ARTICLE IN PRESS
JID: ECOSTA [m3Gsc;May 8, 2023;14:11]
Econometrics and Statistics xxx (xxxx) xxx
Contents lists available at ScienceDirect
Econometrics and Statistics
journal homepage: www.elsevier.com/locate/ecosta

Nonparametric estimation of copulas and copula densities by orthogonal projections

Yves I. Ngounou Bakam[a], Denys Pommeret[a,b,*]

[a] Aix-Marseille University, Ecole Centrale, CNRS, I2M, Campus de Luminy, 13288 Marseille cedex 9, France
[b] ISFA, Univ Lyon, UCBL, LSAF EA2429, Lyon, F-69007, France

ARTICLE INFO

Article history:
Received 5 May 2022
Revised 14 April 2023
Accepted 14 April 2023
Available online xxx

Keywords:
Copula and copula density estimators
Copula coefficients
Orthogonal Legendre polynomials
Nonparametric estimation

ABSTRACT

A nonparametric copula density estimator based on Legendre orthogonal polynomials is proposed. A nonparametric copula estimator is then deduced by integration. Their asymptotic properties are reviewed. Both estimators are based on a sequence of moments that characterize the copulas and that we shall call the *copula coefficients*. A data-driven method is proposed to select the number of copula coefficients to use. An intensive simulation study shows the good performance of both copulas and copula densities estimators compared to a large panel of competitors. Two real datasets illustrate this approach.

1. Introduction

Consider a d-random vector $\mathbf{X} = (X_1, \dots, X_d)^T$ with joint cumulative distribution function (cdf) H and marginal cdf $F_1, \dots, F_d$, that we assumed to be continuous. According to Sklar's Theorem (Sklar, 1959), there exists a unique d-variate function C such that

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)).$$

The function C is called the copula associated to $\mathbf{X}$. The copula is a joint cdf on $[0, 1]^d$, with uniform margins and satisfying $C(u_1, \dots, u_d) = H(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$, where, for $j = 1, \dots, d$, $F_j^{-1}(u_j) = \inf\{x_j; F_j(x_j) \geq u_j\}$, is the quantile function of F_j. Assuming that for $j = 1, \dots, d$, F_j is differentiable, we can express the joint density h of $\mathbf{X}$ (with respect to the Lebesgue measure on $[0, 1]^d$) as

$$h(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{j=1}^{d} f_j(x_j),$$

where for $j = 1, \dots, d$, f_j is the marginal density of X_j and where

$$c := \frac{\partial^d C}{\partial x_1 \cdots \partial x_d},$$

is called the copula density of $\mathbf{X}$.

* Corresponding author.
E-mail address: denys.pommeret@univ-amu.fr (D. Pommeret).

https://doi.org/10.1016/j.ecosta.2023.04.002

Please cite this article as: Y.I. Ngounou Bakam and D. Pommeret, Nonparametric estimation of copulas and copula densities by orthogonal projections, Econometrics and Statistics, https://doi.org/10.1016/j.ecosta.2023.04.002

Copulas and copula densities have a large spectra of applications as described for instance in Joe (2014). They are largely used as a tool to identify a wide variety of properties such as tail-dependencies, heavy-tail, asymmetries or heavy-tail behaviours. Copula estimation has given rise to numerous works. In the parametric statistics literature the maximum likelihood estimation is commonly used, with a general two-stage procedure as in Ko and Hjort (2019), or with a penalized likelihood as in Qu and Yin (2012) and more recently for Archimax copulas in Chatelain et al. (2020). Such parametric methods can suffer from restrictive specification errors with few parameters. Nonparametric copula estimations generally offer a more flexible alternative with several different approaches, such as: empirical processes (see Deheuvels, 1979), kernel estimators (see for instance Geenens et al., 2017; Omelka et al., 2009) where the authors proposed a robust method, B-spline estimators (Kauermann et al., 2013), Bernstein polynomials (Bouezmarni et al., 2010; 2013), dynamic dependence (Krupskii and Joe, 2020), linear wavelet estimators (see Genest et al., 2009), nonlinear wavelet estimators (see Autin et al., 2010) or Legendre multiwavelet estimators (see Chatrabgoun et al., 2017).

In this paper we first propose to revisit the nonparametric method for estimating copula densities by considering an orthogonal shifted Legendre polynomials expansion. In this sense, this approach lies between the Legendre multiwavelet procedure and the Bernstein method. It has been first studied by Gui (2009) who considered the estimation of the copula density in the bivariate case. He used a truncated series of orthogonal Legendre polynomials and obtained some theoretical properties. A few illustrations demonstrated the potential of such approach. Here we propose to adapt the work of Gui (2009) in two directions: i) first, we extend the copula density estimators to the d-multivariate case, with $d \geq 2$; ii) second, by integration we propose a new estimator of the copula.

More precisely, the basic idea developed here is to consider the transformations $U_j = F_j(X_j)$, for $j = 1, \ldots, d$, which yield a vector of uniform random variables denoted by

$$\mathbf{U} := (U_1, \ldots, U_d)^T.$$

Its joint cdf has the form

$$H_{\mathbf{U}}(u_1, \ldots, u_d) = H(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)),$$

which shows that $\mathbf{U}$ and $\mathbf{X}$ have the same structure of dependence with the same copula. We deduce that

$$h_{\mathbf{U}}(u_1, \ldots, u_d) = c(u_1, \ldots, u_d), \tag{1}$$

where $h_{\mathbf{U}}$ denotes the joint density of $\mathbf{U}$ with respect to the Lebesgue measure on $[0, 1]^d$. This basic equality is the start of the construction of the copula density estimator, expressing c in a basis of orthogonal polynomials as done in Gui (2009). In addition, a new copula estimator is derived by integrating the polynomial expansion. Both estimators are based on a characteristic sequence of the copulas, which we refer to as the *copula coefficients*. The properties of these estimators are studied: the copula density estimator satisfies the uniform margins property (see Proposition 3) that almost all nonparametric estimators in the literature suffer. Various asymptotic properties are reviewed for both estimators. As in Gui (2009) we propose a data-driven method to select the number of projections. In the case of independence we obtain the exact form of the copula and our estimator then seems to greatly surpass all the other competitors. Finally, both estimators seem to outperform most others known in the statistical literature for a wide spectrum of scenarios, excepted in the bivariate case where the density estimators of Geenens et al. (2017) are particularly efficient. Moreover, our estimators are very simple and easy to implement and their execution time is very fast. Their interest is also demonstrated on two real datasets: i) first in insurance by considering the Loss-ALAE dataset which contains indemnity payments (Loss) and allocated loss adjustment expens; ii) second in finance where we study the dependence structure of exchange rates. These data are described in the Supplementary Material.

