

$$\mathbb{P}\left(U_1 \leq u_1, \ldots, U_D \leq x_D\right) = C\left(u_1, \ldots, u_D\right) = F\left(F_1^{-1}\left(u_1\right), \ldots, F_D^{-1}\left(u_D\right)\right) \tag{3.7}$$

Analogously, a multivariate survival function could be expressed using the univariate survival functions and the survival copula

$$\mathbb{P}\left(X_1 > x_1, \ldots, X_D > x_D\right) = \bar{F}\left(x_1, \ldots, x_D\right) = \bar{C}\left(\bar{F}_1\left(x_1\right), \ldots, \bar{F}_D\left(x_D\right)\right) \tag{3.8}$$

the survival copula is in this way a distribution function on the hypercube (not a survival function) of the random vector $\bar{\mathbf{U}} = \left(\bar{U}_1, \ldots, \bar{U}_D\right)^T$ with $\bar{U}_i = \bar{F}_i\left(X_i\right)$

$$\mathbb{P}\left(\bar{U}_1 \leq u_1, \ldots, \bar{U}_D \leq x_D\right) = \bar{C}\left(u_1, \ldots, u_D\right) = \bar{F}\left(\bar{F}_1^{-1}\left(u_1\right), \ldots, \bar{F}_D^{-1}\left(u_D\right)\right) \tag{3.9}$$

In this paper, we are concerned with the study of the null hypothesis of rank reflection symmetry, that is represented by the following generalization of (3.5):

$$\begin{aligned} \mathbb{P}\left(U_1 \leq u_1, \ldots, U_D \leq u_D\right) &= \mathbb{P}\left(\bar{U}_1 \leq u_1, \ldots, \bar{U}_D \leq u_D\right) & (3.10)\\ \Leftrightarrow C\left(u_1, \ldots, u_d\right) &= \bar{C}\left(u_1, \ldots, u_d\right) & (3.11) \end{aligned}$$

Since the sample version of $\mathbf{U}$ represents the vector of univariate ranks of the sample divided by the number of observations, and $\bar{\mathbf{U}}$ is $\mathbf{U}$ reflected in the center of the unit hypercube, we choose the name normalized rank reflection symmetry for this property and we, also, remark that this is only a necessary condition for radial symmetry of the multivariate distribution around a point, that must be supplemented by the symmetry of all the marginals.
In terms of probabilities, rank reflection symmetry correspond to:

$$\begin{aligned} &\mathbb{P}\left(X_1 \leq x_1, \ldots, X_D \leq x_D\right) = \mathbb{P}\left(X_1 > \bar{F}_1^{-1}\left(1 - \bar{F}_1\left(x_1\right)\right), \ldots, X_D > \bar{F}_D^{-1}\left(1 - \bar{F}_D\left(x_D\right)\right)\right) \\ \Leftrightarrow &\mathbb{P}\left(X_1 \leq x_1, \ldots, X_D \leq x_D\right) = \mathbb{P}\left(X_1 > F_1^{-1}\left(1 - F_1\left(x_1\right)\right), \ldots, X_D > F_D^{-1}\left(1 - F_D\left(x_D\right)\right)\right) \end{aligned}$$

To have a better understanding of (3.12), we compute the previous relation in the vector of marginal u-th quantiles $\left\{F_d^{-1}\left(u\right)\right\}_{d=1}^{D}$.

$$\mathbb{P}\left(X_1 \leq F_1^{-1}\left(q\right), \ldots, X_D \leq F_D^{-1}\left(u\right)\right) = \mathbb{P}\left(X_1 > F_1^{-1}\left(1 - u\right), \ldots, X_D > F_D^{-1}\left(1 - u\right)\right)$$

The probability of having all variable less than their respective u-th quantile is the same of the probability of having all the variables greater than the complementary marginal quantile. In this way, in particular, upper and lower joint extreme events are equiprobable. Introducing the multivariate generalization of upper and lower tail dependence coefficients [78],[50] allows us to express in another way the same concept :

$$\begin{aligned} \lambda_U &= \lim_{v \to 1} \frac{\bar{C}\left(1 - v, \cdots, 1 - v\right)}{1 - v} = \lim_{u \to 0} \frac{\bar{C}\left(u, \cdots, u\right)}{u} & (3.12)\\ \lambda_L &= \lim_{u \to 0} \frac{C\left(u, \cdots, u\right)}{u} & (3.13) \end{aligned}$$

From (3.12) and (3.13), we have that rank reflection symmetry implies the equality $\lambda_U = \lambda_L$ and this is, also, true for more refined measure of tail dependence as the one introduced in [74]. Given the above interpretation of rank reflection symmetry, testing this property could be evidence of asymmetric dependence an contagion as already discussed in the introduction.For a complete discussion of which parametric models are rank reflection symmetric in the two dimensional case, we refer to the seminal paper of Nelsen [58] (there rank reflection symmetry is named radial symmetry of the copula)and more the recent works [69] and [27]. We, only, stress here, that all elliptical copula are rank reflection symmetric and that Frank [32], explicitly constructed the only rank reflection symmetric bivariate Archimedean copula. For higher dimensions, to our knowledge, only elliptical copulas are known to exhibit this property. In time series context,it is well known, linear models with elliptic innovation are elliptically distributed and, a fortiori, exhibit rank reflection symmetry, because the class of elliptical distributions is closed under affine transformations. In addition, it can be direcly showed unsing the C-convolution contruction introduced in [19] and [18] that the reflection symmetry is preserved for linear process. Generalizing to the multivariate case Proposition 3.2 in [19], it is possible to obtain that, in case of symmetry of both the convolution and the cross-sectional copula the symmetry is preserved. Since in linear models the innovations are assumed to be independent from the variables, the convolution copula is the indipendence copula wich is reflection symmetric, this implies that the simmetry of the innovations is the same of the variables. For what concern multivariate GARCH and copula GARCH models, the covariance process, being quadratic, is invariant by a joint change of sign of innovations for each past time, i.e. for reflection of past innovations, so that the reflection symmetry of the normalized ranks of those models is, completely determined by the symmetry of the Copula joining the innovations.

3.3 Asymptotics with Empirical Processes

In this section, after introducing the empirical copula, the survival empirical and the related empirical processes, we discuss the dependent multiplier bootstrap as a way to approximate the asymptotic distribution and introduce our test statistic.
It can be shown [40],[18], that the multivariate survival copula can be expressed in terms of the corresponding copula through a complex generalizzation of the bivariate relationship:

$$\bar{C}\left(u,v\right) = u + v - 1 + C\left(1-u, 1-v\right) \tag{3.14}$$

This has been the usual way of testing this invariance property [8],[69],[27],[37]and [56]. That approach is difficult to generalize for a number of variables greater than 2. In this paper we propose an alternative route to test the null hypothesis (3.1) by making explicit use of Empirical survival copula process.

3.3.1 Empirical Copula Processes

Let $\{\{X_{id}\}_{i=1}^{n}\}_{d=1}^{D} \equiv \{\mathbf{X}_i\}_{i=1}^{n}$ a D dimensional multivariate sample of size n of strongly-mixing random variables. The empirical distribution function $\hat{F}\left(\mathbf{x}\right)$, the empirical survival function $\hat{\bar{F}}(\mathbf{x})$ and their marginals are

$$\hat{F}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{d=1}^{D}\mathbb{I}(X_{id} \le x_d)) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathbf{X_i} \le \mathbf{x}) \tag{3.15}$$

$$\hat{F}_d(x_d) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(X_{id} \le x_d) \tag{3.16}$$

$$\hat{\bar{F}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{d=1}^{D}\mathbb{I}(X_{id} > x_d) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(\mathbf{X_i} > \mathbf{x}) \tag{3.17}$$

$$\hat{\bar{F}}_d(x_d) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(X_{id} > x_d) \tag{3.18}$$

$$\tag{3.19}$$

Let us define the pseudo observations $\hat{U}_{id} = \hat{F}_d(X_{id})$ and $\hat{\bar{U}}_{id} = \hat{\bar{F}}_d(X_{id}) = 1 - \hat{U}_{id}$. Then, the empirical copula and the empirical survival copula are

$$\begin{aligned}\hat{C}_n(\mathbf{u}) = \hat{F}\left(\hat{F}_1^{-1}(u_1), \dots, \hat{F}_D^{-1}(u_D)\right) &= \frac{1}{n}\sum_{i=1}^{n}\prod_{d=1}^{D}\mathbb{I}\left(X_{id} \le \hat{F}_d^{-1}(u_d)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\left(\hat{\mathbf{U}}_i \le \mathbf{u}\right) \end{aligned}\tag{3.20}$$

$$\begin{aligned}\hat{\bar{C}}_n(\mathbf{u}) = \hat{\bar{F}}\left(\hat{\bar{F}}_1^{-1}(u_1), \dots, \hat{\bar{F}}_D^{-1}(u_D)\right) &= \frac{1}{n}\sum_{i=1}^{n}\prod_{d=1}^{D}\mathbb{I}\left(X_{id} > \hat{\bar{F}}_d^{-1}(u_d)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\left(\hat{\bar{\mathbf{U}}}_i \le \mathbf{u}\right) \end{aligned}\tag{3.21}$$

where we used non increasingness of the marginal empirical survivals.
The empirical processes are:

$$\hat{\mathbb{C}}_n = \sqrt{n}\left(\hat{C}_n(\mathbf{u}) - C(\mathbf{u})\right) \tag{3.22}$$

$$\hat{\bar{\mathbb{C}}}_n = \sqrt{n}\left(\hat{\bar{C}}_n(\mathbf{u}) - \bar{C}(\mathbf{u})\right) \tag{3.23}$$

An application of the functional Delta method [82] on results for central limit theorem (CTL) of multivariate empirical process for strongly mixing data with mixing coefficient $\alpha_n = o(n^{-a})$ for some $a > 0$,that can be found in [67], allow [13] to obtain the following weak convergence result for the empirical copula process

$$\hat{\mathbb{C}}_n \rightsquigarrow \mathbb{C} = \mathbb{B}_C(\mathbf{u}) - \sum_{d=1}^{D}\frac{\partial C(\mathbf{u})}{\partial u_d}\mathbb{B}_{d,C}(u_d) \tag{3.24}$$

where $\mathbb{B}_C$ is a D-dimensional Brownian sheet with covariance function

$$\mathbb{C}\mathrm{ov}\left(\mathbb{B}_C(\mathbf{u}), \mathbb{B}_C(\mathbf{v})\right) = C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})\,C(\mathbf{v}) \tag{3.25}$$

where $\wedge$ is the component-wise minimum.
We derive the weak convergence results for the empirical survival copula process, for strongly mixing data, under an assumption analogous to the one introduced in [76] (see the appendix for the proof):

A 9 *For each $j \in \{1,2,3\}$, the jth first-order partial derivative $\frac{\partial \bar{C}}{\partial u_j}$ exists and is continuous on the set $V_{D,j} := \left\{ u \in [0,1]^D : 0 < u_j < 1 \right\}$.*

proposition 4 *Suppose Conditions **A9** hold and the strongly mixing coefficients α_n of the sample are such that $\alpha_n = o(n^{-a})$ for some $a > 0$. Then the empirical survival copula process $\hat{\bar{\mathbb{C}}}_n = \sqrt{n}\left(\hat{\bar{C}}_n(\mathbf{u}) - \bar{C}(\mathbf{u})\right)$ weakly converges towards a Gaussian field $\bar{\mathbb{C}}$*

$$\hat{\bar{\mathbb{C}}}_n \rightsquigarrow \bar{\mathbb{C}} = \mathbb{B}_{\bar{C}}(\mathbf{u}) - \textstyle\sum_{d=1}^{D} \frac{\partial \bar{C}(\mathbf{u})}{\partial u_d} \mathbb{B}_{d,\bar{C}}(u_d) \qquad in \qquad \ell^\infty [0,1]^D$$

To our knowledge, the weak convergence of the empirical survival copula, under the stated assumptions, is new in literature, the only similar result is obtained for upper tail copula processes in [75] but under more restrictive assumptions on copula derivatives. We ,also, stress that the use of delta method could make the result valid for a variety of weakly dependent conditions for which the CTL on the multivariate empirical processes are known, as argued in [11]. The assumption of strongly mixing processes, assures the validity of the dependent multiplier bootstrap, to be introduced in the next section, known, only, for this kind of mixing condition.

3.3.2 Dependent Multiplier Bootstrap

The dependence of the limiting processes from the unknown copula and related survival Copula, through the covariance of $\mathbb{B}_C$ and $\mathbb{B}_{\bar{C}}$ and derivatives, forbid a closed form inference, based on the their distribution. The multiplier central limit theorem allows to obtain the distribution of the limiting process trough simulations. Here, we will use a recently introduced version of the multiplier central limit theorem valid for strictly stationary strongly mixing data [11]. Define a dependent multiplier sequence $\{\xi_{i,n}\}_{i\in\mathbb{Z}}$ i.e. a sequence that satisfies

1. The sequence $\{\xi_{i,n}\}_{i\in\mathbb{Z}}$ is strictly stationary with $\mathbb{E}(\xi_{0,n}) = 0$, $\mathbb{E}\left(\xi_{0,n}^2\right) = 1$ and $\mathbb{E}(|\xi_{0,n}|^\nu) < \infty$ for $\nu > 2$ and independent from the available sample.

2. There exists a sequence $\ell_n \to \infty$ of strictly positive constants such that $\ell_n = o(n)$ and the sequence $\{\xi_{i,n}\}_{i\in\mathbb{Z}}$ is ℓ_n-dependent i.e $\xi_{i,n}$ is independent from $\xi_{i+h,n}$ for all $h > \ell_n$ and $i \in \mathbb{N}$.

3. There exists a function $\phi : \mathbb{R} \to [0,1]$, symmetric around 0,continous at 0, satisfying $\phi(0) = 1$ and $\phi(x) = 0$ for all $|x| > 1$ such that $\mathbb{E}(\xi_{0,n}\xi_{h,n}) = \phi(h/\ell_n)$ for all $h \in \mathbb{Z}$.

Given M independent copies of the dependent multiplier sequence $\left\{\xi_{i,n}^{[1]}\right\}_{i\in\mathbb{Z}}, \ldots, \left\{\xi_{i,n}^{[M]}\right\}_{i\in\mathbb{Z}}$ we can define the new processes:

$$\tilde{\mathbb{B}}_n^{[m]}(\mathbf{u}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{i,n}^{(m)}\left(\mathbb{I}\left(\hat{\mathbf{U}}_i \le \mathbf{u}\right) - C(\mathbf{u})\right) \tag{3.26}$$

$$\tilde{\bar{\mathbb{B}}}_n^{[m]}(\mathbf{u}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_{i,n}^{(m)}\left(\mathbb{I}\left(\hat{\bar{\mathbf{U}}}_i \le \mathbf{u}\right) - \bar{C}(\mathbf{u})\right) \tag{3.27}$$

$$\tilde{\mathbb{C}}_n^{[m]}(\mathbf{u}) = \tilde{\mathbb{B}}_n^{[m]}(\mathbf{u}) - \sum_{d=1}^{D}\mathcal{D}_{u_d}^{FD}\hat{C}_n(\mathbf{u})\,\tilde{\mathbb{B}}_{d,n}^{[m]}(u_d) \tag{3.28}$$

$$\tilde{\bar{C}}_n^{[m]}(\mathbf{u}) = \tilde{\bar{\mathbb{B}}}_n^{[m]}(\mathbf{u}) - \sum_{d=1}^{D}\mathcal{D}_{u_d}^{FD}\hat{\bar{C}}_n(\mathbf{u})\,\tilde{\bar{\mathbb{B}}}_{d,n}^{[m]}(u_d) \tag{3.29}$$

Where $\mathcal{D}_{u_d}^{FD}$ is the partial finite difference derivative operator, that introduced for derivative estimation in [66], let estimated derivatives satisfy condition 4.1 of [11], originally in [76]. Proposition 4.2 in [11] implies that if $\ell_n = O\left(n^{\frac{1}{2}-\epsilon}\right)$ with $0 < \epsilon < \frac{1}{2}$ and the sample is drawn by a strictly stationary sequence with strongly mixing coefficients $\alpha(r) = O(r^{-a})$, $a = 3 + 3D/2$ we have

$$\left(\hat{\mathbb{C}}_n, \tilde{\mathbb{C}}_n^{[1]}, \ldots, \tilde{\mathbb{C}}_n^{[M]}\right) \rightsquigarrow \left(\mathbb{C}, \mathbb{C}^{[1]}, \ldots, \mathbb{C}^{[M]}\right) \tag{3.30}$$

$$\left(\hat{\bar{\mathbb{C}}}_n, \tilde{\bar{\mathbb{C}}}_n^{[1]}, \ldots, \tilde{\bar{\mathbb{C}}}_n^{[M]}\right) \rightsquigarrow \left(\bar{\mathbb{C}}, \bar{\mathbb{C}}^{[1]}, \ldots, \bar{\mathbb{C}}^{[M]}\right) \tag{3.31}$$

where $\mathbb{C}^{[1]}, \ldots, \mathbb{C}^{[M]}$ are M independent copies of $\mathbb{C}$ and $\bar{\mathbb{C}}^{[1]}, \ldots, \bar{\mathbb{C}}^{[M]}$ of $\bar{\mathbb{C}}$

3.3.3 Test Statistic

We want to test the null hypothesis

$$H_0: \quad C(u_1, \ldots, u_D) = \bar{C}(u_1, \ldots, u_D) \tag{3.32}$$

against the alternative

$$H_1: \quad C(u_1, \ldots, u_D) \neq \bar{C}(u_1, \ldots, u_D) \tag{3.33}$$

We choose a Cramer-Von Mises test statistic under the random measure generated by the empirical copula:

$$\mathbb{T}_n = \int_{(0,1]^D}\left(\hat{C}_n - \hat{\bar{C}}_n\right)^2 d\hat{C}_n = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{C}_n\left(\hat{\mathbf{U}}_i\right) - \hat{\bar{C}}_n\left(\hat{\mathbf{U}}_i\right)\right)^2 \tag{3.34}$$