The paper is organised as follows. In Section 2 we develop the methodology of the estimation procedure, generalizing the copula density estimators of Gui (2009) in a multivariate setting and extending it to copula estimation. Section 3 is devoted to some elementary and asymptotic properties. A selection method of the number of expansion components and a numerical comparison with other nonparametric estimators are proposed in Section 4. Section 5 contains two applications in actuarial and finance science and a conclusion is given in Section 6. All proofs are relegated at the end of the paper.

2. Construction of the estimators

Let us denote by μ the uniform measure on [0,1] and by $\boldsymbol{\mu}^d := \mu \times \cdots \times \mu$ the product of the uniform measure d times on $[0, 1]^d$. We denote by $\{Q_m; m \in \mathbb{N}\}$ the orthonormal basis of shifted Legendre polynomials satisfying

$$\int_{[0,1]} Q_m(x) Q_k(x) \mu(dx) = \delta_{mk},$$

with $\delta_{mk} = 1$ if $m = k$ and 0 otherwise. The orthonormal shifted Legendre polynomials Q_m are defined on [0,1] by

$$Q_m(x) = \sqrt{2m+1}\, L_m(2x-1),$$

where L_m are the well known classical Legendre polynomials defined by

$$L_0 = 1; L_1(x) = x; \text{ and } (m+1)L_{m+1}(x) = (2m+1)xL_m(x) - mL_{m-1}(x).$$

Characterizations and properties of Legendre polynomials can be found in Abramowitz and Stegun (1970). The first three shifted Legendre polynomials used in this paper are given by:

$$Q_1(x) = \sqrt{3}(2x-1), Q_2(x) = \sqrt{5}(6x^2-6x+1), Q_3(x) = \sqrt{7}(20x^3-30x^2+12x-1).$$

For all $\mathbf{x} = (x_1, \dots, x_d)^T \in [0,1]^d$ and for all $\mathbf{m} = (m_1, \dots, m_d)^T \in \mathbb{N}^d$, we write

$$\rho_{\mathbf{m}} := \mathbb{E}\left(\prod_{j=1}^{d} Q_{m_j}(F_j(X_j))\right). \tag{2}$$

The following assumption will be used throughout the paper and says that the copula density c belongs to $\mathbb{L}^2([0,1]^d)$, that is:

$$\int_{[0,1]^d} c^2(u_1, \dots, u_d)\boldsymbol{\mu}^d(du_1, \dots, du_d) < \infty. \tag{3}$$

Assumption (3) is obviously satisfied for any bounded copula density, as the Farlie-Gumbel-Morgenstern and Frank copula densities (see Nelsen (2007) for definitions). When c is the copula density associated to the standard bivariate Gaussian distribution with correlation coefficient $\rho \in (0;1)$, (Beare, 2010) showed that $\|c\|_2^2 = (1-\rho^2)^{-1}$, where $\|.\|_2$ is the $\mathbb{L}^2$ norm on $[0,1]^d$. Moreover, Beare (2010) noted that in the bivariate case, copulas associated to Lancaster type distributions (Lancaster, 1958) satisfy (3). This is the case for bivariate gamma, Poisson, binomial and hypergeometric distributions, and for the compound correlated bivariate Poisson distribution (see for instance Hamdan and Al-Bayyati (1971)). However, copulas exhibiting lower or upper tail dependence (in the sense of McNeil et al. (2015)) do not have square integrable density. In particular, the Gumbel, Clayton, and t-copulas all have upper or lower tail dependence and then do not satisfy condition (3). However we will see in our Simulation study that the proposed nonparametric estimators work very well even if (3) is not satisfied. This is due to the fact that the studied copula have densities in $\mathbb{L}^2([\epsilon, 1-\epsilon]^d)$ for $\epsilon > 0$, and then our estimation procedure gives very good results as long as the variables $F_j(X_j)$ are not too close to zero or 1. But it is important to note that our framework is nonparametric and that the family of the copula is unknown which makes it impossible to verify (3).

Proposition 1. *Assume that (3) holds. Then we have*

$$c(u_1, \dots, u_d) = \sum_{\mathbf{m}\in\mathbb{N}^d} \rho_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j), \tag{4}$$

$$C(u_1, \dots, u_d) = \sum_{\mathbf{m}\in\mathbb{N}^d} \rho_{\mathbf{m}} \prod_{j=1}^{d} I_{m_j}(u_j), \tag{5}$$

where

$$I_{m_j}(u_j) = \int_0^{u_j} Q_{m_j}(x_j)\mu(dx_j).$$

We deduce by Parseval identity that

$$\|c\|_2^2 = \sum_{\mathbf{m}\in\mathbb{N}^d} \rho_{\mathbf{m}}^2 < \infty. \tag{6}$$

It is worth pointing out that the sequence $(\rho_{\mathbf{m}})_{\mathbf{m}\in\mathbb{N}^d}$ characterizes the copula. In this way it will be referred to as the *copula coefficients*. Since $Q_0 = 1$ we have $\rho_{\mathbf{0}} = 1$, where $\mathbf{0}$ stands for the vector with all components equal to zero. The particular case $\rho_{\mathbf{m}} = 0$ for all $\mathbf{m} \neq \mathbf{0}$ coincides with the independent case. As seen in (2), the sequence $(\rho_{\mathbf{m}})_{\mathbf{m}\in\mathbb{N}^d}$ contains all the polynomial correlations between the marginal uniform random variables. In the bivariate case, that is when $d=2$, the element $\rho_{\mathbf{m}}$ simply expresses the correlation between $Q_{m_1}(F_1(X_1))$ and $Q_{m_2}(F_2(X_2))$. In the general case, by orthogonality, we have $\mathbb{E}\left(Q_{m_j}(F_j(X_j))\right) = 0$ for all $m_j > 0$ and then $\rho_{\mathbf{m}} = 0$ as soon as only one component of $\mathbf{m}$ is non-zero. Moreover, if for some integer $j \in \{1, \dots, d\}$, X_j is independent to $(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_d)$, then $\rho_{\mathbf{m}} = 0$ as soon as $m_j > 0$. Then the copula coefficients can be used as an indicator of independence between the components of $\mathbf{X}$. In this sense, we also exhibit a link with the Spearman's rho in Section 3.

For any positive integer vector $\mathbf{N} = (N_1, \dots, N_d)^T$ we define the following $\mathbf{N}$-th order approximations:

$$c^{[\mathbf{N}]}(\mathbf{u}) := \sum_{\mathbf{m}\leq\mathbf{N}} \rho_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j),$$

$$C^{[\mathbf{N}]}(\mathbf{u}) := \sum_{\mathbf{m}\leq\mathbf{N}} \rho_{\mathbf{m}} \prod_{j=1}^{d} I_{m_j}(u_j),$$

where the inequality $\mathbf{m} \le \mathbf{N}$ means that $m_j \le N_j$ for all $j = 1, \ldots, d$. We write $\mathbf{m} \not\le \mathbf{N}$ when $\mathbf{m}$ does not satisfy this inequality. If we observe a n-sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$, of iid random data, with $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})^T$ having joint cdf H, then we can estimate the quantity $\rho_{\mathbf{m}}$ by

$$\widehat{\rho}_{\mathbf{m}} := \begin{cases} 1 & \text{if } \mathbf{m} = \mathbf{0}, \\ 0 & \text{if exactly } d-1 \text{ components of } \mathbf{m} \text{ are zero}, \\ \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\prod_{j=1}^{d} Q_{m_j}(\widehat{F}_j(X_{ij})) & \text{else}, \end{cases}$$

where $\widehat{F}_j(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_{ij} \le x)$.

Remark 1. Since the marginal distributions are continuous, the ties occur with probability zero. So we can also simply write

$$\widehat{\rho}_{\mathbf{m}} = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} Q_{m_j}\left(\frac{R_{ij}}{n}\right),$$

where R_{ij} denotes the rank of X_{ij}.

A $\mathbf{N}$-th order nonparametric estimator of the copula density c is given by

$$\widehat{c}^{[\mathbf{N}]}(u_1, \ldots, u_d) := \sum_{\mathbf{m} \le \mathbf{N}} \widehat{\rho}_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j). \tag{7}$$

In the bivariate case, the estimator $\widehat{c}^{[\mathbf{N}]}(u_1, u_2)$ coincides with the estimator proposed by Gui (2009). We extend this work by introducing a new estimator of the copula C. Indeed, by integration, we get a $\mathbf{N}$-th order nonparametric estimator of the copula function as follows

$$\widehat{C}^{[\mathbf{N}]}(u_1, \ldots, u_d) = \sum_{\mathbf{m} \le \mathbf{N}} \widehat{\rho}_{\mathbf{m}} \prod_{j=1}^{d} I_{m_j}(u_j). \tag{8}$$

Remark 2. In the particular case where the margins $F_1, \ldots, F_d$ are known, the pseudo-estimator of the copula coefficients is given by

$$\tilde{\rho}_{\mathbf{m}} = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij})),$$

and the associated copula and copula density pseudo-estimators are

$$\tilde{c}^{[\mathbf{N}]}(u_1, \ldots, u_d) = \sum_{\mathbf{m} \le \mathbf{N}} \tilde{\rho}_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j),$$

$$\tilde{C}^{[\mathbf{N}]}(u_1, \ldots, u_d) = \sum_{\mathbf{m} \le \mathbf{N}} \tilde{\rho}_{\mathbf{m}} \prod_{j=1}^{d} I_{m_j}(u_j).$$

Remark 3. We can see that $\widehat{c}^{[\mathbf{0}]} = 1$ which coincides exactly with the copula density in the independent case. In such case, we observed in our simulation study that the copula and its estimator are very often equal.

Remark 4. The estimator $\widehat{c}^{[\mathbf{N}]}$ can be negative as an approximation of the expansion of c, and its positivity is not ensured. A way to avoid such negative cases is to consider the maximum $\max(0, \widehat{c}^{[\mathbf{N}]})$ that we can normalize to get

$$\widehat{c}_{+}^{[\mathbf{N}]}(u_1, \ldots, u_d) = \frac{\max(0, \widehat{c}^{[\mathbf{N}]}(u_1, \ldots, u_d))}{\displaystyle\int_{(0,1)^d} \max(0, \widehat{c}^{[\mathbf{N}]}(u_1, \ldots, u_d))\mu(du_1)\cdots\mu(du_d)}.$$

This modification ensures the positivity of $\widehat{c}_{+}^{[\mathbf{N}]}$ which is a dentiy with respect to the Lebesgue measure on $[0, 1]^d$.

In the same way we can normalize $\widehat{C}^{[\mathbf{N}]}$ by using $\min(\max(0, \widehat{C}^{[\mathbf{N}]}), 1)$. Another possibility is to consider $\widehat{C}_{+}^{[\mathbf{N}]} = \min(\int \widehat{c}_{+}^{[\mathbf{N}]}(u_1, \ldots, u_d), 1)$, but this leads to heavier calculations.

3. Some properties of the estimators

3.1. Elementary properties

Proposition 2. *Let $\mathbf{u} \in [0,1]^d$ and fix $\mathbf{N} \in \mathbb{N}^d$. The copula estimator $\widehat{C}^{[\mathbf{N}]}$ given by (8) satisfies the following properties*

(i) *If at least one coordinate of $\mathbf{u}$ is zero then $\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) = 0$.*
(ii) *If $\mathbf{u} = (1,\ldots,1,u_i,1,\ldots,1)$, then $\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) = u_i$.*
(iii) $\widehat{C}^{[\mathbf{0}]}(\mathbf{u}) = \prod_{j=1}^{d} u_j$.