The main motivation for using the random measure associated with the empirical copula, instead of using the the uniform measure, as for example in [27], is that, in the case of goodness of fit test statistics, the use of a Cramer-Von Mises statistic based on the

empirical copula measure leads to a more powerful test than competitors [39]. In [27] and [69], there is a study of the extremes of reflection asymmetry, in the bivariate case. Those results, allow for a proper normalizzation of measures of asymmetries proposed there. The equivalent multivariate results, for our measure, are not available and their derivation, although of great importance, are outside the scope of this paper and will be left for future research. This makes the range of our test statistic unknown, but does not alter the conclusion of our testing procedure.
Under the null we have:

$$n\mathbb{T}_n = \int_{(0,1]^D} \left(\hat{\mathbb{C}}_n - \hat{\bar{\mathbb{C}}}_n\right)^2 d\hat{C}_n \tag{3.35}$$

And we can construct multiplier copies

$$n\tilde{\mathbb{T}}_n^{[m]} = \int_{(0,1]^D} \left(\tilde{\mathbb{C}}_n^{[m]} - \tilde{\bar{\mathbb{C}}}_n^{[m]}\right)^2 d\hat{C}_n \tag{3.36}$$

In the following proposition we obtain the weak limits under the null

proposition 5 *If C is a normalized rank reflection symmetric copula i.e. $C = \bar{C}$ we have*

$$\left(n\mathbb{T}_n, n\tilde{\mathbb{T}}_n^{[1]}, \ldots, n\tilde{\mathbb{T}}_n^{[M]}\right) \rightsquigarrow \left(\mathbb{T}, \mathbb{T}^{[1]}, \ldots, \mathbb{T}^{[M]}\right) \tag{3.37}$$

$$\mathbb{T} = \int_{(0,1]^D} \left(\mathbb{C}_n - \bar{\mathbb{C}}_n\right)^2 dC \tag{3.38}$$

where $\mathbb{T}^{[1]}, \ldots, \mathbb{T}^{[M]}$ are independent copies of $\mathbb{T}$

It follows from proposition 5 that approximate P values for the tests of H_0 based on $\mathbb{T}_n$ are given by

$$\frac{1}{M}\sum_{m=1}^{M} \mathbb{I}\left(\tilde{\mathbb{T}}_n^{[m]} > \mathbb{T}_n\right) \tag{3.39}$$

Our inference procedure is robust, both under the null and under the alternative, for a data generating process whose dynamics can be modeled by strictly stationary strongly mixing processes, thanks to a CTL for strongly mixing empirical processes and the dependent multiplier bootstrap. Even if we are not aware of a data driven procedure for investigating the strongly mixing condition on real data we remark that the most used linear and non linear stationary time series models, usually satisfy this assumption [15]. The residual dependence, in addition, can be described through, the vast majority of copula parametric models, being the Segers condition on derivatives valid for them. For a clear exposition of this point and several examples we refer to the original paper[76]. The proposed, non parametric, statistical methodology, can, in this way,be applied jointly to any number of random variables that satisfy the assumption of most common parametric time series models and for this reason is suited for multivariate financial and economic datasets.

3.4 Simulation Study

In this section, we report the results for the empirical power and size of our non parametric test for different number of observations , for 2 and 10 dimensions. All the test use the dependent multiplier bootstrap of [11], in their moving average approach with the Bartlett kernel and are conducted at the 5% nominal level. We, slightly, modified the bandwidth selection procedure,integrating with respect to the empirical copula measure and not with respect to the uniform one, as in[11]. This choice is, once again, equivalent to averaging on pseudo observations, and we can avoid using an uniform grid for the approximation of the integrals. We postpone the description of the alghorithm used to the appendix B.3 and refer to [11], that introduced this methodology, for additional details. All the copula simulations are obtained by the use of the copula R package [53],with $D = 2, 10$, diferent number of observations N, $Ns = 1000$ simulations,$M = 2500$ multiplier replicates, Kendall's τ equal to $0.1, 0.3, 0.5, 0.7, 0.9$ and five different copula families, two elliptic: the Gaussian and the t with one degree of freedom, and three Archimedean: Frank Clayton and Gumbel. The elliptic copulas are radially symmetric, Gumbel and Clayton are not, the Frank Copula is radially symmetric only in two dimensions. Under Gaussian and t copula, which satisfy the null we are computing the percentage of times the test is rejecting the symmetry hypothesis when it is true, i.e. the size of the test, and this should be close to the 5% nominal level. When simulating under Clayton or Gumbel copula, that are not symmetric, we are computing the number of times the test is rejecting the null under the alternative, i.e. the power of the test. In the case of Frank copula we are computing the size in 2 dimensions and the power in 10 dimensions.
We use three different data generating processes.

3.4.1 i.i.d. Data Generating Process

The first one is the i.i.d. case, where we draw the $U_{i,d}$ with $i \in \{1, \ldots, n\}$ and $d \in \{1, \ldots, D\}$ from the choosen copula model C and put

$$X_{id} = U_{id} \tag{3.40}$$

In table 3.1 we show the results in case the simulated data are i.i.d. The size of the test is close to the nominal value of 5%, also for a moderate number of observation, but is lower for the Frank copula than for elliptical copulas and in general for high values of τ, and it is also lower in 10 dimensions. The empirical power increases, not only, with the number of observations, but also with the number of dimensions.

3.4.2 AR(1) Data Generating Process

For the second DGP again we draw $U_{i,d}$ from a copula C , but now with $i \in \{-100, \ldots, n\}$ and $d \in \{1, \ldots, D\}$. Then we impose Gaussian marginal innovations, applying Φ^{-1} the inverse of standard normal cumulative distribution, and an AR1 dynamic on the marginal

Table 3.1: Power/Size of the test i.i.d. case, Ns=1000 M=2500

D		2						10				
N		Size			Power			Size		Power		
250	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.055	0.077	0.042	0.237	0.055	0.100	0.011	0.018	0.992	0.503	0.949
	0.3	0.048	0.065	0.052	0.894	0.245	0.300	0.019	0.036	1.000	1.000	1.000
	0.5	0.050	0.061	0.048	0.998	0.431	0.500	0.028	0.019	1.000	1.000	1.000
	0.7	0.038	0.049	0.029	1.000	0.404	0.700	0.009	0.002	0.697	1.000	1.000
	0.9	0.015	0.016	0.016	1.000	0.137	0.900	0.000	0.000	1.000	0.000	0.000
500	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.046	0.050	0.042	0.396	0.097	0.1	0.023	0.033	1.000	0.987	1.000
	0.3	0.046	0.042	0.034	0.993	0.557	0.3	0.031	0.040	1.000	1.000	1.000
	0.5	0.049	0.042	0.034	1.000	0.781	0.5	0.032	0.028	1.000	1.000	1.000
	0.7	0.050	0.045	0.031	1.000	0.810	0.7	0.015	0.002	1.000	1.000	1.000
	0.9	0.014	0.028	0.011	1.000	0.498	0.9	0.000	0.000	1.000	0.435	0.000
1000	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.047	0.048	0.051	0.609	0.199	0.1	0.035	0.045	1.000	1.000	1.000
	0.3	0.048	0.052	0.048	1.000	0.877	0.3	0.039	0.047	1.000	1.000	1.000
	0.5	0.037	0.049	0.043	1.000	0.989	0.5	0.028	0.036	1.000	1.000	1.000
	0.7	0.043	0.053	0.040	1.000	0.993	0.7	0.028	0.020	1.000	1.000	1.000
	0.9	0.020	0.033	0.024	1.000	0.952	0.9	0.000	0.000	1.000	1.000	1.000

processes:

$$\begin{aligned}
\epsilon_{i,d} &= \Phi^{-1}\left(U_{i,d}\right) \\
X_{i,d} &= 0.5X_{i-1,d} + \epsilon_{i,d} \\
X_{-100,d} &= \epsilon_{-100,d}
\end{aligned}$$

and then, we discard the first 100 observations from the sample.
In table 3.2 we show the results the AR(1) copula DGP of [11]. All the remarks done for the independent case remain valid, but as can be seen the test is less powerful for dependent data requiring a greater number of observations to be reliable.

3.4.3 GARCH(1,1) Data Generating Process

The third choice of DGP is the same as the second one but now we impose a GARCH dynamics on the marginal processes.

$$\begin{aligned}
\epsilon_{i,d} &= \Phi^{-1}\left(U_{i,d}\right) \\
X_{i,d} &= h_{i,d}^{-\frac{1}{2}}\epsilon_{i,d} \\
h_{i,d} &= \omega + \alpha\epsilon_{i-1,d}^{2} + \beta h_{i-1} \\
h_{i-100} &= \frac{\omega}{1-\alpha-\beta}
\end{aligned}$$

and we discard the first 100 observations as before. The value of parameters used are $\alpha = 0.919$,$\beta = 0.072$,$\omega = 0.012$ as estimated from [51] for the S&P500 and already used in the simulations in [12] and [11].

Table 3.2: Power/Size of the test with marginal AR(1) dependence Ns=1000 M=2500

D		2					10					
N		Size			Power			Size		Power		
250	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.050	0.072	0.049	0.118	0.032	0.1	0.000	0.010	0.486	0.026	0.529
	0.3	0.047	0.053	0.042	0.396	0.080	0.3	0.013	0.021	0.986	0.793	0.992
	0.5	0.043	0.058	0.045	0.552	0.087	0.5	0.014	0.018	0.889	0.940	0.977
	0.7	0.025	0.034	0.018	0.640	0.066	0.7	0.004	0.001	0.104	0.686	0.444
	0.9	0.010	0.010	0.006	0.263	0.012	0.9	0.000	0.000	1.000	0.000	0.000
500	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.037	0.067	0.060	0.160	0.053	0.1	0.007	0.016	0.976	0.151	0.992
	0.3	0.040	0.071	0.035	0.604	0.125	0.3	0.011	0.026	1.000	1.000	1.000
	0.5	0.051	0.070	0.037	0.861	0.167	0.5	0.020	0.027	1.000	1.000	1.000
	0.7	0.032	0.032	0.023	0.934	0.144	0.7	0.006	0.004	0.921	1.000	0.998
	0.9	0.006	0.013	0.005	0.690	0.019	0.9	0.000	0.000	1.000	0.000	0.000
1000	τ	Gaussian	t	Frank	Clayton	Gumbel	τ	Gaussian	t	Frank	Clayton	Gumbel
	0.1	0.047	0.048	0.043	0.645	0.210	0.1	0.044	0.036	1.000	0.776	1.000
	0.3	0.056	0.055	0.047	1.000	0.891	0.3	0.034	0.038	1.000	1.000	1.000
	0.5	0.047	0.040	0.048	1.000	0.981	0.5	0.038	0.041	1.000	1.000	1.000
	0.7	0.047	0.051	0.043	1.000	0.993	0.7	0.018	0.028	1.000	1.000	1.000
	0.9	0.019	0.029	0.016	1.000	0.966	0.9	0.000	0.000	1.000	0.194	0.003

We report in a different tables results with innovations, linked by t copula with 1 and 5 degrees of freedom. Using 1 degree of freedom, so drawing from a Cauchy copula, it seems that the symmetry is changed by the GARCH dynamics. This is in contrast to what we remarked at the end section 3.2. This rather surprising result is in our opinion due to the fact that we the parameters used for the GARCH with those kind of innovations leads to a non stationary process (see [60] and [3]).
Table 3.3 reports results for marginal GARCH(1,1) processes. The greater persistence in those models leads to an additional power reduction and we need, still more observations for proper inference, in particular for Gumbel distributed innovations.
We report, in table 3.4, the results with t student innovations. In this case, it seems that GARCH dynamics alters the simmetry properties of the innovations, in fact, the number of times we reject the symmetry is, in the vast majority of the times, higher than the nominal level and this discrepance grows with the number of observations.

In table 3.4, the results with t student innovations with 1 and 5 degrees of freedom, shows instead that reducing the tail dependence to values, usually seen in financial data, [31] results in a better size for the test. We conjecture that this is impacting the stationarity of the GARCH process,for the actual values of parameters choosen ([60], [3]), bringing us outside the hypothesis of validity of our methodology, but futrher investigations of those issues are deferred to future research.

3.5 Data Applications

In this section, we report the application of our testing procedure to real data. Dataset have been choosen in order to have more than 1000 observation, an high number of variables , and for some of them temporal dependence. In this way, we hope to highlight, the advantages of using a non parametric, high dimensional and dependence robust, test of reflection symmetry. Bandwidth selection is done as in previous section and Pvalues are

Table 3.3: Power/Size of the test with GARCH(1,1) marginal dependence Ns=1000 NB=2500

D		2				10				
N		Size		Power			Size	Power		
250	τ	Gaussian	Frank	Clayton	Gumbel	τ	Gaussian	Frank	Clayton	Gumbel
	0.1	0.050	0.042	0.122	0.038	0.1	0.003	0.642	0.062	0.638
	0.3	0.056	0.048	0.425	0.084	0.3	0.011	0.997	0.946	0.998
	0.5	0.049	0.031	0.605	0.064	0.5	0.017	0.969	0.979	0.998
	0.7	0.031	0.033	0.697	0.061	0.7	0.006	0.315	0.801	0.699
	0.9	0.006	0.008	0.331	0.015	0.9	0.000	1.000	0.000	0.000
500	τ	Gaussian	Frank	Clayton	Gumbel	τ	Gaussian	Frank	Clayton	Gumbel
	0.1	0.062	0.053	0.173	0.052	0.1	0.008	0.997	0.342	0.997
	0.3	0.062	0.039	0.673	0.097	0.3	0.024	1.000	1.000	1.000
	0.5	0.055	0.046	0.895	0.140	0.5	0.033	1.000	1.000	1.000
	0.7	0.037	0.044	0.958	0.134	0.7	0.011	0.995	1.000	1.000
	0.9	0.007	0.018	0.818	0.033	0.9	0.000	1.000	0.000	0.000
1000	τ	Gaussian	Frank	Clayton	Gumbel	τ	Gaussian	Frank	Clayton	Gumbel
	0.1	0.055	0.053	0.253	0.058	0.1	0.016	1	0.918	1
	0.3	0.054	0.045	0.909	0.210	0.3	0.035	1	1	1
	0.5	0.064	0.058	0.996	0.320	0.5	0.049	1	1	1
	0.7	0.047	0.089	0.999	0.332	0.7	0.028	1	1	1
	0.9	0.016	0.047	0.999	0.141	0.9	0	1	0.274	0.09
1500	τ	Gaussian	Frank	Clayton	Gumbel	τ	Gaussian	Frank	Clayton	Gumbel
	0.1	0.045	0.057	0.371	0.1	0.1	0.024	1	1	1
	0.3	0.051	0.037	0.987	0.297	0.3	0.043	1	1	1
	0.5	0.066	0.054	1	0.493	0.5	0.076	1	1	1
	0.7	0.049	0.104	1	0.517	0.7	0.059	1	1	1
	0.9	0.016	0.072	1	0.269	0.9	0	1	1	0.998

Table 3.4: Test with GARCH(1,1) dependence and t-student innovations, Ns=1000 NB=2500

D		2	10
N			
250	τ	t	t
	0.1	0.095	0.030
	0.3	0.087	0.033
	0.5	0.088	0.028
	0.7	0.052	0.004
	0.9	0.017	0.000
500	τ		
	0.1	0.097	0.067
	0.3	0.092	0.095
	0.5	0.082	0.070
	0.7	0.061	0.023
	0.9	0.029	0.000
1000	τ		
	0.1	0.135	0.265
	0.3	0.130	0.277
	0.5	0.124	0.19
	0.7	0.079	0.076
	0.9	0.032	0
1500	τ		
	0.1	0.165	0.485
	0.3	0.176	0.456
	0.5	0.155	0.283
	0.7	0.104	0.163
	0.9	0.052	0

Table 3.5: Test with GARCH(1,1) dependence and t-student innovations, Ns=1000 NB=2500 with 1 and 5 degree of freedom i 2 dimensions

D.O.F		1	5
N			
250	τ		
	0.1	0.095	0.064
	0.3	0.087	0.058
	0.5	0.088	0.045
	0.7	0.052	0.040
	0.9	0.017	0.012
500	τ		
	0.1	0.097	0.058
	0.3	0.092	0.066
	0.5	0.082	0.069
	0.7	0.061	0.040
	0.9	0.029	0.006
1000	τ		
	0.1	0.135	0.063
	0.3	0.130	0.071
	0.5	0.124	0.062
	0.7	0.079	0.062
	0.9	0.032	0.025

approximated using the same number of multipliers $M = 2500$. As a first, simple illustration of the methodologies described , we apply the rank reflection test to the dataset introduced in [21],[22], as an example of non elliptically distributed multivariate data. This seven-dimensional data set contains the log-concentration of uranium (U), lithium (Li),cobalt (Co), potassium (K), cesium (Cs), scandium (Sc) and Titanium (Ti) measured in n = 655 data samples taken near Grand Junction, Colorado, US. Recently, the pairs of this data set have been tested for a possible modeling by certain selected Archimedean copula models (Ali–Mikhail–Haq, Clayton, Frank, Gumbel–Hougaard) in [38]. [4] and [65] carried out various tests on the same data set for the hypothesis that the pairs can be modeled by an arbitrary extreme-value copula. In table 3.6, we report the Pvalues for

Table 3.6: Pairwise and all variables test Pvalues for the Cook & Johnson Database

	Li	Co	K	Cs	Sc	Ti
U	0.709	0.178	0.064	0.286	0.469	0.168
Li		0.109	0.176	0.189	0.120	0.214
Co			0.008	0.010	0.572	0.241
K				0.076	0.000	0.034
Cs					0.001	0.000
Sc						0.010
All	0.238					

all pairwise rank reflection test and the pvalue for the test of rank reflection on the joint distribution of all variables. As can be seen we are not able to reject the rank reflection symmetry for most of the pairs with the exception of pairs with titanium and scandium. If we make a comparison with the goodness of fit test reported in [38], where the only rank reflection symmetric copula considered is the Frank copula, we see that we accept the null of radial symmetry for the couples (U,Li),(U,SC),(Li,Ti) and (Co,Ti) where the best model chosen by their test is the Frank Copula. For two out of four cases when the best model chosen by their procedure is not radially symmetric i.e. (Co,Cs)and (Cs,Sc) we reject the null at 5% level and for the other two cases i.e. (U,Co) and (Li,Sc), even if we are not able to reject the null at 5% we have moderately low p-value. We cannot, also , reject the rank reflection symmetry for the distribution of all the variables. Next, We consider financial data that are usually assumed to be dependent. First of all, we use the dataset of international financial indexes, introduced in [16], where he propose a moment based test in a dynamic copula framework to discriminate between different dynamic copula models. The Dataset includes log returns from six stock indexes taken from yahoo! finance: CAC40(France), The FTSE1000(UK),The Hang Seng Index (China) , Nikkei (Japan), S&P500(US),Russel2000(US) and TWSE(Taiwan) from July 7 1997 to December 30,2003.[1]
In table 3.7 we show the pvalues for the pairwise test of rank reflection. We have three rejection of rank reflection at 5% level. The first one is Russel2000-TWSE that is one of the

[1]In the original article [16] the sample was from January 1, 1995 to December 31, 2003 but we were not able to find in yahoo! finance the time series of TWSE preceding July 7 1997. We think this sample reduction cannot qualitatively change the results

four couples for which a Gaussian copula (rank reflection symmetric) is rejected in [16]. The second one S&P500-TWSE is one of the three couples for which in the same paper a Gumbel(not rank reflection symmetric) copula is accepted. The third case, HangSeng-S&P500, they accept the Survival Gumbel (not rank reflection symmetric). So at least for the rejected couples we are in agreement with the conclusion of [16]. As an additional

Table 3.7: Pairwise variables test Pvalues for the international Stock indexes returns used in [16]

		UK	China	Japan	US Broad	US Top	Taiwan
		FTSE	Hang Seng	Nikkei	Russel2000	S&P500	TWSE
France	CAC	0.723	0.053	0.055	0.846	0.554	0.207
UK	FTSE		0.561	0.252	0.962	0.783	0.668
China	HangSeng			0.297	0.058	0.009	0.690
Japan	Nikkei				0.255	0.111	0.452
US Broad	Russel2000					0.078	0.000
US Top	S&P500						0.001

check we conducted the test also for the selected terns of indexes they choose. Results are in table 3.8 together with the test on all the variables. The table shows three rejec-

Table 3.8: Selected Terns and all variables test Pvalues for the Stock index returns used in [16]

S&P500	Russel2000	FTSE100	0.657
S&P500	Russel2000	Nikkei	0.044
Russel2000	FTSE100	CAC40	0.866
Russel2000	FTSE100	Nikkei	0.110
Russel2000	FTSE100	TWSE	0.023
FTSE100	CAC40	Nikkei	0.075
FTSE100	CAC40	TWSE	0.294
CAC40	Nikkei	Hang Seng	0.020
CAC40	Nikkei	TWSE	0.092
		All	0.016

tion at 5% level, S&P500-Russel2000-Nikkei, Russel2000-FTSE100-TWSE and CAC40-Nikkei-HangSeng. Of those three only for S&P500-Russel2000-Nikkei some of the test for trivariate Gaussian copula reject the hypothesis in [16], so we have a partial agreement. We must stress that all their test with t copula accept the null and that they do not test in the trivariate setting for non elliptical copulas. In addition the test for all the variables reject the rank reflection symmetry so that overall our conclusion is the opposite of that reported in [**?**] and is in favor of asymmetric dependence in stock returns as reported by older studies [57], [1] and [49].
To further explore this issue, we apply the proposed test to log returns of stocks indexes of Belgium,France,Germany,Greece,Ireland,Italy,Netherlands,Portugal,Spain United Kingdom from January 2,2007 to May 20, 2014,during the past subprime and sovereign debt

crisis. The prices are total return index net of dividends, taken from Bloomberg. As can

Table 3.9: Pairwise and all variables test Pvalues for European Stock index returns

		Germany	UK	Netherlands	Belgium	Ireland	Greece	Spain	Italy	Portugal
		DAX	FSTE	AEX	BEL20	ISEQ	ASE	IBEX	FTSEMIB	PSI20
France	CAC	0.913	0.496	0.451	0.608	0.814	0.013	0.692	0.453	0.006
Germany	DAX		0.281	0.384	0.489	0.791	0.090	0.632	0.174	0.022
UK	FSTE			0.598	0.755	0.094	0.004	0.835	0.559	0.029
Netherlands	AEX				0.415	0.494	0.024	0.983	0.882	0.035
Belgium	BEL20					0.260	0.019	0.985	0.686	0.001
Ireland	ISEQ						0.063	0.413	0.496	0.448
Greece	ASE							0.019	0.011	0.083
Spain	IBEX								0.512	0.001
Italy	FTSEMIB									0.055
All	0.018									

be seen in 3.9, at 5% level, most of the couples with Greece and Portugal are not rank reflection symmetric and the test for all the variables is rejected. This is again in favor of dependence asymmetry and contagion in the euro area during the last crisis. To further investigate this finding, we computed our statistic from all the european index on rolling window of 1500 observations on a dayly basis, starting from 1/5/1998 to 5/20/2014. As

Figure 3.1: Daily Rolling Statistic

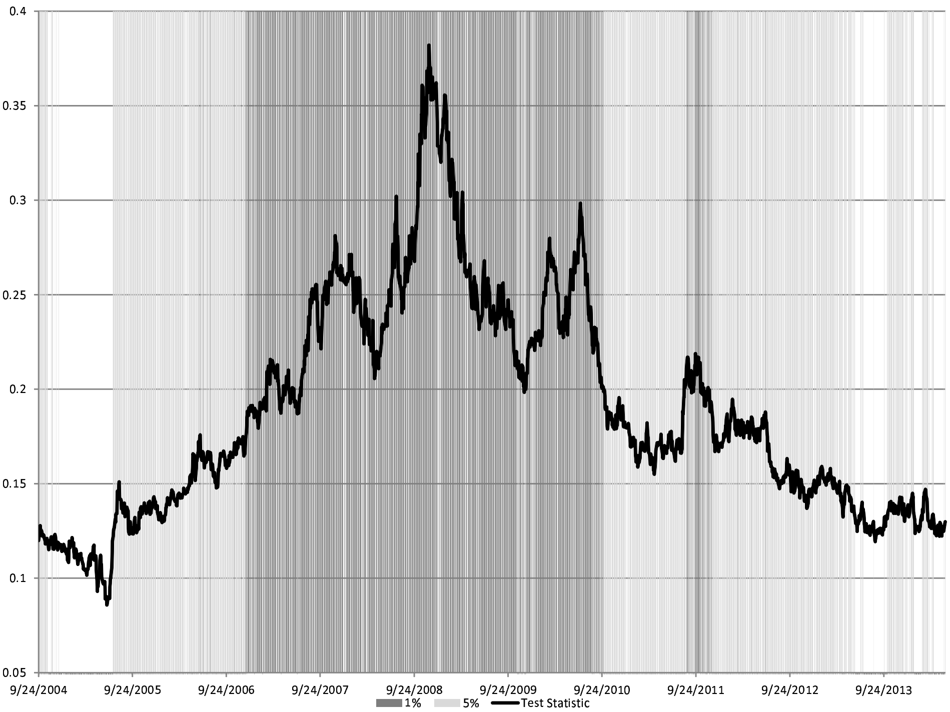

fig 3.1 shows for most of the sample data shows asymmetry at the 5% level and during the crisis period the test is significant 1% level. In addition, the statistic has a smooth behaviour, increasing inthe first part of the sample peaking in the end of 2008 after the AIG bailout and other FED intervention measure and then decrease. It, then, reincreases,

again, during the sovereign debt crisis, in 2010 and, to a lesser extent, in the second part of 2011. Those preliminar results, seems to suggest that a properly normalized asymmetry measure could be used as a stress indicator for contagion in financial markets. Further investigation of this possibility, on different time periods and datasets, are left for future research.

3.6 Conclusions

In this article we propose a test for normalized rank reflection symmetry. This property of copula functions, was known under the name of radial symmetry. We motivate our choice of name, by distinguishing it from the stronger radial symmetry property of multivariate distribution, giving a probabilistic interpretation and some of its consequences. Our test is easier to extend to more than two dimensions than previously proposed test of this type, in addition the use of probabilistic result for the CLT of multivariate empirical processes for strongly mixing data and of a new multiplier bootstrap procedure allows the application of the test to dependent data, covering most of known stationary parametric models. This is particularly important in light of the application of the test, we have done, to financial contagion issues. Our extensive simulation study showed that the test is enough powerful with a moderate number of observation and that increasing the dimension make the test more powerful. Dependence in the data lower the power, but with more than 1000 observations the test appears reliable for most of the DGP used. We must remark that the use of the dependent bootstrap in case of i.e. data come at the cost of losing power with respect to the simple multiplier bootstrap [11]. Even with this cost in mind, in our opinion this approach is favorable to the usual one that tend to apply the test to the residuals of some parametric model. In our application to financial time series, we showed that our test could be used to detect dependence asymmetries. Our application to European Stock indexes hints, also, to a link to financial contagion, being able to isolate Greece and Portugal as two of the most contagious countries during the sovereign debt crisis. This connection with contagion and the consequent possibility of using the test statistic as stress indicator for contagion is, partly, explored computing our statistic on a rolling window and will explored further in future research, given the encouraging behaviour showed in fig. 3.1. We hope ,that the importance of the normalized rank reflection symmetry, in correctingly evaluating, the risk of the joint extreme events and the reliability of our test procedure for dependent high dimensional data, showed in our simulation study and our empirical applications, will make this test widespread and used.

Chapter 4

Discrete Orthogonal Polynomials Derivatives for Empirical Copula process

4.1 Introduction

The empirical copula process has important statistical applications in non parametric testing. Independence [26], goodness of fit [39] ,radial symmetry [8], global symmetry, associativity, archimedianicity , max-stability ([56] and reference therein) are only some examples of test procedures proposed in the literature. The underlying assumption for those inference procedures are usually weaker than most of the non copula based methods and usually are more powerful with respect to the alternative test based on the multivariate empirical process.
With the exception of independence and goodness of fit tests, in all the other cases, the true copula is unknown, even under the null hypothesis, and inference procedures should be based on bootstrap. Although several approaches has been proposed for bootstrapping from the asymptotic distribution of the process, the most reliable is the multiplier bootstrap approach with derivative estimation, even if it requires the estimation of the unknown copula derivatives [10]. In the literature, the only approach to copula derivatives estimation is through central two point finite differences :

$$\mathcal{D}_x^{FD} f(x) \equiv \frac{(f(x+h) - f(x-h))}{2h} = \frac{\partial f(x)}{\partial x} + o(h) \tag{4.1}$$

with step h proportional to the square root of the number of observations N. In this paper, instead, we propose the use of discrete orthogonal polynomials derivatives, proposed in[28], as a generalization of Savitzky-Golay method [72]. The main theoretical advantages of this method are the possibility of considering more than two point in the approximation, and the possibility of weighting, through the orthogonality measure, different points in different ways. Those theoretical advantages will be checked through simulation on copula derivatives estimation through the empirical copula and multiplier bootstrap method on radial symmetry test.
The paper is structured as follows: In Section 4.2 we summarize recent findings of [28], derive the convergence in probability of our proposed estimator and outline the algorithm

for its computation. Then in Section 4.3 we evaluate the efficiency of the proposed method through an extensive simulation study, discuss results and suggest possible extension of our efforts.

4.2 Orthogonal polynomials Derivatives and Empirical copula process

In this section we recall the definition of orthogonal polynomials and introduce the orthogonal polynomials derivative. Then, after introducing the copula, the empirical copula and the empirical copula process, we will state in proposition 6 that the Diekema derivative applied on the empirical copula converge in probability to the copula derivative, and give details of the algorithmic procedure to estimate the derivative on data.
Given a positive Borel measure on $\mathbb{R}$,$\mu(x)$ such that $\int_{\mathbb{R}} |x|^n \, d\mu(x)$, with $n \geq 0$, orthogonal polynomials with respect to μ are defined as the set polynomials $p_i \equiv q_c = \sum_{k=0}^{\infty} c_k x^k = \sum_{k=0}^{i} c_k x^k$ of degree exactly i s.t.

$$\langle p_i, p_j \rangle = \int_{\mathbb{R}} p_i(x)\, p_j(x)\, d\mu(x) = 0 \qquad i \neq j \tag{4.2}$$

If the measure has finite discrete support $d\mu(x) = \sum_{k=1}^{K} \delta(x - x_k)\, w_k dx$, with $\delta(x - x_k)$ the point mass Dirac delta centered in x_k, we can use sums instead of integrals:

$$\langle p_i, p_j \rangle = \int_{\mathbb{R}} p_i(x)\, p_j(x)\, d\mu(x) = \sum_{k=1}^{K} p_i(x_k)\, p_j(x_k)\, w_k = 0 \qquad i \neq j \tag{4.3}$$

and, in this case, there are only K orthogonal polynomials.
The orthonormal polynomials used in this paper satisfy the three term recurrence relation:

$$\sqrt{\beta_{j+1}} p_{j+1}(x) = (x - \alpha_j)\, p_j(x) - \sqrt{\beta_j} p_{j-1}(x), \qquad p_{-1}(x) = 0, \qquad p_0(x) = \frac{1}{\sqrt{\beta_0}}.$$

Analogous recurrences for polynomial derivatives can be obtained by deriving both sides of the three terms recurrence.
Using orthonormal polynomials and their derivatives it is possible to define an m-degree polynomial approximation for the first derivative $\partial_x f(x)$ and prove, under the existence of m-th derivative at x, the following [28]:

$$\mathcal{D}^{OP}_{x,m} f(x) \equiv \sum_{j=1}^{m} \frac{1}{h} \left.\frac{\partial p_j}{\partial x}\right|_{x=0} \int_{\mathbb{R}} f(x + h\xi)\, p_j(\xi)\, d\mu(\xi) = \frac{\partial f(x)}{\partial x} + o\left(h^{m-1}\right) \tag{4.4}$$

We, now, recall some definition and results concerning the empirical copula. Given D random variables $\{X_i\}_{i=1}^{D}$, with marginals $\mathbb{P}(X_i \leq x_i) = F_i(x_i)$, $i = 1, \ldots, D$ and joint cumulative distribution $\mathbb{P}(X_1 \leq x_1, \ldots, X_D \leq x_D) = F(x_1, \ldots, x_D)$ by Sklar' s theorem in D dimensions [59] we know it exists a copula function $C : [0,1]^D \mapsto [0,1]$ such that $F(x_1, \ldots, x_D) = C(F_1(x_1), \ldots, F_D(x_D))$.

Let $(X_{11}, \ldots, X_{N1}) \ldots (X_{1N}, \ldots, X_{DN})$ be a random sample, distributed according to F, the empirical distribution function and its margins are

$$\hat{F}_N(x_1, \ldots, x_D) = \frac{1}{N} \sum_{j=1}^{N} \prod_{i=1}^{D} \mathbb{I}(X_{ij} \leq x_i) \tag{4.5}$$

$$\hat{F}_{Ni}(x_i) = \frac{1}{N} \sum_{j=1}^{N} \mathbb{I}(X_{ij} \leq x_i) \tag{4.6}$$

The empirical copula is :

$$\hat{C}_N(u_1, \ldots, u_D) = \frac{1}{N} \sum_{j=1}^{N} \prod_{i=1}^{D} \mathbb{I}\left(\hat{U}_{Nij} \leq u_i\right) \tag{4.7}$$

with the pseudo-observations given by $\hat{U}_{Nij} = \hat{F}_{Ni}(X_{ij})$. Several authors ([76] and references therein) have studied the weak convergence of the associated empirical process

$$\hat{\mathbb{C}}_N(u_1, \ldots, u_3) = \sqrt{N}\left(\hat{C}_N(u_1, \ldots, u_D) - C(u_1, \ldots, u_D)\right) \tag{4.8}$$

, and under nonrestrictive assumptions on copula derivatives they prove for $N \to \infty$:

$$\hat{\mathbb{C}}_N(u_1, \ldots, u_D) \rightsquigarrow \alpha_C(u_1, \ldots, u_D) + \sum_{i=1}^{D} \beta_{iC}(u_i)\, \partial_{u_i} C(u_1, \ldots, u_D) \tag{4.9}$$

where $\rightsquigarrow$ denotes weak convergence and $\alpha_C(u_1, \ldots, u_D)$ is a C-Brownian Bridge on $[0,1]^D$ and $\beta_{iC}(u_i)$ are its margins. From the last expression it is evident how a good test on copula characteristics depends on good estimations of derivatives and this becomes the more important the higher is the copula dimension D.
We state our main proposition on orthogonal polynomials derivative estimator:

proposition 6 *Assuming $h = N^{-\frac{1}{2}}$,let μ_N be a sequence of measure with limiting measure μ such that for each N $supp(\mu_N) \subset [a, b]$ with $a, b \in \mathbb{R}$ not dependent on N and a constant $B \in \mathbb{R}$, under the existence of $\frac{\partial^m C}{\partial u_i^m}$ on the set $V_{D,i} \equiv \left\{u \in [0,1]^D : 0 < u_i < 1\right\}$, we have:*

$$\left|\mathcal{D}^{OP}_{u_i,m} \hat{C}_N\right| \leq B \qquad \sup_{u \in V_{D,i}} \left|\mathcal{D}^{OP}_{u_i,m} \hat{C}_N - \partial_{u_i} C\right| \xrightarrow{\mathbb{P}} 0 \qquad i, u_1, \ldots, u_D \tag{4.10}$$

The hypothesis on copula derivatives for $m > 1$ are slightly more restrictive than in [76] but needed for the validity of (4.4) and satisfied by the copula families and values of m used in the following. The use of this proposition in conjunction with proposition 3.2 in[76] justify the use of multiplier based inference procedures adopted in the second part of the simulation study.
We, now, introduce our choice of discrete orthogonal measure. The use of a discrete measure is, mainly, motivated by numerical methods for obtaining orthogonal polynomials.