Proposition 3. *The copula density estimator $\widehat{c}^{[\mathbf{N}]}$ given by (7) satisfies the following properties*

(i) $\int_{[0,1]^d} \widehat{c}^{[\mathbf{N}]}(\mathbf{u})\boldsymbol{\mu}^d(d\mathbf{u}) = 1.$
(ii) *For all $u_j \in [0,1]$, $j=1,\ldots,d$, writing $\mathbf{x}_{-j} = (x_1,\ldots,x_{j-1},u_j,x_{j+1},\ldots,x_d)$, we have*

$$\int_{[0,1]^{d-1}} \widehat{c}^{[\mathbf{N}]}(\mathbf{x}_{-j})\boldsymbol{\mu}^{d-1}(\mathrm{d}x_1\cdots\mathrm{d}x_{j-1}\mathrm{d}x_{j+1}\cdots\mathrm{d}x_d) = 1.$$

(iii) $\widehat{c}^{[\mathbf{0}]}(\mathbf{u}) = 1$, *for all $\mathbf{u} \in [0,1]^d$.*

Propositions 2 and 3 imply that the copula estimator is grounded and satisfies the uniform margin property.

Remark 5. Let us recall that for any continuous bivariate random variable (X_1, X_2) with copula C, the Spearman's rho can be express as (see Nelsen (2007)):

$$\rho_C = 12\int_0^1\int_0^1 C(u,v)dudv - 3.$$

In the bivariate case, if (X_1, X_2) is a continuous bivariate random variable with copula C, then the Spearman's rho coincides with the first copula coefficient defined by (2), that is:

$$\rho_C = \rho_{11}.$$

We can immediately deduce an estimator of the Spearman's rho as follows:

$$\hat{\rho}_C = \hat{\rho}_{11} = \frac{3}{n}\sum_{i=1}^{n}\left(2\widehat{F}_1(X_{i1}) - 1\right)\left(2\widehat{F}_2(X_{i2}) - 1\right).$$

This estimator can be compared to the ones given in Pérez and Prieto-Alaiz (2016) but such a study exceeds the scope of this paper.

3.2. Asymptotic properties

The following result shows the proximity between our estimators and those using pseudo-estimators when the margins are known.

Proposition 4. *Fix $\mathbf{N} \in \mathbb{N}^d$, independent of n. For all $\mathbf{u} \in [0,1]^d$, we have*

(i) $\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) = \tilde{c}^{[\mathbf{N}]}(\mathbf{u}) + \mathcal{O}_{\mathbb{P}}\left(\sqrt{n^{-1}}\right)$.
(ii) $\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) = \tilde{C}^{[\mathbf{N}]}(\mathbf{u}) + \mathcal{O}_{\mathbb{P}}\left(\sqrt{n^{-1}}\right)$.

Here the notation $Z_n = \mathcal{O}_{\mathbb{P}}(k_n)$ means that for all $\epsilon > 0$, there exists η and n_0 such that $\mathbb{P}(|Z_n/k_n| > \eta) < \epsilon$, for all $n > n_0$. For any integer $\mathbf{N} \in \mathbb{N}^d$, the notation $\mathbf{N} \to \infty$ means $N_j \to \infty$ for all $j = 1,\ldots,d$, and $\max_{j=1,\ldots,d}(N_j)$ will be denoted by N_{max}.

We now consider the Mean Integrated Squared Error (MISE) as a rule criteria to decide which degree of approximation we use. This criteria is very popular and was also used in Gui (2009) to choose the degree of the bivariate copula density approximation. We write

$$MISE(\widehat{c}^{[\mathbf{N}]}) := \mathbb{E}\|\widehat{c}^{[\mathbf{N}]} - c\|_2^2 = \mathbb{E}\int_{[0,1]^d}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) - c(\mathbf{u})\right)^2\boldsymbol{\mu}^d(d\mathbf{u}).$$

Proposition 5. *Assume that (3) holds. For all $\mathbf{N} \in \mathbb{N}^d$ we have*

$$MISE(\widehat{c}^{[\mathbf{N}]}) = \mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}}(\widehat{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2\right) + \sum_{\mathbf{m}\not\le\mathbf{N}}\rho_{\mathbf{m}}^2.$$

Thus $\mathbf{N}$ represents a smoothing parameter which controls the trade-off between the bias-squared and the variance.

Corollary 1. *Fix $N \in \mathbb{N}^d$, independent of n. Then we have*

$$MISE(\widehat{c}^{[\mathbf{N}]}) = MISE(\widetilde{c}^{[\mathbf{N}]}) + \mathcal{O}(n^{-1}).$$

From now on, we assume that the approximation degree depends of the sample size, that is $\mathbf{N} = \mathbf{N}(n)$, and we consider the assumption

$$(\mathbf{H}) \qquad\qquad N_{max}^{2d+4} = o(n).$$

Corollary 2. *Assume that (6) and **(H)** hold. The copula density estimator $\widehat{c}^{[\mathbf{N}]}$ is consistent in the integrated mean squared sense, that is:*

$$MISE(\widehat{c}^{[\mathbf{N}]}) \xrightarrow[n\longrightarrow\infty]{} 0.$$

Proposition 6. *Assume that (3) and **(H)** hold. For any $\mathbf{u} \in [0,1]^d$, we have as $n \to \infty$*

$$\mathbb{E}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right) \longrightarrow c(\mathbf{u}) \text{ and } \mathbb{E}\left(\widehat{C}^{[\mathbf{N}]}(\mathbf{u})\right) \longrightarrow C(\mathbf{u}).$$

We conclude this section by the asymptotic normality of $\widehat{c}^{[N]}(\mathbf{u})$ that can be deduced from Radulović et al. (2017) (see also Fermanian et al. (2004) for the bivariate case). Fix $\mathbf{u} \in [0,1]^d$ and $\mathbf{N} \in \mathbb{N}^d \backslash \mathbf{0}$ and rewrite

$$\begin{aligned}
c^{[\mathbf{N}]}(\mathbf{u}) &= \sum_{\mathbf{m}\leq\mathbf{N}} \rho_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j) \\
&= \sum_{\mathbf{m}\leq\mathbf{N}} \mathbb{E}\left(\prod_{j=1}^{d} Q_{m_j}\left(F_j(X_j)\right)\right) \prod_{j=1}^{d} Q_{m_j}(u_j) \\
&= \mathbb{E}\left(\sum_{\mathbf{m}\leq\mathbf{N}} \prod_{j=1}^{d} Q_{m_j}\left(F_j(X_{ij})\right) Q_{m_j}(u_j)\right) \\
&= \mathbb{E}\Big(g_{\boldsymbol{u},\boldsymbol{N}}(F_1(X_1),\ldots,F_d(X_d))\Big),
\end{aligned}$$

where $g_{\boldsymbol{u},\boldsymbol{N}}$ is the polynomial function defined on $[0,1]^d$ by

$$g_{\boldsymbol{u},\boldsymbol{N}}(\boldsymbol{v}) := \sum_{\mathbf{m}\leq\mathbf{N}} \prod_{j=1}^{d} Q_{m_j}(v_j) Q_{m_j}(u_j).$$

Similarly write

$$\begin{aligned}
\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) &= \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{\mathbf{m}\leq\mathbf{N}} \prod_{j=1}^{d} Q_{m_j}\left(\widehat{F}_j(X_{ij})\right) Q_{m_j}(u_j)\right) \\
&= \frac{1}{n} \sum_{i=1}^{n} g_{\boldsymbol{u},\boldsymbol{N}}\left(\widehat{F}_{i,1}(X_1),\ldots,\widehat{F}_d(X_{i,d})\right).
\end{aligned}$$

We obtain

$$\begin{aligned}
&\sqrt{n}\left(\widehat{c}^{[bfN]}(\mathbf{u}) - c^{[\mathbf{N}]}(\mathbf{u})\right) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ g_{\boldsymbol{u},\boldsymbol{N}}\left(\widehat{F}_1(X_{i1}),\ldots,\widehat{F}_d(X_{id})\right) - \mathbb{E} g_{\boldsymbol{u},\boldsymbol{N}}\left(F_1(X_{i1}),\ldots,F_d(X_{id})\right)\right\} \\
&= \sqrt{n} \int_{[0,1]^d} g_{\boldsymbol{u},\boldsymbol{N}}(\boldsymbol{v}) d(\mathbb{C}_n - C)(\boldsymbol{v}) \\
&:= \bar{\mathbb{Z}}_n(g_{\boldsymbol{u},\boldsymbol{N}}),
\end{aligned}$$

where

$$\mathbb{C}_n(\boldsymbol{v}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{\widehat{F}_1(X_{i,1}) \leq v_1,\ldots,\widehat{F}_d(X_{i,d}) \leq v_d\}.$$

The process $\bar{\mathbb{Z}}_n(g_{\boldsymbol{u},\boldsymbol{N}})$ is indexed by the functions $g_{\boldsymbol{u},\boldsymbol{N}}$ on the unit hypercube. Radulović et al. (2017) studied such processes for general classes of functions under several conditions. We consider two of them below:

- **Assumption D**. The class $\mathcal{G} := \{g_{\boldsymbol{u},\boldsymbol{N}}; \mathbf{u} \in [0,1]^d, \mathbf{N} \in \mathbb{N}^d\}$ is C-Donsker.
- **Assumption P**. For each $k \in \{1,\dots,d\}$, the kth first-order partial derivative of the copula function C exists and is continuous on the set $\{\mathbf{u} \in [0,1]^d, u_k \in (0,1)\}$.

Theorem 1. *Let Assumption D or P holds. Then the empirical process $\bar{\mathbb{Z}}_n(g_{\boldsymbol{u},\boldsymbol{N}})$ indexed by $\mathcal{G}$ congerves weakly in $\ell^\infty(\mathcal{G})$ to a Gaussian limit.*

4. Numerical analyses and comparisons

In this section, we present simulation results which demonstrate the performance of our approach compared to several estimators. All computations were performed using the R software.

Since the results depend on the degree of the approximations we first present a data-driven method to select **N**.

4.1. Data-driven degree selection

We propose to use a data-driven procedure based on the Least-Squares Cross-Validation (LSCV) to select the optimal parameter $\widehat{\mathbf{N}}_{opt}$. The LSCV procedure has been introduced by Rudemo (1982) and Bowman (1984) to select the smoothing bandwidth for Kernel density estimation and it has been adapted to orthogonal series estimators by Taylor (1990). In the general case, the smoothing parameter is the minimizer of the following function

$$LSCV(\mathbf{N}) = \int_{[0,1]^d} \left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right)^2 d\mathbf{u} - \frac{2}{n}\sum_{i=1}^n \widehat{c}_{-i}^{[\mathbf{N}]}(F_1(X_{i1}),\dots,F_d(X_{id})),$$

which can be estimated by

$$\widehat{LSCV}(\mathbf{N}) = \int_{[0,1]^d} \left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right)^2 d\mathbf{u} - \frac{2}{n}\sum_{i=1}^n \widehat{c}_{-i}^{[\mathbf{N}]}\left(\widehat{F}_1(X_{i1}),\dots,\widehat{F}_d(X_{id})\right), \tag{9}$$

yielding to the following estimator of **N**:

$$\widehat{\mathbf{N}}_{opt} = \underset{\mathbf{N}\in\mathbb{N}^d}{\operatorname{argmin}}\, \widehat{LSCV}(\mathbf{N}),$$

where $\widehat{c}_{-i}^{[\mathbf{N}]}$ is the leave-one-out copula density estimator without the data point $\mathbf{X}_i = (X_{i1},\dots,X_{id})$. Note that the expression (9) has a similar form as the LSCV criterion used by Bouezmarni et al. (2013).

The proposition below gives an abridged form of $\widehat{LSCV}(\mathbf{N})$ which is very useful to decrease its computation time in the numerical study.

Proposition 7. *Assume that (3) holds. We have*

$$\widehat{LSCV}(\mathbf{N}) = \frac{1}{n^2}\sum_{\mathbf{m}\le\mathbf{N}}\left(\sum_{i=1}^n\prod_{j=1}^d Q_{m_j}^2\left(\widehat{F}_j(X_{ij})\right) - \frac{n+1}{n-1}\sum_{k\neq i}\prod_{j=1}^d Q_{m_j}\left(\widehat{F}_j(X_{ij})\right)Q_{m_j}\left(\widehat{F}_j(X_{kj})\right)\right).$$

The form of the $\widehat{LSCV}$ given in Proposition 7 has the advantage to be easily evaluated numerically and it will be used in our simulation to select the value of **N** by minimising this expression.