Apart from classical orthogonal polynomials, for which everything is known, methods for determination of recurrence coefficients in (4.4) from their orthogonality measure are computationally feasible only for measures with discrete supports, measures over a continuum, usually, are discretized and then those methods are applied [34]. In this paper the recurrence coefficients are obtained, using the matlab implementation of Lanczos method described in [34].
First of all to satisfy the hypothesis of the previous proposition 6, and to compare directly with finite differences (4.1) we choose the point of the support such that $-1 \leq \xi_k \leq 1$. In this way finite difference can be interpreted as $\xi_k \in \{-1, 0, 1\}$ and constant orthogonality measure. Instead we consider a subsample of points equally spaced with space $N^{-1/2}$ so that the empirical copula is computed on points spaced by N^{-1} and this is the minimum spacing that results in different value of this function. In particular since it is well known that higher order polynomial approximation in presence of noise are subject to the Runge phenomenon, that could greatly deteriorate the approximation, we choose the subsample that better resembles Chebycev points according to the mock Chebycev scheme proposed in [9]. In this paper we report only results obtained by the orthogonality measure characterized by the weights w_k:

$$\begin{aligned} \tilde{w}_k &= \left| \hat{C}_N \left(\ldots, u_i + \xi_k h, \ldots \right) - \hat{C}_N \left(\ldots, u_i, \ldots \right) \right| + \frac{1}{N} \\ w_k &= \frac{\tilde{w}_k}{\sum_{k'=-\lfloor \sqrt{N} \rfloor}^{\lfloor \sqrt{N} \rfloor} \tilde{w}_{k'}} \end{aligned} \tag{4.11}$$

In the previous expression, the first term on the right hand side weights more points with an empirical copula value more different from the empirical copula value at which the derivative is computed and the second term is introduced to avoid numerical issues associated with zero measure point and is set to the minimum value of the first term different form zero. This measure is a first heuristic attempt to discriminate noise induced variation of empirical copula values from true ones. We stress that our proposition 6 is valid for a wide range of continous and discrete orthogonality measures and that the investigation of different measures from the one used in this paper could lead to even better gain in the reliability of test procedures.

4.3 Simulation Study

In this section we report the results from the simulation study comparing finite difference derivatives with our proposed orthogonal polynomials derivatives. All simulations are obtained by the use of the copula R package [48],with $D = 2$, $N = 100$, 1000 simulations, Kendall's τ equal to $0.1, 0.5, 0.9$ and four different copula families, two elliptic: the Gaussian and the t with one degree of freedom, and three Archimedean: Frank Clayton and Gumbel. The first comparison is done computing the difference between true copula derivative $\frac{\partial C(u,v)}{\partial u}$ and two derivatives approximations (4.1),(4.4) applied empirical copula, that we, with an abuse of notation, denote together by $\mathcal{D}_u \hat{C}_N(u,v)$, using the L^2

distance based on the empirical copula measure:

$$\left\| \frac{\partial C}{\partial u} - \mathcal{D}_u \hat{C}_N \right\|_{2,\hat{C}_N} = \sqrt{\int_{[0,1]^2} \left(\frac{\partial C(u,v)}{\partial u} - \mathcal{D}_u \hat{C}_N(u,v) \right)^2 d\hat{C}_N(u,v)} \tag{4.12}$$

The resulting measure is then averaged over the simulations.
For the second comparison we report the size/power of the reflection symmetry test based

Table 4.1: Average over 1000 simulations of $\left\| \frac{\partial C}{\partial u} - \mathcal{D}_u \hat{C}_N \right\|_{2,\hat{C}_N}$ with $D = 2$, $N = 100$

		Gaussian		t		Frank		Clayton		Gumbel	
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.1	1	0.1046	0.1008	0.1086	0.1050	0.1079	0.1053	0.1082	0.1082	0.1051	0.1032
	2	0.1046	0.1025	0.1086	0.1065	0.1079	0.1053	0.1082	0.1082	0.1051	0.1032
	3	0.1046	0.1550	0.1086	0.1574	0.1079	0.1503	0.1082	0.1702	0.1051	0.1504
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.5	1	0.1098	0.1084	0.1216	0.1200	0.1060	0.1122	0.1251	0.1200	0.1090	0.1064
	2	0.1098	0.1074	0.1216	0.1183	0.1060	0.1122	0.1251	0.1200	0.1090	0.1064
	3	0.1098	0.1588	0.1216	0.1626	0.1060	0.1122	0.1251	0.1523	0.1090	0.1484
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.9	1	0.1540	0.1594	0.1907	0.1977	0.1529	0.1406	0.1726	0.1612	0.1582	0.1473
	2	0.1540	0.1442	0.1907	0.1811	0.1529	0.1406	0.1726	0.1612	0.1582	0.1473
	3	0.1540	0.1589	0.1907	0.1856	0.1529	0.1317	0.1726	0.1436	0.1582	0.1364

on the $\mathbb{A}_{5,n}$ statistic of [8] computed using $B = 1000$ multiplier bootstrap replicates. The elliptic copulas and the Frank copula are reflection symmetric, Gumbel and Clayton are not. From Table 4.1, we can see that OP derivatives generally outperforms FD for at

Table 4.2: Reflection symmetry test size/power $D = 2$, $N = 100$, $B = 1000$,1000 simulations

		Size						Power			
		Gaussian		t		Frank		Clayton		Gumbel	
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.1	1	0.048	0.046	0.065	0.067	0.045	0.044	0.100	0.098	0.06	0.058
	2	0.048	0.048	0.065	0.067	0.045	0.044	0.100	0.099	0.06	0.058
	3	0.048	0.026	0.065	0.04	0.045	0.025	0.100	0.064	0.06	0.036
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.5	1	0.043	0.04	0.05	0.051	0.043	0.041	0.775	0.775	0.208	0.199
	2	0.043	0.042	0.05	0.049	0.043	0.044	0.775	0.776	0.208	0.201
	3	0.043	0.021	0.05	0.021	0.043	0.022	0.775	0.682	0.208	0.125
τ	m	FD	OP	FD	OP	FD	OP	FD	OP	FD	OP
0.9	1	0.011	0.006	0.03	0.012	0.018	0.008	0.638	0.369	0.09	0.057
	2	0.011	0.013	0.03	0.031	0.018	0.019	0.638	0.672	0.09	0.097
	3	0.011	0	0.03	0.002	0.018	0.001	0.638	0.136	0.09	0.003

least one choice of m, the only exception being the Frank copula with $\tau = 0.5$. For what concern the choice of m Archimedean copulas gave the same results for $m = 1, 2$

and that outperform $m = 3$ with low and medium association but do worst for $\tau = 0.9$. Considering also the elliptic copulas makes $m = 2$ the most convenient choice. Table 4.2 confirms that $m = 2$ is the best choice, but the OP derivative outperform FD only for $\tau = 0.9$. However differences in size/power are really small across all the table and further investigation in terms of different values of N,D and different test statistics are needed to better understand the impact of different approximations on inference procedures.

4.4 Concluding Remarks

The use of empirical copula based test is becoming, quickly, an active part of multivariate statistic. For this reason, better ways of evaluating the limiting asymptotic process are becoming, incresingly, important. Since the best known method is the multiplier method with derivative estimation, we tried to contribute proposing a new estimator, based on orthogonal polynomials.In addition, recent methodologies [45] are directly based on conditional copulas for which the estimator is directly an estimator of copula derivatives and in those cases is even more important to have a reliable derivative estimator.In this paper we were able to show that, at least for high association, that is the case where the use of copulas matter the most, our proposed orthogonal polynomials derivative do better than the finite difference benchmark, and given the possibility of using different orthogonality measures, the application of orthogonal polynomials derivatives to the empirical copula is an idea worth of further exploration.

Appendix A

Appendix of Chapter 2

A.1 Derivatives

In the following, we introduce our definition of finite difference derivatives and then show they satisfy the requirements to be copula consistent derivatives. This is done to show that the simplest possible estimator of derivatives can be used for constructing the projected empirical copula process, if the true copula satisfies **A3** and **A4**.
Slightly adapting our definitions from [76] we have

$$\begin{aligned}
\partial_i C(\mathbf{u}) &= \limsup_{h\downarrow 0}\left\{\frac{C(\mathbf{u}+\mathbf{e}_i 2h)}{2h}\mathbb{I}(u_i=0)\right.\\
&+ \frac{C(\mathbf{u}+\mathbf{e}_i h)-C(\mathbf{u}-\mathbf{e}_i h)}{2h}\mathbb{I}(0<u_i<1)\\
&+ \left.\frac{C(\mathbf{u})-C(\mathbf{u}-\mathbf{e}_i 2h)}{2h}\mathbb{I}(u_i=1)\right\} \qquad \text{(A.1)}
\end{aligned}$$

where $\mathbf{u}=(u_1,\ldots,u_D)$ and $\mathbf{e}_i$ is the i-th D-dimensional basis vector.
In this way, it is easy to obtain the finite difference approximation used in [36]:

$$\begin{aligned}
\mathcal{D}^{(1)}_{N,i}C(\mathbf{u}) &= \left\{\frac{C(\mathbf{u}+\mathbf{e}_i(2h-u_i))}{2h}\mathbb{I}(u_i\le h)\right.\\
&+ \frac{C(\mathbf{u}+\mathbf{e}_i h)-C(\mathbf{u}-\mathbf{e}_i h)}{2h}\mathbb{I}(h<u_i<1-h)\\
&+ \left.\frac{C(\mathbf{u}-\mathbf{e}_i(u_i-1))-C(\mathbf{u}-\mathbf{e}_i(u_i-(1-2h)))}{2h}\mathbb{I}(u_i\ge 1-h)\right\}
\end{aligned}$$

If we define forward, backward and central differences as

$$\Delta_i^{2h}f(\mathbf{u}) = \frac{f(\mathbf{u}+\mathbf{e}_i 2h)-f(\mathbf{u})}{2h} \qquad \text{(A.2)}$$

$$\nabla_i^{2h}f(\mathbf{u}) = \frac{f(\mathbf{u})-f(\mathbf{u}-\mathbf{e}_i 2h)}{2h} \qquad \text{(A.3)}$$

$$\delta_i^{2h}f(\mathbf{u}) = \frac{f(\mathbf{u}+\mathbf{e}_i h)-f(\mathbf{u}-\mathbf{e}_i h)}{2h} \qquad \text{(A.4)}$$

(A.1) could be rewritten

$$\begin{aligned}\partial_i C(\mathbf{u}) &= \limsup_{h\downarrow 0}\left\{\frac{\Delta_i^{2h}C(\mathbf{u})}{2h}\mathbb{I}(u_i=0)+\frac{\delta_i^{2h}C(\mathbf{u})}{2h}\mathbb{I}(0<u_i<1)\right.\\ &+\left.\frac{\nabla_i^{2h}C(\mathbf{u})}{2h}\mathbb{I}(u_i=1)\right\} \qquad (A.5)\end{aligned}$$

(A.5) can be generalized to:

$$\begin{aligned}\partial_i^{(n)} C(\mathbf{u}) &= \limsup_{h\downarrow 0}\left\{\frac{\left(\Delta_i^{2h}\right)^n C(\mathbf{u})}{(2h)^n}\mathbb{I}(u_i=0)+\frac{\left(\delta_i^{2h}\right)^n C(\mathbf{u})}{(2h)^n}\mathbb{I}(0<u_i<1)\right.\\ &+\left.\frac{\left(\nabla_i^{2h}\right)^n C(\mathbf{u})}{(2h)^n}\mathbb{I}(u_i=1)\right\} \qquad (A.6)\end{aligned}$$

From this expression the approximation is:

$$\begin{aligned}\mathcal{D}_{N,i}^{(n)} C(\mathbf{u}-\mathbf{e}_i u_i) &= \left\{\frac{\left(\Delta_i^{2h}\right)^n C(\mathbf{u}-\mathbf{e}_i u_i)}{(2h)^n}\mathbb{I}(u_i\le nh)+\frac{\left(\delta_i^{2h}\right)^n C(\mathbf{u})}{(2h)^n}\mathbb{I}(nh<u_i<1-nh)\right.\\ &+\left.\frac{\left(\nabla_i^{2h}\right)^n C(\mathbf{u}-\mathbf{e}_i(u_i-1))}{(2h)^n}\mathbb{I}(u_i\ge 1-nh)\right\} \qquad (A.7)\end{aligned}$$

A.1.1 Finite difference Approximation satisfies A5 and A6

In this paragraph, we will show that finite difference approximations are uniform approximations of the derivatives for derivable functions (i.e. assumption **A5**) and that are a consistent estimators for copula derivatives (i.e. assumption **A6**)

Let us start with **A5**, it is well known that for n times differentiable functions G

$$\left(\frac{\left(\Delta_i^{2h}\right)^n G(\mathbf{u})}{(2h)^n}-\partial_i^{(n)}G(\mathbf{u})\right)=o(1)$$

$$\left(\frac{\left(\nabla_i^{2h}\right)^n G(\mathbf{u})}{(2h)^n}-\partial_i^{(n)}G(\mathbf{u})\right)=o(1)$$

$$\left(\frac{\left(\delta_i^{2h}\right)^n G(\mathbf{u})}{(2h)^n}-\partial_i^{(n)}G(\mathbf{u})\right)=o(1)$$

thus

$$\begin{aligned}\left(\mathcal{D}_{N,i}^{(n)}C(\mathbf{u})-\partial_i^{(n)}C(\mathbf{u})\right) &= \qquad (A.8)\\ &= \left(o(1)+\partial_i^{(n)}C(\mathbf{u}-\mathbf{e}_i u_i)-\partial_i^{(n)}C(\mathbf{u})\right)\mathbb{I}(u_3\le nh)\\ &+ o(1)\,\mathbb{I}(nh<u_3<1-nh)\\ &+ o(1)\left(\partial_i^{(n)}C(\mathbf{u}-\mathbf{e}_i(u_i-1))-\partial_i^{(n)}C(\mathbf{u})\right)\mathbb{I}(u_3>1-nh)\end{aligned}$$

When $u_i \in \{0,1\}$ the derivative difference (A.8) is 0 otherwise is $o(1)$ by the continuity in u_i, so in the end the difference between the approximation and the derivative is $o(1)$. Then, using the previous result, **A6** follows:

$$
\begin{aligned}
& \left(\mathcal{D}_{N,i}^{(n)} \hat{C}_N(\mathbf{u}) - \partial_i^{(n)} C(\mathbf{u})\right) \\
= & \left(\mathcal{D}_{N,i}^{(n)} \hat{C}_N(\mathbf{u}) - \mathcal{D}_{N,i}^{(n)} C(\mathbf{u})\right) + o(1) \\
= & \frac{1}{\sqrt{N}} \mathcal{D}_{N,i}^{(n)} \hat{\mathbb{C}}_N(\mathbf{u}) + o(1) \\
= & \left\{ \frac{\left(\Delta_i^{2h}\right)^n \hat{\mathbb{C}}_N(\mathbf{u} - \mathbf{e}_i u_i)}{\sqrt{N}(2h)^n} \mathbb{I}(u_i \le nh) \right. \\
+ & \frac{\left(\delta_i^{2h}\right)^n \hat{\mathbb{C}}_N(\mathbf{u})}{\sqrt{N}(2h)^n} \mathbb{I}(nh < u_i < 1 - nh) \\
+ & \left. \frac{\left(\nabla_i^{2h}\right)^n \hat{\mathbb{C}}_N(\mathbf{u} - \mathbf{e}_i(u_i - 1))}{\sqrt{N}(2h)^n} \mathbb{I}(u_i \ge 1 - nh) \right\} + o(1) \\
= & \\
= & \left\{ \frac{\sum_{j=1}^n (-1)^j \binom{n}{j} \hat{\mathbb{C}}_N(\mathbf{u} - \mathbf{e}_i(u_i + (n-i)2h))}{\sqrt{N}(2h)^n} \mathbb{I}(u_i \le nh) \right. \\
+ & \frac{\sum_{j=1}^n (-1)^j \binom{n}{j} \hat{\mathbb{C}}_N\left(\mathbf{u} - \mathbf{e}_i\left(\frac{n}{2} - j\right)2h\right)}{\sqrt{N}(2h)^n} \mathbb{I}(nh < u_i < 1 - nh) \\
+ & \left. \frac{\sum_{j=1}^n (-1)^j \binom{n}{j} \hat{\mathbb{C}}_N(\mathbf{u} - \mathbf{e}_i(u_i - 1 - j2h))}{\sqrt{N}(2h)^n} \mathbb{I}(u_i \ge 1 - nh) \right\} + o(1)
\end{aligned}
$$

Since $\sum_{j=1}^n (-1)^j \binom{n}{j} = 0$,if $h^n = O\left(\frac{1}{\sqrt{N}}\right)$, it goes to zero in probability by the continuity of the paths of $\mathbb{C}$

A.1.2 Finite Difference Approximations satisfies A8

Finite difference approximation satisfies also the integrated difference rate of convergence condition **A8**.

Under hypothesis **A3** and **A4** on copula derivatives needed for theorem 1, by Taylor expansion we can show:

$$\left(\frac{\left(\Delta_i^{2h}\right)C\left(\mathbf{u}\right)}{(2h)} - \partial_i C\left(\mathbf{u}\right)\right) = O\left(h\right)$$
$$\left(\frac{\left(\nabla_i^{2h}\right)C\left(\mathbf{u}\right)}{(2h)} - \partial_i C\left(\mathbf{u}\right)\right) = O\left(h\right)$$
$$\left(\frac{\left(\delta_i^{2h}\right)C\left(\mathbf{u}\right)}{(2h)} - \partial_i C\left(\mathbf{u}\right)\right) = o\left(h\right)$$

Then we have

$$\begin{aligned}
& \sqrt{N}\int_0^{u_3} dv_3 \left(\mathcal{D}_{N,3}^{(1)}C\left(u_1,u_2,v_3\right) - \partial_3 C\left(u_1,u_2,v_3\right)\right) \\
= \; & \sqrt{N}\int_0^{u_3\wedge h} dv_3 O\left(h\right) + \mathbb{I}\left(u_3 \geq 1-h\right)\sqrt{N}\int_{1-h}^{u_3} dv_3 O\left(h\right) + \sqrt{N}o\left(h\right) \\
= \; & \sqrt{N}u_3 \wedge hO\left(h\right) + \mathbb{I}\left(u_3 \geq 1-h\right)\sqrt{N}\left(\left(1-h\right)-u_3\right)O\left(h\right) + \sqrt{N}o\left(h\right) \\
= \; & \sqrt{N}o\left(h\right) = o\left(1\right)
\end{aligned}$$

A.1.3 Finite Difference Approximations Allows A7

The most challenging requirement is asymptotic integration by part given in the assumption **A7**. We, now, show that finite difference approximation are asymptotically integrable by part.

Let $F(u_1,u_2,u_3) = \int_0^{u_3} f(u_1,u_2,v_3)\,dv_3 + k(u_1,u_2)$, then:

$$
\begin{aligned}
& \int_0^{u_3} dv_3 f(u'_1,u'_2,v_3)\,\mathcal{D}^{(1)}_{N,3}\hat{\mathbb{C}}_N(u_1,u_2,v_3) \\
+ & \int_0^{u_3} dv_3 \mathcal{D}^{(1)}_{N,3} f(u'_1,u'_2,v_3)\,\hat{\mathbb{C}}_N(u_1,u_2,v_3) \\
- & f(u'_1,u'_2,u_3)\,\hat{\mathbb{C}}_N(u_1,u_2,u_3) \\
= & \int_0^{u_3\wedge h} dv_3 f(u'_1,u'_2,v_3)\,\frac{\hat{\mathbb{C}}_N(u_1,u_2,2h)}{2h} \\
+ & \int_h^{u_3\wedge(1-h)} dv_3 f(u'_1,u'_2,v_3)\,\frac{\hat{\mathbb{C}}_N(u_1,u_2,v_3+h)-\hat{\mathbb{C}}_N(u_1,u_2,v_3-h)}{2h} \\
+ & \mathbb{I}(u_3 \geq (1-h)) \int_{(1-h)}^{u_3} dv_3 f(u'_1,u'_2,v_3)\,\frac{\hat{\mathbb{C}}_N(u_1,u_2,1)-\hat{\mathbb{C}}_N(u_1,u_2,1-h)}{2h} \\
+ & \int_0^{u_3\wedge h} dv_3 \frac{f(u'_1,u'_2,2h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) \\
+ & \int_h^{u_3\wedge(1-h)} dv_3 \frac{f(u'_1,u'_2,v_3+h)-f(u'_1,u'_2,v_3-h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) \\
+ & \mathbb{I}(u_3 \geq (1-h)) \int_{(1-h)}^{u_3} dv_3 \frac{f(u'_1,u'_2,1)-f(u'_1,u'_2,1-h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) \\
- & f(u'_1,u'_2,u_3)\,\hat{\mathbb{C}}_N(u_1,u_2,u_3)
\end{aligned}
$$

$$
\begin{aligned}
= & \hat{\mathbb{C}}_N(u_1,u_2,2h)\,\frac{F(u'_1,u'_2,u_3\wedge h)-F(u'_1,u'_2,0)}{u_3\wedge h}\frac{u_3\wedge h}{2h} && \text{(A.9)} \\
+ & \int_0^{u_3\wedge(1-h)-h} dv_3 f(u'_1,u'_2,v_3-h)\,\frac{\hat{\mathbb{C}}_N(u_1,u_2,v_3)}{2h} && \text{(A.10)} \\
- & \int_{2h}^{u_3\wedge(1-h)+h} dv_3 f(u'_1,u'_2,v_3+h)\,\frac{\hat{\mathbb{C}}_N(u_1,u_2,v_3)}{2h} && \text{(A.11)} \\
+ & \left(\hat{\mathbb{C}}_N(u_1,u_2,1)-\hat{\mathbb{C}}_N(u_1,u_2,1-h)\right)\mathbb{I}(u_3 \geq (1-h)) \times \\
\times & \frac{f(u'_1,u'_2,(1-h))-f(u'_1,u'_2,u_3)}{(1-h)-u_3}\frac{(1-h)-u_3}{2h} && \text{(A.12)} \\
+ & \int_0^{u_3\wedge h} dv_3 \frac{f(u'_1,u'_2,2h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) && \text{(A.13)} \\
+ & \int_h^{u_3\wedge(1-h)} dv_3 \frac{f(u'_1,u'_2,v_3+h)-f(u'_1,u'_2,v_3-h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) && \text{(A.14)} \\
+ & \mathbb{I}(u_3 \geq (1-h)) \int_{(1-h)}^{u_3} dv_3 \frac{f(u'_1,u'_2,1)-f(u'_1,u'_2,1-h)}{2h}\hat{\mathbb{C}}_N(u_1,u_2,v_3) && \text{(A.15)} \\
- & f(u'_1,u'_2,u_3)\,\hat{\mathbb{C}}_N(u_1,u_2,u_3) && \text{(A.16)}
\end{aligned}
$$

When $N \to \infty$ we have that

$$\begin{aligned}
\frac{u_3 \wedge h}{2h} &\to \frac{1}{2} \\
\frac{F(u_1', u_2', u_3 \wedge h) - F(u_1', u_2', 0)}{u_3 \wedge h} &\to f(u_1', u_2', 0) \\
\hat{\mathbb{C}}_N(1, u_2, 2h) &\rightsquigarrow \mathbb{C}(1, u_2, 0) = 0
\end{aligned}$$

where the last term follows from the continuity of the sample paths of $\mathbb{C}$. So (A.9) goes to zero.
Analogously (A.12) goes to zero since:

$$\begin{aligned}
\mathbb{I}(u_3 \geq 1-h) &\to \mathbb{I}(u_3 = 1) \\
\frac{u_3 - 1 + h}{2h}\mathbb{I}(u_3 \geq 1-h) &\to \frac{1}{2}\mathbb{I}(u_3 = 1) \\
\frac{F(u_1', u_2', u_3) - F(u_1', u_2', 1-h)}{u_3 - 1 + h}\mathbb{I}(u_3 \geq 1-h) &\to f(u_1', u_2', 1) \\
\hat{\mathbb{C}}_N(u_1, u_2, 1) &\rightsquigarrow \mathbb{C}(1, u_2, 1) \\
\hat{\mathbb{C}}_N(u_1, u_2, 1-2h) &\rightsquigarrow \mathbb{C}(u_1, u_2, 1)
\end{aligned}$$

For (A.13) and (A.15) we recall a rough bound

$$\left|\hat{\mathbb{C}}_N\right| \leq \sqrt{N}\left(\left|\hat{C}_N\right| + |C|\right) \leq 2\sqrt{N}$$

so that

$$\begin{aligned}
&\left|\frac{f(u_1', u_2', 2h) - f(u_1', u_2', 0)}{2h}\int_0^{u_3 \wedge h} dv_3 \hat{\mathbb{C}}_N(u_1, u_2, v_3)\right| \\
\leq\ & 2\left|f(u_1', u_2', 2h) - f(u_1', u_2', 0)\right| \frac{\sqrt{N}}{2h} u_3 \wedge h \\
&\left|\mathbb{I}(u_3 \geq (1-h)) \frac{f(u_1', u_2', 1) - f(u_1', u_2', 1-h)}{2h}\int_{(1-h)}^{u_3} \hat{\mathbb{C}}_N(u_1, u_2, v_3)\, dv_3\right| \\
\leq\ & \left|f(u_1', u_2', 1) - f(u_1', u_2', 1-h)\right| \frac{\sqrt{N}}{2h} 2\left((1-h) - u_3\right)\mathbb{I}(u_3 \geq (1-h))
\end{aligned}$$

By the continuity of f the limit is zero in both cases.
If we sum (A.10),(A.11) and (A.14) we obtain

$$\begin{aligned}
&\int_{u_3 \wedge (1-h) - h}^{u_3 \wedge (1-h)} dv_3 f(u_1', u_2', v_3 - h) \frac{\hat{\mathbb{C}}_N(u_1, u_2, v_3)}{2h} \\
-\ & \int_h^{2h} dv_3 f(u_1', u_2', v_3 - h) \frac{\hat{\mathbb{C}}_N(u_1, u_2, v_3)}{2h} \\
+\ & \int_{u_3 \wedge (1-h)}^{u_3 \wedge (1-h) + h} dv_3 f(u_1', u_2', v_3 + h) \frac{\hat{\mathbb{C}}_N(u_1, u_2, v_3)}{2h} \\
-\ & \int_0^h dv_3 f(u_1', u_2', v_3 + h) \frac{\hat{\mathbb{C}}_N(u_1, u_2, v_3)}{2h} \qquad \text{(A.17)}
\end{aligned}$$

All those terms can be written as particular instances of the following integral

$$
\begin{aligned}
& \frac{1}{k}\int_a^{a+k} dv_3 f\left(u_1',u_2',v_3\right)\mathbb{C}_N\left(u_1,u_2,v_3\right) \\
= & \frac{1}{k}\left[\sqrt{N}\int_a^{a+k} dv_3 f\left(u_1',u_2',v_3\right)C\left(u_1,u_2,v_3\right)\right. \\
- & \left.\sqrt{N}\int_a^{a+k} dv_3 f\left(u_1',u_2',v_3\right)\hat{C}_N\left(u_1,u_2,v_3\right)\right]
\end{aligned}
$$

Since both C and $\hat{C}_N$ are bounded monotonic non decreasing non negative function, and f is a bounded integrable function, we can use the second mean value theorem [42]. For some $\eta, \eta' \in [0,k]$ we have:

$$
\begin{aligned}
= & \frac{1}{k}\left[\sqrt{N}C\left(u_1,u_2,a+k\right)\int_{a+\eta}^{a+k} dv_3 f\left(u_1',u_2',v_3\right)\right. \\
- & \left.\sqrt{N}\hat{C}_N\left(u_1,u_2,a+k\right)\int_{a+\eta'}^{a+k} dv_3 f\left(u_1',u_2',v_3\right)\right] \\
= & \frac{1}{k}\left[\hat{\mathbb{C}}_N\left(u_1,u_2,a+k\right)\int_a^{a+k} dv_3 f\left(u_1',u_2',v_3-k\right)\right. \\
- & \sqrt{N}C\left(u_1,u_2,a+k\right)\int_a^{a+\eta} dv_3 f\left(u_1',u_2',v_3\right) \\
+ & \left.\sqrt{N}\hat{C}_N\left(u_1,u_2,a+k\right)\int_a^{a+\eta'} dv_3 f\left(u_1',u_2',v_3\right)\right] \\
= & \frac{1}{k}\left[\hat{\mathbb{C}}_N\left(u_1,u_2,a+k\right)k\left[f\left(u_1',u_2',a\right)+O\left(k\right)\right]\right. \\
- & \sqrt{N}C\left(u_1,u_2,a+k\right)\eta\left[f\left(u_1',u_2',a\right)+O\left(\eta\right)\right] \\
+ & \left.\sqrt{N}\hat{C}_N\left(u_1,u_2,a+k\right)\eta'\left[f\left(u_1',u_2',a\right)+O\left(\eta'\right)\right]\right] \\
= & \hat{\mathbb{C}}_N\left(u_1,u_2,a+k\right)\left[f\left(u_1',u_2',a\right)+O\left(k\right)\right]
\end{aligned}
$$

Applying this result to (A.17), subtracting (A.14) we get a zero limit by the continuity of the paths of $\mathbb{C}$, so that **A7** is verified.

A.1.4 Explicit Expressions

In this section, we compute explicit expression for the Empirical Projected Copula and a Cramer Von Mises statistics for testing conditional independence, when we use finite difference derivative estimators

$$\begin{aligned}
& \hat{\Pi}_{N|3}\left(\hat{C}_N(u_1,u_2,u_3)\right) \\
= & \int_0^{u_3} dv_3 \mathcal{D}^{(1)}_{N,3}\hat{C}_N(u_1,1,v_3)\,\mathcal{D}^{(1)}_{N,3}\hat{C}_N(1,u_2,v_3) \\
= & \sum_{i,j=1}^{N} \mathbb{I}(U_{1i}\le u_1)\,\mathbb{I}(U_{2j}\le u_2) \\
& \left\{\frac{1}{(2Nh)^2}\mathbb{I}(U_{3i}\le 2h)\,\mathbb{I}(U_{3j}\le 2h)\int_0^{u_3}\mathbb{I}(v_3<h)\,dv_3\right. \\
+ & \frac{1}{(2Nh)^2}\int_0^{u_3}\mathbb{I}(U_{3i}-h\le v_3<U_{3i}+h)\,\mathbb{I}(U_{3j}-h\le v_3<U_{3j}+h)\,\mathbb{I}(h\le v_3<1-h)\,dv_3 \\
+ & \left.\frac{1}{(2Nh)^2}\mathbb{I}(1-2h\le U_{3i}<1)\,\mathbb{I}(1-2h\le U_{3i}<1)\int_0^{u_3}\mathbb{I}(v_3\ge 1-h)\,dv_3\right\}
\end{aligned}$$

$$\begin{aligned}
= & \sum_{i,j=1}^{N}\mathbb{I}(U_{1t}\le u_1)\,\mathbb{I}(U_{2s}\le u_2) \\
& \left\{\frac{1}{(2Nh)^2}\mathbb{I}(U_{3i}\le 2h)\,\mathbb{I}(U_{3j}\le 2h)\,u_3\wedge h\right. \\
+ & \frac{1}{(2Nh)^2}\left[u_3\wedge(1-h)\wedge(U_{3i}\wedge U_{3j}+h)-h\vee(U_{3i}\vee U_{3j}-h)\right]\vee 0 \\
+ & \left.\frac{1}{(2Nh)^2}\mathbb{I}(1-2h\le U_{3i})\,\mathbb{I}(1-2h\le U_{3i})\,(u_3-1+h)\wedge 0\right\}
\end{aligned}$$

$$= \sum_{i,j=1}^{N}\mathbb{I}(U_{1i}\le u_1)\,\mathbb{I}(U_{2j}\le u_2)\,\mathbb{K}(u_3,U_{3i},U_{3j},h)$$

We have used the following integral of a variable $v\in[0,1]$

$$\int_0^u dv\prod_{i=1}^{m}\mathbb{I}(a_i\le v<b_i) = [1\wedge u\wedge b_1\wedge\ldots\wedge b_m - 0\vee a_1\vee\ldots\vee a_m]\vee 0$$

The Cramer Von Mises statistic is:

$$\begin{aligned}
\int_{[0,1]^3} d^3u\left(\hat{C}_N-\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right)^2 &= \int_{[0,1]^3} d^3u\left(\hat{C}_N\right)^2 \\
- \ 2\int_{[0,1]^3} d^3u\left(\hat{C}_T\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right) &+ \int_{[0,1]^3} d^3u\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right)^2
\end{aligned}$$

$$\begin{aligned}
\int_{[0,1]^3} d^3u\left(\hat{C}_N\right)^2 &= \frac{1}{N^2}\sum_{i,j=1}^{N}\prod_{d=1}^{3}\int_0^1 du_d\mathbb{I}(U_{di}\le u_d)\,\mathbb{I}(U_{dj}\le u_d) \\
&= \frac{1}{N^2}\sum_{i,j=1}^{N}\prod_{d=1}^{3}[1-U_{di}\vee U_{dj}]
\end{aligned}$$

$$
\begin{aligned}
&\int_{[0,1]^3} d^3u \left(\hat{C}_N \hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right) \\
= &\ \frac{1}{N}\sum_{k,i,j=1}^{N} \int_0^1 \mathbb{I}\left(U_{1i} \le u_1\right)\mathbb{I}\left(U_{1k} \le u_1\right) du_1 \int_0^1 \mathbb{I}\left(U_{2k} \le u_2\right)\mathbb{I}\left(U_{2j} \le u_2\right) du_2 \\
&\int_0^1 \mathbb{I}\left(U_{3k} \le u_3\right)\left\{\frac{1}{(2Nh)^2}\mathbb{I}\left(U_{3i} \le 2h\right)\mathbb{I}\left(U_{3j} \le 2h\right) u_3 \wedge h\right. \\
+ &\ \frac{1}{(2Nh)^2}\left[u_3 \wedge (1-h) \wedge \left(U_{3i} \wedge U_{3j} + h\right) - h \vee \left(U_{3i} \vee U_{3j} - h\right)\right] \wedge 0 \\
+ &\ \left.\frac{1}{(2Nh)^2}\mathbb{I}\left(1-2h \le U_{3i}\right)\mathbb{I}\left(1-2h \le U_{3i}\right)\left(u_3 - 1 + h\right) \wedge 0\right\} \\
= &\ \frac{1}{N}\sum_{k,i,j=1}^{N} \left[1 - U_{1t} \vee U_{1r}\right]\left[1 - U_{2r} \vee U_{2s}\right] \\
&\int_0^1 \mathbb{I}\left(U_{3k} \le u_3\right)\left\{\frac{1}{(2Nh)^2}\mathbb{I}\left(U_{3i} \le 2h\right)\mathbb{I}\left(U_{3j} \le 2h\right) u_3 \wedge h\right. \\
+ &\ \frac{1}{(2Nh)^2}\left[u_3 \wedge (1-h) \wedge \left(U_{3i} \wedge U_{3j} + h\right) - h \vee \left(U_{3i} \vee U_{3j} - h\right)\right] \wedge 0 \\
+ &\ \left.\frac{1}{(2Nh)^2}\mathbb{I}\left(1-2h \le U_{3i}\right)\mathbb{I}\left(1-2h \le U_{3i}\right)\left(u_3 - 1 + h\right) \wedge 0\right\}
\end{aligned}
$$