Proposition 8. *Fix $\mathbf{N} \in \mathbb{N}^d$, independent of n and assume that (3) holds. We have*

$$\mathbb{E}\left(\widehat{LSCV}(N)\right) = MISE(\widehat{c}^{[\mathbf{N}]}) - \|c\|_2^2 + \mathcal{O}(n^{-1}).$$

Remark 6. In the case where the margins are known, we have

$$\mathbb{E}\left(\widehat{LSCV}(N)\right) = MISE(\widehat{c}^{[\mathbf{N}]}) - \|c\|_2^2.$$

Table 1
Relative MIAE ×100 for bivariate copulas. Values in bold indicate the minimum of relative MIAE and underlined values indicate the minimum of standard deviations (in brackets).

Sample size n = 500								
Copulas		Methods						$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	1.78(0.44)	1.63(0.47)	1.75(0.44)	2.62(0.93)	**1.51**(0.68)	1.53(0.50)	2
	Frank	1.76(0.41)	1.63(0.44)	1.74(0.42)	2.58(0.87)	1.50(0.67)	**1.29**(0.45)	1
	Student	1.80(0.45)	1.66(0.48)	1.78(0.45)	2.62(0.95)	**1.53**(0.70)	1.55(0.52)	2
	Gumbel	1.81(0.45)	1.68(0.49)	1.79(0.45)	2.60(0.94)	1.56(0.68)	**1.50**(0.54)	3
	Joe	1.80(0.47)	1.67(0.50)	1.79(0.48)	2.74(0.92)	1.65(0.73)	**1.61**(0.53)	7
	Clayton	1.79(0.47)	1.64(0.49)	1.76(0.47)	2.68(0.88)	1.60(0.68)	**1.53**(0.44)	5
$\tau = 0.55$	Gauss	1.14(0.23)	0.98(0.26)	1.07(0.24)	4.02(0.52)	1.67(0.55)	**0.93**(0.29)	5
	Frank	1.15(0.23)	1.02(0.25)	1.11(0.23)	3.97(0.47)	1.69(0.48)	**0.89**(0.29)	3
	Student	1.17(0.25)	1.01(0.27)	1.11(0.25)	4.0(0.54)	1.68(0.57)	**0.96**(0.30)	5
	Gumbel	1.18(0.25)	1.04(0.28)	1.13(0.25)	3.97(0.57)	1.68(0.30)	**0.99**(0.30)	6
	Joe	1.58(0.26)	**1.02**(0.29)	1.11(0.26)	3.96(0.57)	1.75(0.56)	**1.03**(0.29)	10
	Clayton	1.17(0.27)	**1.01**(0.29)	1.09(0.27)	3.93(0.57)	1.72(0.56)	**1.03**(0.30)	9
$\tau = 0.8$	Gauss	0.31(0.044)	**0.23**(0.051)	0.26(0.044)	4.76(0.092)	1.97(0.098)	0.27(0.050)	10
	Frank	0.51(0.058)	**0.33**(0.072)	0.38(0.057)	4.75(0.10)	1.98(0.11)	0.41(0.073)	8
	Student	0.52(0.064)	**0.33**(0.083)	0.38(0.064)	4.76(0.14)	1.98(0.15)	0.41(0.079)	10
	Gumbel	0.54(0.069)	0.36(0.085)	**0.41**(0.067)	4.75(0.15)	1.97(0.17)	0.43(0.083)	10
	Joe	0.52(0.081)	**0.36**(0.096)	0.40(0.076)	4.71(0.17)	1.97(0.18)	0.44(0.090)	16
	Clayton	0.54(0.077)	**0.35**(0.099)	0.39(0.079)	4.70(0.17)	1.96(0.18)	0.46(0.085)	17
Independent		2.30(0.59)	2.14(0.63)	2.28(0.59)	1.30(0.62)	1.64(0.66)	**4.81**e-**32**(0.00)	0
Sample size n = 1000								
Copulas		Methods						$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	1.25(0.32)	1.18(0.34)	1.24(0.33)	2.61(0.67)	1.26(0.57)	**1.02**(0.39)	3
	Frank	1.25(0.30)	1.82(0.31)	1.24(0.30)	2.57(0.63)	1.28(0.54)	**0.87**(0.39)	2
	Student	1.26(0.33)	**1.20**(0.34)	1.26(0.33)	2.60(0.68)	1.27(0.57)	1.26(0.34)	2
	Gumbel	1.28(0.32)	1.21(0.34)	1.27(0.32)	2.59(0.70)	1.30(0.56)	**1.11**(0.37)	5
	Joe	1.23(0.30)	1.16(0.32)	1.16(0.31)	2.59(0.63)	1.29(0.53)	**1.09**(0.34)	7
	Clayton	1.24(0.31)	1.18(0.32)	1.24(0.31)	2.58(0.65)	1.30(0.51)	**1.16**(0.33)	12
$\tau = 0.55$	Gauss	0.79(0.17)	0.72(0.18)	0.77(0.17)	0.40(0.37)	1.65(0.40)	**0.67**(0.20)	6
	Frank	0.81(0.16)	0.75(0.17)	0.79(0.16)	3.97(0.34)	1.67(0.36)	**0.67**(0.19)	3
	Student	0.80(0.18)	0.74(0.19)	0.79(0.18)	4.00(0.38)	1.60(0.42)	**0.69**(0.21)	5
	Gumbel	0.82(0.17)	0.76(0.18)	0.81(0.17)	3.98(0.40)	1.66(0.43)	**0.75**(**0.19**)	12
	Joe	0.79(0.18)	**0.73**(0.18)	0.77(0.174)	3.91(0.39)	1.64(0.41)	0.74(0.18)	17
	Clayton	0.79(0.18)	**0.72**(0.19)	0.77(0.18)	3.91(0.39)	1.64(0.41)	0.73 (0.19)	13
$\tau = 0.8$	Gauss	0.31(0.04)	**0.23**(**0.05**)	0.26(0.04)	4.76(0.09)	1.97(0.10)	0.27(0.05)	14
	Frank	0.32(0.04)	**0.25**(0.05)	0.27(0.04)	4.74(0.07)	1.97(0.08)	0.28(0.05)	9
	Student	0.32(0.04)	**0.24**(0.05)	0.27(0.05)	4.75(0.10)	1.97(0.10)	0.28(0.05)	13
	Gumbel	0.34(0.05)	**0.26**(0.06)	0.29(0.05)	4.75(0.11)	1.97(0.12)	0.30(0.06)	16
	Joe	0.33(0.05)	0.25(0.06)	**0.28**(0.05)	4.70(0.14)	1.95(0.13)	0.29(0.06)	20
	Clayton	0.33(0.05)	**0.25**(0.06)	0.28(0.05)	4.70(0.11)	1.96(0.12)	0.30(0.06)	20
Independent		1.62(0.43)	1.55(0.45)	1.62(0.44)	0.92(0.45)	1.16(0.49)	**4.81**e-**15**(0.00)	0

4.2. Finite-sample performance of copula estimators

To simplify our numerical study we fix $N_1 = \cdots = N_d := N$. It would be therefore possible to gain in precision by taking the time to choose a best combination of the components of the degree approximation $\mathbf{N}$.

We use Monte-Carlo simulations to demonstrate the potential of the copula estimator $\widehat{C}^{[\mathbf{N}]}$ given in (8) where the optimal parameter $\mathbf{N}$ is selected as described in Section 4.1. We estimate N by its mode over all replications. We compare the performance with four competitors, namely:

- The empirical copula (Deheuvels, 1979), denoted by Emp;
- The empirical checkerboard copula (Carley and Taylor, 2002), denoted by Check;
- The empirical Bernstein copula (Sancetta and Satchell, 2004), with smoothing parameter $k = 10$ and $k = 25$, denoted by Berns10 and Berns25, respectively;
- The empirical beta copula (Segers et al., 2017), denoted by Beta.

We consider the classic copulas below in our simulation:

Table 2
Relative MISE $\times 10^4$ for bivariate copulas. Values in bold indicate the minimum of relative MISE and underlined values indicate the minimum of standard deviations (in brackets).

Sample size n = 500								
Copulas		Methods						$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	3.76(2.15)	3.26(2.14)	3.71(2.17)	2.78(4.33)	2.78(2.55)	**2.58**($\underline{2.01}$)	2
	Frank	3.61(1.88)	3.13(1.88)	3.57(1.89)	7.17(4.46)	7.17(2.63)	**1.77**($\underline{0.16}$)	1
	Student	3.82(2.18)	3.32(2.18)	3.77(2.20)	6.62(4.48)	2.85(2.66)	**2.66**($\underline{2.08}$)	2
	Gumbel	3.80(2.12)	3.32(2.10)	3.76(2.12)	6.78(4.32)	2.90(2.56)	**2.64**($\underline{2.09}$)	3
	Joe	3.77($\underline{2.20}$)	3.30($\underline{2.20}$)	3.72(2.21)	8.06(4.64)	3.22(2.84)	**3.06**(2.21)	7
	Clayton	3.89(2.29)	3.38($\underline{2.27}$)	3.82(2.29)	7.92(4.62)	3.17(2.28)	**2.92**(2.28)	5
$\tau = 0.55$	Gauss	1.86(0.92)	1.53(0.92)	1.81(0.92)	16.0(4.22)	3.45(2.16)	**1.23**($\underline{0.90}$)	5
	Frank	1.79(0.82)	1.48(0.83)	1.75(0.81)	18.2(3.99)	3.99(2.12)	**0.99**($\underline{0.73}$)	3
	Student	1.91(0.96)	1.58(0.97)	1.86(0.96)	16.0(4.36)	3.62(2.26)	**1.28**($\underline{0.95}$)	5
	Gumbel	1.91(0.94)	1.59(0.94)	1.86($\underline{0.93}$)	17.0(4.42)	3.72(2.28)	**1.37**($\underline{0.93}$)	6
	Joe	1.88(1.01)	1.58(1.01)	1.84($\underline{1.00}$)	18.0(4.45)	4.24(2.37)	**1.53**(1.01)	10
	Clayton	1.99(1.14)	1.66(1.15)	1.92(1.14)	4.32(4.72)	4.32(2.52)	**1.59**($\underline{1.13}$)	9
$\tau = 0.8$	Gauss	0.21(0.84)	0.17(0.88)	0.20(0.08)	0.32(1.15)	6.66(0.64)	**0.15**($\underline{0.08}$)	10
	Frank	0.43(0.14)	0.31(0.15)	0.38($\underline{0.13}$)	0.35(1.31)	7.40(0.72)	**0.26**($\underline{0.13}$)	8
	Student	0.47(0.17)	0.33(0.19)	0.42(0.17)	0.33(1.72)	6.78(0.96)	**0.29**($\underline{0.16}$)	10
	Gumbel	0.49(0.17)	0.35(0.19)	0.44(0.17)	0.33(1.84)	6.87(1.02)	**0.32**($\underline{0.16}$)	10
	Joe	0.50($\underline{0.21}$)	**0.38**(0.23)	0.45(0.20)	0.34(1.88)	7.50(1.05)	0.38($\underline{0.21}$)	16
	Clayton	0.51($\underline{0.21}$)	**0.38**(0.24)	0.46(0.22)	0.34(1.95)	7.58(1.08)	0.40($\underline{0.21}$)	17
Independent		5.38(3.10)	4.72(3.10)	5.32(3.10)	1.88(1.90)	2.89(2.50)	**8.02**e-**29**($\underline{0.00}$)	0
Sample size n = 1000								
Copulas		Methods						$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	1.88(1.14)	1.71(1.14)	1.87(1.15)	5.90(3.00)	1.85(1.65)	**1.13**($\underline{1.13}$)	3
	Frank	1.66(1.00)	1.66(0.99)	1.82(1.00)	6.69(3.08)	2.03(1.69)	**0.94**($\underline{0.89}$)	2
	Student	1.90($\underline{1.15}$)	1.73($\underline{1.15}$)	1.89($\underline{1.15}$)	5.99($\underline{1.15}$)	1.89($\underline{1.15}$)	**1.67**($\underline{1.15}$)	2
	Gumbel	1.90($\underline{1.11}$)	1.74($\underline{1.11}$)	1.90($\underline{1.11}$)	6.23(3.06)	1.94(1.66)	**1.46**($\underline{1.11}$)	5
	Joe	1.74(0.98)	1.58(0.97)	1.73(0.98)	6.95(2.83)	1.93(1.54)	**1.40**($\underline{0.97}$)	7
	Clayton	1.91(1.06)	1.74($\underline{1.05}$)	1.90(1.06)	7.03(3.0)	2.03(1.56)	**1.67**($\underline{1.05}$)	12
$\tau = 0.55$	Gauss	0.92(0.50)	0.81(0.50)	0.91(0.50)	1.62(3.00)	3.14(1.49)	**0.66**($\underline{0.49}$)	6
	Frank	0.91(0.42)	0.81(0.42)	0.91(2.85)	0.18(2.85)	3.69(1.46)	**0.56**($\underline{0.37}$)	3
	Student	0.94(0.50)	0.83(0.50)	0.93(4.99)	0.16(3.05)	3.19(1.53)	**0.65**($\underline{0.49}$)	5
	Gumbel	0.96($\underline{0.47}$)	0.85($\underline{0.47}$)	0.95($\underline{0.47}$)	0.17(3.13)	3.35(1.57)	**0.81**($\underline{0.47}$)	12
	Joe	0.90(0.46)	**0.80**(0.46)	0.89(0.46)	0.18(2.94)	3.63(1.51)	**0.80**($\underline{0.46}$)	17
	Clayton	0.02($\underline{0.49}$)	0.01($\underline{0.49}$)	0.95($\underline{0.49}$)	4.00(3.09)	3.74(1.57)	**0.82**($\underline{0.49}$)	13
$\tau = 0.8$	Gauss	0.32(4.39)	0.17(0.09)	0.20(0.08)	0.32(1.53)	6.66(0.64)	**0.15**($\underline{0.08}$)	14
	Frank	0.32(3.91)	0.25(4.52)	0.27($\underline{0.06}$)	0.34(0.90)	7.33(0.50)	**0.13**($\underline{0.06}$)	9
	Student	0.22(0.08)	0.17(0.09)	0.21(0.08)	0.33(1.21)	6.68(0.66)	**0.15**($\underline{0.08}$)	13
	Gumbel	0.23($\underline{0.09}$)	0.19(0.09)	0.22($\underline{0.09}$)	0.33(1.29)	6.82(0.71)	**0.18**($\underline{0.09}$)	16
	Joe	0.23(0.09)	**0.19**(0.10)	0.22($\underline{0.09}$)	0.34(1.26)	7.32(0.71)	**0.19**(0.09)	20
	Clayton	0.24(0.10)	0.19(0.11)	0.23(0.10)	0.34(0.11)	7.47(0.72)	**0.19**($\underline{0.10}$)	20
Independent		2.71(43.0)	2.48(1.6)	2.70(2.0)	0.96(29.0)	1.46(34.0)	**8.02**e-**29**($\underline{0.00}$)	0