So let's consider integral of the type

$$
\begin{aligned}
k(a,b,c,d) &= \int_0^1 \mathbb{I}\left(a \le v < b\right)\left[v \wedge c - d\right] \vee 0 \\
&= \int_0^1 \mathbb{I}\left(a \le v < b\right)\mathbb{I}\left(v \le c\right)\mathbb{I}\left(v - d \ge 0\right)\left[v - d\right] dv \\
&+ \int_0^1 \mathbb{I}\left(a \le v < b\right)\mathbb{I}\left(v > c\right)\mathbb{I}\left(c - d \ge 0\right)\left[c - d\right] dv
\end{aligned}
$$

$$
\begin{aligned}
= &\ \int_0^1 \mathbb{I}\left(a \vee d \le v < b \wedge c\right)\left[v - d\right] dv \\
+ &\ 0 \vee \left[b - c \vee a\right]\left[c - d\right] \vee 0 \\
= &\ \frac{\left[(b \wedge c)^2 - (a \vee d)^2\right] \vee 0}{2} - d\left[b \wedge c - a \vee d\right] \vee 0 \\
+ &\ 0 \vee \left[b - c \vee a\right]\left[c - d\right] \vee 0
\end{aligned}
$$

We have

$$\int_{[0,1]^3} d^3u \left(\hat{C}_N \hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right) = \frac{1}{N}\sum_{k,i,j=1}^{N} \left[1 - U_{1t} \vee U_{1r}\right]\left[1 - U_{2r} \vee U_{2s}\right]$$
$$\left\{\mathbb{I}\left(U_{3i} \le 2h\right)\mathbb{I}\left(U_{3j} \le 2h\right)\frac{k\left(U_{3k}, 1, h, 0\right)}{(2Nh)^2}\right.$$
$$+ \frac{k\left(U_{3k}, 1, (1-h) \wedge \left(U_{3i} \wedge U_{3j} + h\right), h \vee \left(U_{3i} \vee U_{3j} - h\right)\right)}{(2Nh)^2}$$
$$\left. + \frac{\mathbb{I}\left(1 - 2h \le U_{3i}\right)\mathbb{I}\left(1 - 2h \le U_{3i}\right)k\left(U_{3k}, 1, 0, (1-h)\right)}{(2Nh)^2}\right\}$$

Finally

$$\int_{[0,1]^3} d^3u \left(\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right)$$
$$= \sum_{l,k,i,j=1}^{N} \left[1 - U_{1i} \vee U_{1l}\right]\left[1 - U_{2k} \vee U_{2j}\right]$$
$$\int_0^1 \left\{\frac{1}{(2Nh)^2}\mathbb{I}\left(U_{3k} \le 2h\right)\mathbb{I}\left(U_{3l} \le 2h\right)u_3 \wedge h\right.$$
$$+ \frac{1}{(2Nh)^2}\left[u_3 \wedge (1-h) \wedge \left(U_{3k} \wedge U_{3l} + h\right) - h \vee \left(U_{3k} \vee U_{3l} - h\right)\right] \wedge 0$$
$$\left. + \frac{1}{(2Nh)^2}\mathbb{I}\left(1 - 2h \le U_{3k}\right)\mathbb{I}\left(1 - 2h \le U_{3k}\right)\left(u_3 - 1 + h\right) \wedge 0\right\}$$
$$\left\{\frac{1}{(2Nh)^2}\mathbb{I}\left(U_{3i} \le 2h\right)\mathbb{I}\left(U_{3j} \le 2h\right)u_3 \wedge h\right.$$
$$+ \frac{1}{(2Nh)^2}\left[u_3 \wedge (1-h) \wedge \left(U_{3i} \wedge U_{3j} + h\right) - h \vee \left(U_{3i} \vee U_{3j} - h\right)\right] \wedge 0$$
$$\left. + \frac{1}{(2Nh)^2}\mathbb{I}\left(1 - 2h \le U_{3i}\right)\mathbb{I}\left(1 - 2h \le U_{3i}\right)\left(u_3 - 1 + h\right) \wedge 0\right\} du_3$$

let's define

$$\begin{aligned} b_{lm} &= (1-h) \wedge \left(U_{3l} \wedge U_{3m} + h\right) \\ a_{lm} &= h \vee \left(U_{3l} \vee U_{3m} - h\right) \end{aligned} \tag{A.18}$$

The integrals are all of the following type

$$\begin{aligned}
k^2(a,b,c,d) &= \int_0^1 0\vee[v\wedge b-a][v\wedge c-d]\vee 0dv \\
&= \int_0^1 0\vee[v\wedge b-a]\,\mathbb{I}(v>c)[c-d]\vee 0dv \\
&+ \int_0^1 0\vee[v\wedge c-d]\,\mathbb{I}(v>b)[b-a]\vee 0dv \\
&+ \int_0^1 \mathbb{I}(v\le c)\,\mathbb{I}(v\le b)\,\mathbb{I}(v-a\ge 0)\,\mathbb{I}(v-d\ge 0)[v-a][v-d]\,dv \\
&= k(c,1,b,a)[c-d]\vee 0+k(b,1,c,d)[b-a]\vee 0 \\
&+ \mathbb{I}(a\vee d<b\wedge c)\int_{a\vee d}^{b\wedge c}[v-a][v-d]\,dv \\
&= k(c,1,b,a)[c-d]\vee 0+k(b,1,c,d)[b-a]\vee 0 \\
&+ \mathbb{I}(a\vee d<b\wedge c)\left\{\frac{1}{3}\left[(b\wedge c)^3-(a\vee d)^3\right]\right. \\
&- \left.(a+d)\frac{1}{2}\left[(b\wedge c)^2-(a\vee d)^2\right]+ad\left[(b\wedge c)-(a\vee d)\right]\right\}
\end{aligned}$$

$$\begin{aligned}
&\int_{[0,1]^3} d^3u\left(\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\hat{\Pi}_{N|3}\left(\hat{C}_N\right)\right) \\
=\ & \frac{1}{(2Nh)^4}\sum_{l,k,i,j=1}^{N}[1-U_{1t}\vee U_{1q}][1-U_{2r}\vee U_{2s}]\{ \\
& \mathbb{I}(U_{3k}\vee U_{3l}\le 2h)\left[\mathbb{I}(U_{3i}\vee U_{3j}\le 2h)\,k^2(0,h,0,h)\right. \\
+\ & k^2(0,h,a_{ts},b_{ts}) \\
+\ & \left.\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(0,h,1-h,0)\right] \\
+\ & \left[\mathbb{I}(U_{3i}\vee U_{3j}\le 2h)\,k^2(a_{rq},b_{rq},0,h)\right. \\
+\ & k^2(a_{rq},b_{rq},a_{ts},b_{ts}) \\
+\ & \left.\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(a_{rq},b_{rq},1-h,0)\right] \\
+\ & \mathbb{I}(U_{3k}\wedge U_{3l}\ge 1-2h)\left[\mathbb{I}(U_{3i}\vee U_{3j}\le 2h)\,k^2(1-h,0,0,h)\right. \\
+\ & k^2(1-h,0,a_{ts},b_{ts}) \\
+\ & \left.\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(1-h,0,1-h,0)\right]\} \\
=\ & \frac{1}{(2Nh)^4}\sum_{l,k,i,j=1}^{N}[1-U_{1t}\vee U_{1q}][1-U_{2r}\vee U_{2s}]\{ \\
& \mathbb{I}(U_{3k}\vee U_{3l}\le 2h)\,\mathbb{I}(U_{3i}\vee U_{3j}\le 2h)\,k^2(0,h,0,h) \\
+\ & 2\mathbb{I}(U_{3k}\vee U_{3l}\le 2h)\,k^2(0,h,a_{ts},b_{ts}) \\
+\ & 2\mathbb{I}(U_{3k}\vee U_{3l}\le 2h)\,\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(0,h,1-h,0) \\
+\ & k^2(a_{rq},b_{rq},a_{ts},b_{ts}) \\
+\ & 2\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(a_{rq},b_{rq},1-h,0) \\
+\ & \mathbb{I}(U_{3k}\wedge U_{3l}\ge 1-2h)\,\mathbb{I}(U_{3i}\wedge U_{3j}\ge 1-2h)\,k^2(1-h,0,1-h,0)\}
\end{aligned}$$

A.2 Granger Causality and Conditional Independence

This short section is intended to show in an informal way the relationship beetwin Conditional independence as expressed in (2.1) and the lag 1 Granger non causality. Applying the definition of Granger non causality given in [43] for 1-lag we have:

$$\mathbb{P}\left(X_t \leq x_t \left| X_{t-1} = x_{t-1}, Y_{t-1} = y_{t-1}\right.\right) = \mathbb{P}\left(X_t \leq x_t \left| X_{t-1} \leq x_{t-1}\right.\right) \tag{A.19}$$

As noted in [30] a necessary condition for Granger non causality is a conditional independence requirement of the type of (2.1).
If $X_t \perp Y_{t-1}|X_{t-1}$ we have

$$\begin{aligned}
& \mathbb{P}\left(X_t \leq x_t \left| X_{t-1} = x_{t-1}, Y_{t-1} = y_{t-1}\right.\right) \\
= & \lim_{dy_{t-1}\to 0} \frac{\mathbb{P}\left(X_t \leq x_t, y_{t-1} \leq Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)}{\mathbb{P}\left(y_{t-1} \leq Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)} \\
= & \lim_{dy_{t-1}\to 0} \frac{\mathbb{P}\left(X_t \leq x_t, Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right) - \mathbb{P}\left(X_t \leq x_t, Y_{t-1} \leq y_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)}{\mathbb{P}\left(Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right) - \mathbb{P}\left(Y_{t-1} \leq y_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)} \\
= & \lim_{dy_{t-1}\to 0} \frac{\mathbb{P}\left(X_t \leq x_t \left| X_{t-1} = x_{t-1}\right.\right) \mathbb{P}\left(Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)}{\mathbb{P}\left(Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right) - \mathbb{P}\left(Y_{t-1} \leq y_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)} \\
- & \frac{\mathbb{P}\left(X_t \leq x_t \left| X_{t-1} = x_{t-1}\right.\right) \mathbb{P}\left(Y_{t-1} \leq y_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)}{\mathbb{P}\left(Y_{t-1} \leq y_{t-1} + dy_{t-1} \left| X_{t-1} = x_{t-1}\right.\right) - \mathbb{P}\left(Y_{t-1} \leq y_{t-1} \left| X_{t-1} = x_{t-1}\right.\right)} \\
= & \mathbb{P}\left(X_t \leq x_t \left| X_{t-1} = x_{t-1}\right.\right)
\end{aligned} \tag{A.20}$$

By this relationship, the projected empirical copula process, if valid at least for weakly dependent random variables, could be also important in testing for Granger non causality
.

Appendix B

Appendix of Chapter 3

B.1 Inverse of Survival and Univariate Symmetry

To find the inverse of the survival function in terms of the distribution we must solve the equation

$$G\left(\bar{F}(x)\right) = G(1 - F(x)) = x \tag{B.1}$$

This is accomplished by computing (B.1) in $F^{-1}(q)$

$$\begin{aligned} G\left(1 - F\left(F^{-1}(q)\right)\right) &= F^{-1}(q) \\ \Leftrightarrow G(1-q) &= F^{-1}(q) \\ \Leftrightarrow G(p) &= F^{-1}(1-p) \quad p = 1-q \end{aligned}$$

$$\bar{F}^{-1}\left(\bar{F}(x)\right) = F^{-1}\left(1 - \bar{F}(x)\right) = F^{-1}(1 - (1 - F(x))) = F^{-1}(F(x)) = x \tag{B.2}$$

so we have

$$\bar{F}^{-1}(u) = F^{-1}(1-u) \qquad u \in [0,1] \tag{B.3}$$

Using this relation we can prove the equivalence

$$\begin{cases} F_i(c_i - x_i) = 1 - F_i(c_i + x_i) = \bar{F}_i(c_i + x_i) \\ F_i(c_i) = 1 - F_i(c_i) = \bar{F}_i(c_i) = \dfrac{1}{2} \end{cases} \Leftrightarrow \mathbb{P}(U_i \leq u_i) = \mathbb{P}\left(\bar{U}_i \leq u_i\right) \tag{B.4}$$

First the left to right implication. Let us call $u_i = F_i(c_i - x_i) = \bar{F}_i(c_i + x_i)$. Using the fact that the distribution function is non decreasing and the survival function is non increasing, We have:

$$\mathbb{P}(X_i \leq c_i - x_i) = \mathbb{P}(F_i(X_i) \leq F_i(c_i - x_i)) = \mathbb{P}(U_i \leq u_i) \tag{B.5}$$

$$\mathbb{P}(X_i > c_i + x_i) = \mathbb{P}\left(\bar{F}_i(X_i) \leq \bar{F}_i(c_i + x_i)\right) = \mathbb{P}\left(\bar{U}_i \leq u_i\right) \tag{B.6}$$

For the right to left implication let $u_i = F_i(c_i - x_i)$ we have:

$$\mathbb{P}(U_i \leq u_i) = \mathbb{P}\left(F_i^{-1}(U_i) \leq F_i^{-1}(u_i)\right) = \mathbb{P}(X_i \leq c_i - x_i) \tag{B.7}$$

and

$$\mathbb{P}\left(\bar{U}_i \leq u_i\right) = \mathbb{P}\left(\bar{F}_i^{-1}\left(\bar{U}_i\right) > \bar{F}_i^{-1}(u_i)\right) = \mathbb{P}\left(X_i > \bar{F}_i^{-1}\left(F_i\left(c_i - x_i\right)\right)\right) \tag{B.8}$$

It remains only to evaluate $\bar{F}_i^{-1}\left(F_i\left(c_i - x_i\right)\right)$:

$$\begin{aligned}
\bar{F}_i^{-1}\left(F_i\left(c_i - x_i\right)\right) &= F_i^{-1}\left(1 - F_i\left(c_i - x_i\right)\right) && \text{(B.9)}\\
&= F_i^{-1}\left(1 - \left(F_i\left(c_i\right) - F_i\left(x_i\right)\right)\right) && \text{(B.10)}\\
&= F_i^{-1}\left(\left(\bar{F}_i\left(c_i\right) + F_i\left(x_i\right)\right)\right) && \text{(B.11)}\\
&= F_i^{-1}\left(\left(F_i\left(c_i\right) + F_i\left(x_i\right)\right)\right) && \text{(B.12)}\\
&= F_i^{-1}\left(F_i\left(c_i + x_i\right)\right) = c_i + x_i && \text{(B.13)}
\end{aligned}$$

B.2 Proofs

B.2.1 Proof of proposition 4

Following [80] we can express the empirical survival copula in a more convenient way. Consider a Sample from the uniform random variables $\bar{\mathbf{U}}$ that are distributed according to $\bar{C}$.
We define

$$\begin{aligned}
\hat{\bar{G}}_n(\mathbf{u}) &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\left(\bar{\mathbf{U}}_i \leq \mathbf{u}\right) && \text{(B.14)}\\
&= \frac{1}{n}\sum_{i=1}^{n}\prod_{d=1}^{D} \mathbb{I}\left(\hat{X}_{id} > \bar{F}_d^{-1}(u_d)\right) && \text{(B.15)}\\
&= \hat{\bar{F}}\left(\bar{F}_1^{-1}(u_1), \ldots, \bar{F}_D^{-1}(u_D)\right) && \text{(B.16)}
\end{aligned}$$

the last equality follows from the fact that, by the probability integral transform, $\hat{X}_{id} = \bar{F}_d^{-1}\left(\bar{U}_{id}\right)$ has the same distribution as X_{id}. In this way we have:

$$\begin{aligned}
\hat{\bar{G}}_n\left(\bar{F}_1(x_1) \ldots \bar{F}_1(x_1)\right) &= \hat{\bar{F}}(x_1, \ldots, x_D) && \text{(B.17)}\\
\hat{\bar{G}}_{nd}\left(\bar{F}_d(x_d)\right) &= \hat{\bar{F}}_d(x_d) && \text{(B.18)}\\
\bar{F}_d^{-1}\left(\hat{\bar{G}}_{nd}^{-1}(u_d)\right) &= \hat{\bar{F}}_d^{-1}(u_d) && \text{(B.19)}
\end{aligned}$$

And we get for the empirical survival copula

$$\bar{C}(u_1, \ldots, u_d) = \hat{\bar{G}}_n\left(\hat{\bar{G}}_{n1}^{-1}(u_1), \ldots, \hat{\bar{G}}_{nd}^{-1}(u_d)\right) \tag{B.20}$$

We use the map introduced in [13] :

$$\Phi : \begin{cases} \mathbb{D}_\Phi \mapsto \ell^\infty [0,1]^D \\ H \mapsto H\left(H_1^{-1}, \ldots, H_D^{-1}\right) \end{cases} \tag{B.21}$$

where $\mathbb{D}_\Phi$ denotes the set of all distribution functions H on $[0,1]^D$ whose marginal cdfs H_d satisfy $H_d(0) = 0$. Using this map, the empirical survival copula process can be expressed as maps from the multivariate empirical processes on $[0,1]^D$, $\bar{\mathbb{G}}_n = \sqrt{n}\left(\hat{\bar{G}}_n - \bar{C}\right)$:

$$\hat{\bar{\mathbb{C}}}_n = \sqrt{n}\left(\Phi\left(\hat{\bar{G}}_n(\mathbf{u})\right) - \Phi\left(\bar{C}(\mathbf{u})\right)\right) \tag{B.22}$$

since $\bar{\mathbb{G}}_n$ is a multivariate empirical process (not a multivariate survival empirical process) the already cited results in [68] for strongly mixing data lead directly to the following weak convergence limit:

$$\bar{\mathbb{G}}_n(\mathbf{u}) \rightsquigarrow \mathbb{B}_{\bar{C}}(\mathbf{u}) \tag{B.23}$$

$$\mathbb{C}\text{ov}\left(\mathbb{B}_{\bar{C}}(\mathbf{u}), \mathbb{B}_{\bar{C}}(\mathbf{v})\right) = \bar{C}(\mathbf{u} \wedge \mathbf{v}) - \bar{C}(\mathbf{u})\bar{C}(\mathbf{v}) \tag{B.24}$$

Theorem 2.4 in [13] implies it that, under **A 3**, Φ is Hadamard differentiable at $\bar{C}$ and the application of the functional delta method to (B.22) yields the result.

B.2.2 Proof of proposition 5

Let $C[0,1]^D$ the space of function $f : [0,1]^D \to \mathbb{R}$ that are continuous $D[0,1]^D$ the space of cadlag function on $[0,1]^D$ and $BV1[0,1]^D$ as the subspace of $D[0,1]^D$ consisting of the functions with total variation bounded by one. For notational convenience we consider only one multiplier replicate, the generalization to M being straightforward. From continuous mapping theorem we get

$$\left(\left(\hat{\mathbb{C}}_n - \hat{\bar{\mathbb{C}}}_n\right)^2, \left(\tilde{\mathbb{C}}_n^{[1]} - \tilde{\bar{\mathbb{C}}}_n^{[1]}\right)^2, \hat{\mathbb{C}}_n\right) \rightsquigarrow \left(\left(\mathbb{C} - \bar{\mathbb{C}}\right)^2, \left(\mathbb{C}^{[1]} - \bar{\mathbb{C}}^{[1]}\right)^2, \mathbb{C}\right) \tag{B.25}$$

on $\left[\ell^\infty[0,1]^D\right]^4$ Because we can write

$$\left(\left(\hat{\mathbb{C}}_n - \hat{\bar{\mathbb{C}}}_n\right)^2, \left(\tilde{\mathbb{C}}_n^{[1]} - \tilde{\bar{C}}_n^{[1]}\right)^2, \hat{\mathbb{C}}_n\right) = \sqrt{n}\left(\left(\hat{\mathbb{A}}_n, \hat{\mathbb{A}}_n^{[1]}, \hat{C}_n\right) - \left(\mathbb{A}, \mathbb{A}^{[1]}, C\right)\right) \tag{B.26}$$

where $\hat{\mathbb{A}}_n = \sqrt{n}\left(\hat{C}_n - \hat{\bar{C}}_n\right)^2$, $\hat{\mathbb{A}}_n^{[1]} = \frac{1}{\sqrt{n}}\left(\left(\tilde{\mathbb{C}}_n^{[1]} - \tilde{\bar{\mathbb{C}}}_n^{[1]}\right)^2\right)$ and $\mathbb{A} = \mathbb{A}^{[1]} = 0$ Let us introduce the map $\Psi : \ell^\infty[0,1]^D \times \ell^\infty[0,1]^D \times BV1[0,1]^D \longrightarrow \mathbb{R}^2$ defined by

$$\Psi(\alpha, \tilde{\alpha}, \beta) = \left(\int_{(0,1]^D} \alpha d\beta, \int_{(0,1]^D} \tilde{\alpha} d\beta\right) \tag{B.27}$$

we have then

$$\left(n\hat{\mathbb{T}}_n, n\tilde{\mathbb{T}}_n^{[1]}\right) = \sqrt{n}\left(\Psi\left(\hat{\mathbb{A}}_n, \hat{\mathbb{A}}_n^{[1]}, \hat{C}_n\right) - \Psi\left(\mathbb{A}, \mathbb{A}^{[1]}, C\right)\right) \tag{B.28}$$

we state the Hadamard differentiability of Ψ tangentially to $C[0,1]^D \times C[0,1]^D \times D[0,1]^D$ at each $(\alpha, \tilde{\alpha}, \beta)$ in $\ell^\infty[0,1]^D \times \ell^\infty[0,1]^D \times BV1[0,1]^D$ such that $\int |d\alpha| < \infty$ and $\int |d\tilde{\alpha}| < \infty$ in the lemma 3 below. Then an application of the functional delta method gives

$$\left(n\hat{\mathbb{T}}_n, n\tilde{\mathbb{T}}_n^{[1]}\right) \rightsquigarrow \Psi'_{A,A^{[1]},C}\left(\left(\mathbb{C} - \bar{\mathbb{C}}\right)^2, \left(\mathbb{C}^{[1]} - \bar{\mathbb{C}}^{[1]}\right)^2, \mathbb{C}\right) \tag{B.29}$$

with

$$\Psi'_{A,A^{[1]},C}\left(\left(\mathbb{C}-\bar{\mathbb{C}}\right)^2,\left(\mathbb{C}^{[1]}-\bar{\mathbb{C}}^{[1]}\right)^2,\mathbb{C}\right) \tag{B.30}$$

$$=\left(\int_{(0,1]^D}\mathbb{A}d\mathbb{C}+\int_{(0,1]^D}\left(\mathbb{C}-\bar{\mathbb{C}}\right)^2dC,\int_{(0,1]^D}\mathbb{A}^{[1]}d\mathbb{C}+\int_{(0,1]^D}\left(\mathbb{C}^{[1]}-\bar{\mathbb{C}}^{[1]}\right)^2dC\right) \tag{B.31}$$

$$=\left(\int_{(0,1]^D}\left(\mathbb{C}-\bar{\mathbb{C}}\right)^2dC,\int_{(0,1]^D}\left(\mathbb{C}^{[1]}-\bar{\mathbb{C}}^{[1]}\right)^2dC\right)=\left(\mathbb{T},\mathbb{T}^{[1]}\right) \tag{B.32}$$

Lemma 3 *The map Ψ defined in (B.27) is Hadamard Differentiable tangentially to $C\,[0,1]^D\times C\,[0,1]^D\times D\,[0,1]^D$ at each $(\alpha,\tilde{\alpha},\beta)$ in $\ell^\infty\,[0,1]^D\times\ell^\infty\,[0,1]^D\times BV1\,[0,1]^D$ such that $\int|d\alpha|<\infty$ and $\int|d\tilde{\alpha}|<\infty$ with derivative given by*

$$\Psi'_{A,\tilde{A},B}\left(\alpha,\tilde{\alpha},\beta\right)=\left(\int_{(0,1]^D}Ad\beta+\int_{(0,1]^D}\alpha dB,\int_{(0,1]^D}\tilde{A}d\beta+\int_{(0,1]^D}\tilde{\alpha}dB\right) \tag{B.33}$$

where $\int\alpha\beta,\int\tilde{\alpha}\beta$ are defined via the D-dimensional integration by parts formula exemplified for 2 dimension in Theorem 8.8 of [47] if β is not of bounded variation.

Lemma 3 is a vectorized D-dimensional version of lemma 3.9.17 in [82] (see also lemma 4.3 of [14])and since the proof is similar, it will be omitted.

B.3 Bandwidth Selection and generation of dependent multiplier sequences

B.3.1 Bandwidth ℓ_n

In this appendix we give more details on the concrete implementation of the dependent multiplier bootstrap by considering the bandwidth selection and the generation of the dependent multiplier sequence. We adapt the procedure introduced in [11] for the estimation of the bandwidth parameter ℓ_n to be more coherent with our test statistic. In their paper, the optimal bandwidth is obtained by minimizing the integrated MSE of an estimator of

$$\sigma_C\left(\mathbf{u},\mathbf{v}\right)=\mathbb{C}\mathrm{ov}\left(\mathbb{B}_C\left(\mathbf{u}\right),\mathbb{B}_C\left(\mathbf{v}\right)\right) \tag{B.34}$$

given by

$$\hat{\sigma}_n\left(\mathbf{u},\mathbf{v}\right)=\sum_{k=-L}^{L}k_{F,0.5}\left(k/L\right)\hat{\gamma}_n\left(k,\mathbf{u},\mathbf{v}\right) \tag{B.35}$$

where L is an integer >1 to be chosen in the following,

$$\hat{\gamma}_n\left(k,\mathbf{u},\mathbf{v}\right)=\begin{cases}n^{-1}\sum_{i=1}^{n-k}\left\{\mathbb{I}\left(\hat{U}_i<u\right)-\hat{C}_n\left(u\right)\right\}\left\{\mathbb{I}\left(\hat{U}_{i+k}<v\right)-\hat{C}_n\left(v\right)\right\} & k\geq 0\\ n^{-1}\sum_{i=1-k}^{n}\left\{\mathbb{I}\left(\hat{U}_i<u\right)-\hat{C}_n\left(u\right)\right\}\left\{\mathbb{I}\left(\hat{U}_{i+k}<v\right)-\hat{C}_n\left(v\right)\right\} & k<0\end{cases} \tag{B.36}$$

and $k_{F,c}$ is the flat top kernel

$$k_{F,c} = \{[(1-|x|)/(1-c)] \vee 0\} \wedge 1 \tag{B.37}$$

In order to obtain the optimal ℓ_n they minimize

$$IMSE_U(\hat{\sigma}_n(\mathbf{u},\mathbf{v})) = \int_{[0,1]^{2D}} MSE(\hat{\sigma}_n(\mathbf{u},\mathbf{v}))\, d\mathbf{u}d\mathbf{v} \tag{B.38}$$

approximating the integral with a finite grid. To avoid the arbitrary choice of the grid and to be coherent with our test statistic we choose to minimize

$$IMSE_{\hat{C}_n}(\hat{\sigma}_n(\mathbf{u},\mathbf{v})) = \int_{[0,1]^{2D}} MSE(\hat{\sigma}_n(\mathbf{u},\mathbf{v}))\, d\hat{C}_n(\mathbf{u})\, d\hat{C}_n(\mathbf{v}) \tag{B.39}$$

$$= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} MSE\left(\hat{\sigma}_n\left(\hat{\mathbf{U}}_i,\hat{\mathbf{U}}_j\right)\right) \tag{B.40}$$

In complete analogy with their computations our optimal bandwidth is

$$\ell_{opt} = \left(\frac{4\hat{\hat{\Gamma}}_{n,\hat{C}_n}}{\hat{\hat{\Delta}}_{n,\hat{C}_n}}\right) n^{1/5} \tag{B.41}$$

$$\hat{\hat{\Gamma}}_{n,\hat{C}_n} = \frac{1}{4}\left.\frac{d^2\phi(x)}{dx^2}\right|_{x=0} n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=-L}^{L} k_{F,0.5}k^2\hat{\gamma}_n\left(k,\hat{\mathbf{U}}_i,\hat{\mathbf{U}}_j\right) \tag{B.42}$$

$$\hat{\hat{\Delta}}_{n,\hat{C}_n} = \left\{\int_{-1}^{1}\phi(x)^2\,dx\right\}\left[\left(n^{-1}\sum_{i=1}^{n}\hat{\sigma}_n\left(\hat{\mathbf{U}}_i,\hat{\mathbf{U}}_i\right)\right)^2 - n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\sigma}_n\left(\hat{\mathbf{U}}_i,\hat{\mathbf{U}}_j\right)\right] \tag{B.43}$$

Then we choose L as the minimum lag for which autocorrelation of all series becomes negligible using the automatic procedure proposed in [63] in the matlab implementation that can be found on Andrew Patton website.

B.3.2 Dependent Multiplier sequence $\xi_{i,n}$

Once we have chosen the bandwidth we are ready to generate a dependent multiplier sequence according to the moving average method introduced in [12] and discussed in detail in [11]. Let k be some positive bounded real function such that $k(x) > 0$ for all $|x| < 1$. let b_n be a sequence of integers such that $b_n \to 0$, $b_n = o(n)$ and $b_n \geq 1$ for all $n \in \mathbb{N}$. Let $Z_1, \ldots, Z_{n+2b_n-2}$ be i.e. random variables independent of the sample such that $\mathbb{E}(Z_1) = 0$,$\mathbb{E}(Z_1^2) = 1$ and $\mathbb{E}(|Z_1|^\nu) < \infty$ for all $\nu > 2$. Then let $\ell_n = 2b_n - 1$, for any $j \in \{1, \ldots, \ell_n\}$, let $w_{j,n} = k((j-b_n)/b_n)$ and $\tilde{w_{j,n}} = w_{j,n}\left(\sum_{j=1}^{\ell_n} w_{j,n}^2\right)^{-\frac{1}{2}}$. For each $i \in \{1, \ldots, n\}$ they show that the sequence

$$\xi_{i,n} = \sum_{j=1}^{\ell_n}\tilde{w}_{j,n}Z_{j+i-1} \tag{B.44}$$

is a dependent multiplier sequence as defined in subsection 3.3.2 with function ϕ given by

$$\phi(x) = \frac{k \star k(2x)}{k \star k(2x)} \tag{B.45}$$

B.3.3 Additional Details

In all the simulations and data application performed, we draw Z_j from a standard normal distribution and we choose $k(x)$ to be the Bartlett kernel

$$k(x) = k_B(x) = (1 - |x|) \vee 0 \tag{B.46}$$

with the previous choice it follows that ϕ is the Parzen kernel

$$\begin{aligned} \phi(x) = k_P(x) &= \left(1 - 6x^2 - 6|x|^3\right) \mathbb{I}(|x| \leq 1/2) \\ &+ 2\left(1 - |x|^3\right) \mathbb{I}(1/2 < |x| \leq 1) \end{aligned} \tag{B.47}$$

and that the quantities needed for the bandwidth estimation are

$$\left.\frac{d^2\phi(x)}{dx^2}\right|_{x=0} = -12 \tag{B.48}$$

$$\int_{-1}^{1} \phi(x)^2 \, dx = 151/280 \tag{B.49}$$

Appendix C

Appendix of Chapter 4

C.1 Proof of Proposition 6

First of all we need a bound on j-th recursive coefficents α_j, β_j. For doing this we use the matrix relation among the matrix of weights and support points and the matrix of recursion coefficients of orthonormal polynomials (c.f. [33])

$$
\begin{aligned}
Q^T \begin{pmatrix} 1 & \sqrt{w_1} & \sqrt{w_2} & \dots & \sqrt{w_N} \\ \sqrt{w_1} & x_1 & 0 & \dots & 0 \\ \sqrt{w_2} & 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{w_2} & 0 & 0 & \dots & x_n \end{pmatrix} Q &= \begin{pmatrix} 1 & \sqrt{\beta_0} & 0 & \dots & 0 \\ \sqrt{\beta_0} & \alpha_0 & \sqrt{\beta_1} & \dots & 0 \\ 0 & \sqrt{\beta_1} & \alpha_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_n \end{pmatrix} \\
Q^T T Q &= J
\end{aligned}
$$

with $Q^T Q = I_N$. Analogous result with in general different orthogonal matrices holds also for the j-th principal minors of the two matrices that involves only weights support points and recursion coefficent till j:

$$
Q^T_{(j)} T_{(j)} Q_{(j)} = J_{(j)} \tag{C.1}
$$

Using some matrix norm inequality we can bound the recursion coefficents by knowing that the sum of weights is bounded by K_1 and that the support points are also bounded by a and b. We have

$$
\max_{k,l=1,\dots,j} J_{(j)kl} \leq \left\| J_{(j)} \right\|_2 \leq \left\| (T_{(j)} \right\|_2 \leq \sqrt{\left\| T_{(j)} \right\|_1 \left\| T_{(j)} \right\|_\infty} \tag{C.2}
$$

where $\| \|_2$ is the spectral radius that it is equal for the two matrices being them linked by an orthogonal transformation and $\| \|_1, \| \|_\infty$ are the max row sum and the max column sum respectively. In the end we have

$$
\begin{aligned}
\max_{k,l=1,\dots,j} J_{(j)kl} &\leq \max_{l=1,\dots,j} \left(\max\left((1 + \sum_{m=1}^{j} \sqrt{w_m}), \sqrt{w_l} + x_l \right) \right)^2 \\
&\leq K_2 \\
K_2 &\equiv \max\left(\left(1 + \sqrt{j}\sqrt{K_1} \right), \sqrt{K_1} + \max\left(|a|, |b| \right) \right)^2
\end{aligned} \tag{C.3}
$$

This means that the recursion coefficent are bounded by K_2. It can be shown that the j-th orthogonal polinomial is a polinomial of j-th degree with coefficents of the power k given by rational function of product of k recursion coefficents so that the modulus of the j-th orthogonal polinomial can be roughly bounded by $jK_2^j \max(|a|, |b|)^j$ that for j fixed does not depend from N and an analogous bound can be obtained for the derivatives of the polynomials that being polnomials of lower degree can be bounded by $(j-m) K_2^j \max(|a|, |b|)^{j-m}$.

For notational convenience we prove the theorem only for the two dimensional case using u for u_1 and v for u_2. The extension to the general case does not present any additional difficulty.

Regarding the first part of the theorem, from $0 \le \left|\hat{C}_N\right| \le 1$, follows that $\hat{C}_N \in L^2(\mathbb{R}, d\mu_N)$

$$\left\|\hat{C}_N\right\|_2^2 = \int_{\mathbb{R}} \left|\hat{C}_N(u+h\xi, v)\right|^2 d\mu_N(\xi) \le \int_{\mathbb{R}} d\mu_N(\xi) \tag{C.4}$$

And since by construction $p_j \in L^2(\mathbb{R}, d\mu_N)$ we can use Cauchy–Schwarz inequality to bound $\left|\mathcal{D}^{OP}_{u,m}\hat{C}_N(u,v)\right|$

$$\begin{aligned}\left|\mathcal{D}^{OP}_{u,m}\hat{C}_N(u,v)\right| &= \left|\frac{1}{h}\sum_{j=1}^{m}\int_{\mathbb{R}} \hat{C}_N(u+h\xi, v)\, p_j(\xi)\, d\mu_N(\xi) \left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \\ &= \left|\sum_{j=1}^{m}\int_{\mathbb{R}} \frac{\hat{C}_N(u+h\xi, v) - \hat{C}_N(u,v)}{h\xi}\xi p_j(\xi)\, d\mu_N(\xi) \left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right|\end{aligned}$$

The last expression come from the following relationship that uses the orthogonality of polynomials and that the sum over j starts from 1:

$$\hat{C}_N(u,v)\frac{1}{h}\sum_{j=1}^{m}\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\int_{\mathbb{R}} p_j(\xi)\, d\mu_N(\xi) = 0 \tag{C.5}$$

and since finite difference derivative is bounded for every ξ it is in $L^2(\mathbb{R}, d\mu_N)$ and we get the result by Cauchy-Swartz inequality and the bounds on polynomials derived earlier :

$$\begin{aligned}\left|\mathcal{D}^{OP}_{u,m}\hat{C}_N(u,v)\right| &\le \sum_{j=1}^{m}\left\|\mathcal{D}^{FD}_{u,m,h\xi}\hat{C}_N(u,v)\,\xi p_j(\xi)\right\|_1 \left|\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \\ &\le \left\|\mathcal{D}^{FD}_{u,m,h\xi}\hat{C}_N(u,v)\right\|_2 \sum_{j=1}^{m}\|\xi p_j(\xi)\|_2 \left|\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \\ &\le \left\|\mathcal{D}^{FD}_{u,m,h\xi}\hat{C}_N(u,v)\right\|_2 \|\xi p_1(\xi)\|_2 \left|\left.\frac{\partial p_1}{\partial x}\right|_{x=0}\right|\end{aligned}$$

For the second part, we begin proving that that $\Delta_{h\xi}\hat{\mathbb{C}}_N(u,v) \equiv \hat{\mathbb{C}}_N(u+h\xi, v) -$

$\hat{\mathbb{C}}_N(u,v) \in L^2(\mathbb{R}, d\mu_N)$. We have:

$$\left\|\Delta_{h\xi}\hat{\mathbb{C}}_N\right\|_2 = \sqrt{\int_{\mathbb{R}} \left|\Delta_{h\xi}\hat{\mathbb{C}}_N(u,v)\right|^2 d\mu_N(\xi)} \tag{C.6}$$

$$\leq \sqrt{\sqrt{N}2\int_a^b d\mu_N(\xi)} \tag{C.7}$$

where we used the fact that both the copula and the empirical copula are non negative and bounded by one. Under the assumptions of the proposition we can apply the first Diekema derivative to the copula obtaining

$$\begin{aligned} \mathcal{D}^{OP}_{u,m}C(u,v) &= \frac{1}{h}\sum_{j=1}^{m}\int_{\mathbb{R}} C(u+h\xi,v)\,p_j(\xi)\,d\mu_N(\xi)\left.\frac{\partial p_j}{\partial x}\right|_{x=0} \\ &= \partial_u C(u,v) + o\left(h^{2-1}\right) \end{aligned} \tag{C.8}$$

Then we must study the probability limit of:

$$\mathcal{D}^{OP}_{u,m}\left(\hat{C}_N(u,v) - C(u,v)\right) = \mathcal{D}^{OP}_{u,m}\frac{\hat{\mathbb{C}}_N(u,v)}{\sqrt{N}}$$

$$= \frac{1}{h}\sum_{j=1}^{m}\int_{\mathbb{R}}\frac{\hat{\mathbb{C}}_N(u+h\xi,v)}{\sqrt{N}}p_j(\xi)\,d\mu_N(\xi)\left.\frac{\partial p_j}{\partial x}\right|_{x=0} \tag{C.9}$$

using $h = \frac{1}{\sqrt{N}}$ we obtain:

$$\sum_{j=1}^{m}\int_{\mathbb{R}}\hat{\mathbb{C}}_N(u+h\xi,v)\,p_j(\xi)\,d\mu(\xi)\left.\frac{\partial p_j}{\partial x}\right|_{x=0} \tag{C.10}$$

in the above formula we can rewrite :

$$\hat{\mathbb{C}}_N(u+h\xi,v) = \hat{\mathbb{C}}_N(u,v) + \Delta_{h\xi}\hat{\mathbb{C}}_N(u,v) \tag{C.11}$$

and we have for a given N:

$$\left|\sum_{j=1}^{m}\int_{\mathbb{R}}\hat{\mathbb{C}}_N(u+h\xi,v)\,p_j(\xi)\,d\mu(\xi)\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \tag{C.12}$$

$$\leq \left|\hat{\mathbb{C}}_N(u,v)\right|\left|\sum_{j=1}^{m}\int_{\mathbb{R}}p_j(\xi)\,d\mu(\xi)\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \tag{C.13}$$

$$+\sum_{j=1}^{m}\left\|\Delta_{h\xi}\hat{\mathbb{C}}_N\right\|_2\left|\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \tag{C.14}$$

$$\leq \left\|\Delta_{h\xi}\hat{\mathbb{C}}_N\right\|_2\sum_{j=1}^{m}\left|\left.\frac{\partial p_j}{\partial x}\right|_{x=0}\right| \tag{C.15}$$

Where we have used the orthogonality with respect to a polynomial of minor degree and again the Cauchy-swhartz inequality. Since by the bounds obtained before the sum is bounded the result follows by the continuity of the limiting process of $\hat{\mathbb{C}}_N$:

$$\left\|\Delta_{h\xi}\hat{\mathbb{C}}_N\right\|_2 = O_P\left(1\right)\sqrt{\int_{\mathbb{R}} d\mu} \tag{C.16}$$

Bibliography

[1] Andrew Ang and Joseph Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494, 2002.

[2] Brendan K Beare and Juwon Seo. Time irreversible copula-based markov models. *Econometric Theory*, pages 1–38, 2012.

[3] Fabio Bellini and Leonardo Bottolo. Stationarity domains for δ-power garch process with heavy tails. *Statistics & probability letters*, 77(13):1418–1427, 2007.

[4] Noomen Ben Ghorbal, Christian Genest, and Johanna Nešlehová. On the ghoudi, khoudraji, and rivest test for extreme-value dependence. *Canadian Journal of Statistics*, 37(4):534–552, 2009.

[5] Patrick Billingsley. *Convergence of probability measures*, volume 493. John Wiley & Sons, 2009.

[6] Taoufik Bouezmarni, Jeroen VK Rombouts, and Abderrahim Taamouti. Nonparametric copula-based test for conditional independence with applications to granger causality. *Journal of Business & Economic Statistics*, 30(2):275–287, 2012.

[7] Eric Bouyé and Mark Salmon. Dynamic copula quantile regressions and tail area dynamic dependence in forex markets. *The European Journal of Finance*, 15(7-8):721–750, 2009.

[8] Salim Bouzebda and Mohamed Cherfi. Test of symmetry based on copula function. *Journal of Statistical Planning and Inference*, 142(5):1262–1271, 2012.

[9] John P Boyd and Fei Xu. Divergence (runge phenomenon) for least-squares polynomial approximation on an equispaced grid and mock–chebyshev subset interpolation. *Applied Mathematics and Computation*, 210(1):158–168, 2009.

[10] Axel Bücher and Holger Dette. A note on bootstrap approximations for the empirical copula process. *Statistics & probability letters*, 80(23):1925–1932, 2010.

[11] Axel Bücher and Ivan Kojadinovic. A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *arXiv preprint arXiv:1306.3930*, 2013.

[12] Axel Bücher and Martin Ruppert. Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis*, 116:208–229, 2013.

[13] Axel Bücher and Stanislav Volgushev. Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis*, 119:61–70, 2013.

[14] Alberto Carabarin-Aguirre and B Gail Ivanoff. Estimation of a distribution under generalized censoring. *Journal of Multivariate Analysis*, 101(6):1501–1519, 2010.

[15] Marine Carrasco and Xiaohong Chen. Mixing and moment properties of various garch and stochastic volatility models. *Econometric Theory*, 18(01):17–39, 2002.

[16] Yi-Ting Chen. Moment-based copula tests for financial returns. *Journal of Business & Economic Statistics*, 25(4):377–397, 2007.

[17] Umberto Cherubini, Elisa Luciano, and Walter Vecchiato. *Copula methods in finance*. John Wiley & Sons, 2004.

[18] Umberto Cherubini, Sabrina Mulinacci, Fabio Gobbi, and Silvia Romagnoli. *Dynamic Copula methods in finance*, volume 625. John Wiley & Sons, 2011.

[19] Umberto Cherubini, Sabrina Mulinacci, and Silvia Romagnoli. A copula-based model of speculative price dynamics in discrete time. *Journal of Multivariate Analysis*, 102(6):1047–1063, 2011.

[20] Umberto Cherubini and Silvia Romagnoli. The dependence structure of running maxima and minima: results and option pricing applications. *Mathematical Finance*, 20(1):35–58, 2010.

[21] R. Dennis Cook and Mark E. Johnson. A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(2):pp. 210–218, 1981.

[22] R Dennis Cook and Mark E Johnson. Generalized burr-pareto-logistic distributions with applications to a uranium exploration data set. *Technometrics*, 28(2):123–131, 1986.

[23] H.S.M. Coxeter. *Introduction to Geometry*. Wiley Classics Library. Wiley, 1989.

[24] William F Darsow, Bao Nguyen, Elwood T Olsen, et al. Copulas and markov processes. *Illinois Journal of Mathematics*, 36(4):600–642, 1992.

[25] A Philip Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–31, 1979.

[26] Paul Deheuvels. La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.(5)*, 65(6):274–292, 1979.

[27] Azam Dehgani, Ali Dolati, and Manuel Úbeda-Flores. Measures of radial asymmetry for bivariate random vectors. *Statistical Papers*, 54(2):271–286, 2013.

[28] Enno Diekema and Tom H Koornwinder. Differentiation by integration using orthogonal polynomials, a survey. *Journal of Approximation Theory*, 164(5):637–667, 2012.

[29] Jean-David Fermanian, Dragan Radulovic, Marten Wegkamp, et al. Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860, 2004.

[30] Jean-Pierre Florens and Michel Mouchart. A note on noncausality. *Econometrica: Journal of the Econometric Society*, pages 583–591, 1982.

[31] Gabriel Frahm, Markus Junker, and Alexander Szimayer. Elliptical copulas: applicability and limitations. *Statistics & Probability Letters*, 63(3):275–286, 2003.

[32] Maurice J Frank. On the simultaneous associativity off (x, y) andx+y- f (x, y). *Aequationes Mathematicae*, 19(1):194–226, 1979.

[33] W. Gautschi. *Orthogonal Polynomials: Computation and Approximation*. Numerical mathematics and scientific computation. Oxford University Press, 2004.

[34] Walter Gautschi. Orthogonal polynomials (in matlab). *Journal of Computational and Applied Mathematics*, 178(1):215–234, 2005.

[35] Christian Genest and Anne-Catherine Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering*, 12(4):347–368, 2007.

[36] Christian Genest, Johanna Nešlehová, and Jean-François Quessy. Tests of symmetry for bivariate copulas. *Annals of the Institute of Statistical Mathematics*, 64(4):811–834, 2012.

[37] Christian Genest and Johanna G Nešlehová. On tests of radial symmetry for bivariate copulas. *Statistical Papers*, pages 1–13, 2013.

[38] Christian Genest, Jean-François Quessy, and Bruno Rémillard. Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33(2):337–366, 2006.

[39] Christian Genest, Bruno Rémillard, and David Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2):199–213, 2009.

[40] Pierre Georges, Arnaud-Guilhem Lamy, Emeric Nicolas, Guillaume Quibel, and Thierry Roncalli. Multivariate survival modelling: A unified approach with copulas. 2001.

[41] Gene H Golub and Gérard Meurant. *Matrices, moments and quadrature with applications*. Princeton University Press, 2009.

[42] Israel S Gradshteyn and IM Ryzhik. Table of integrals, series, and products. translated from the russian. translation edited and with a preface by alan jeffrey and daniel zwillinger, 2000.

[43] Newbold P. Granger, C.W.J. *Forecasting Economic Time Series(2nd edition)*. Academic Press, 1986.

[44] László Györfi and Harro Walk. Strongly consistent nonparametric tests of conditional independence. *Statistics & Probability Letters*, 82(6):1145–1150, 2012.

[45] Ingrid Hobaek Haff and Johan Segers. Nonparametric estimation of pair-copula constructions with the empirical pair-copula. *arXiv preprint arXiv:1201.5133*, 2012.

[46] Michael Harder and Ulrich Stadtmüller. Maximal non-exchangeability in dimension d. *Journal of Multivariate Analysis*, 124:31–41, 2014.

[47] T.H. Hildebrandt. *Introduction to the theory of integration*. Pure and applied mathematics. Academic Press, 1963.

[48] Marius Hofert, Ivan Kojadinovic, Martin Maechler, and Jun Yan. copula: Multivariate dependence with copulas. *R package version 0.999-7, URL http://CRAN. R-project. org/package= nacopula*, 2013.

[49] Yongmiao Hong, Jun Tu, and Guofu Zhou. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies*, 20(5):1547–1581, 2007.

[50] Harry Joe. *Multivariate models and multivariate dependence concepts*, volume 73. CRC Press, 1997.

[51] Eric Jondeau, Ser-Huang Poon, and Michael Rockinger. *Financial modeling under non-Gaussian distributions*. Springer, 2007.

[52] Erich Peter Klement, Radko Mesiar, and Endre Pap. Invariant copulas. *Kybernetika*, 38(3):275–286, 2002.

[53] Ivan Kojadinovic, Jun Yan, et al. Modeling multivariate distributions with continuous margins using the copula r package. *Journal of Statistical Software*, 34(9):1–20, 2010.

[54] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[55] Michael R Kosorok. *Introduction to empirical processes and semiparametric inference*. Springer, 2007.

[56] Bo Li and Marc G Genton. Nonparametric identification of copula structures. *Journal of the American Statistical Association*, 108(502):666–675, 2013.

[57] Francois Longin and Bruno Solnik. Extreme correlation of international equity markets. *The Journal of Finance*, 56(2):649–676, 2001.

[58] Roger B Nelsen. Some concepts of bivariate symmetry. *Journal of Nonparametric Statistics*, 3(1):95–101, 1993.

[59] Roger B Nelsen. *An introduction to copulas*. Springer, 2007.

[60] Daniel B Nelson. Stationarity and persistence in the garch (1, 1) model. *Econometric theory*, 6(03):318–334, 1990.

[61] Roger B Nelson. Extremes of nonexchangeability. *Statistical Papers*, 48(2):329–336, 2007.

[62] A Patton. On the importance of skewness and asymmetric dependence in stock returns for asset allocation. *Manuscript, UCSD*, 2001.

[63] Dimitris N Politis and Halbert White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70, 2004.

[64] David Pollard. Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*, pages i–86, 1990.

[65] Jean-François Quessy. Testing for bivariate extreme dependence using kendall's process. *Scandinavian Journal of Statistics*, 39(3):497–514, 2012.

[66] Bruno Rémillard and Olivier Scaillet. Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100(3):377–386, 2009.

[67] Emmanuel Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31. Springer, 1999.

[68] Emmanuel Rio. Inequalities and limit theorems for weakly dependent sequences. 2013.

[69] JF Rosco and Harry Joe. Measures of tail asymmetry for bivariate copulas. *Statistical Papers*, 54(3):709–726, 2013.

[70] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, 42(1):43, 1956.

[71] Ludger Ruschendorf. Asymptotic distributions of multivariate rank order statistics. *The Annals of Statistics*, pages 912–923, 1976.

[72] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.

[73] Olivier Scaillet. A kolmogorov-smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics*, 33(3):415–427, 2005.

[74] Friedrich Schmid and Rafael Schmidt. Multivariate extensions of spearman's rho and related statistics. *Statistics & Probability Letters*, 77(4):407–416, 2007.

[75] Rafael Schmidt and Ulrich Stadtmüller. Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33(2):307–335, 2006.

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

[77] Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*, volume 59. Siam, 2009.

[78] Masaaki Sibuya. Bivariate extreme statistics, i. *Annals of the Institute of Statistical Mathematics*, 11(2):195–210, 1959.

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

[80] Hideatsu Tsukahara. Semiparametric estimation in copula models. *Canadian Journal of Statistics*, 33(3):357–375, 2005.

[81] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

[82] Aad W Van Der Vaart and Jon A Wellner. *Weak Convergence*. Springer, 1996.

[83] Halbert White and Xun Lu. Granger causality and dynamic structural systems. *Journal of Financial Econometrics*, 8(2):193–243, 2010.