- Clayton copula;
- Frank copula;
- Gaussian copula;
- Gumbel copula;
- Independence copula;
- Joe copula;
- Student t-copula with degrees of freedom $\nu = 17$.

The readers may refer to Nelsen (2007) for the explicit functional forms and properties of these copulas. As mentioned in Introduction, Gumbel, Clayton and t-copulas do not satisfy condition (3). However their copula densities belong to $\mathbb{L}^2([\epsilon, 1-\epsilon]^d)$ for $\epsilon > 0$. Then our method could yield estimators of such (truncated) copulas as soon as the components of $\mathbf{u}$ are not too close to the corners of $[0,1]^d$. In our simulation we consider a grid with points from 0.01 to 0.99.

Table 3
Relative MKSE ×100 for bivariate copulas. Values in bold indicate the minimum of relative MKSE and underlined values indicate the minimum of standard deviations (in brackets).

Sample size n = 500								
Copulas			Methods					$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	2.51(0.53)	2.03(0.51)	2.40(0.53)	1.68(0.54)	1.38(0.50)	**1.29**(0.46)	2
	Frank	2.37(0.51)	1.99(0.49)	2.36(0.59)	1.97(0.59)	1.42(0.54)	**0.93**(0.30)	1
	Student	2.41(0.53)	2.03(0.51)	2.39(0.53)	1.72(0.53)	1.39(0.51)	**1.30**(0.46)	2
	Gumbel	2.42(0.54)	2.04(0.52)	2.40(0.53)	1.92(0.45)	1.42(0.50)	**1.41**(0.46)	3
	Joe	2.39(0.53)	2.02(0.52)	2.37(0.53)	2.42(0.40)	**1.53**(0.51)	1.70(0.51)	7
	Clayton	2.41(0.47)	1.64(0.49)	17.6(0.47)	2.68(0.88)	1.60(0.68)	**1.53**(0.54)	5
$\tau = 0.55$	Gauss	2.13(0.45)	1.76(0.44)	2.12(0.45)	3.00(0.41)	1.72(0.47)	**1.23**(0.41)	5
	Frank	2.09(0.43)	1.71(0.42)	2.08(0.43)	3.70(0.42)	1.98(0.49)	**1.01**(0.34)	3
	Student	2.13(0.45)	1.75(0.42)	2.11(0.45)	3.05(0.42)	1.75(0.48)	**1.23**(0.41)	5
	Gumbel	2.11(0.46)	1.73(0.44)	2.09(0.46)	3.35(0.36)	1.86(0.44)	**1.32**(0.43)	6
	Joe	2.10(0.46)	1.74(0.45)	2.09(0.46)	4.11(0.28)	2.18(0.37)	**1.56**(0.45)	10
	Clayton	2.15(0.48)	1.77(0.46)	2.13(0.48)	4.05(0.30)	2.15(0.38)	**1.55**(0.45)	9
$\tau = 0.8$	Gauss	1.09(0.22)	0.91(0.22)	1.08(0.22)	5.30(0.12)	2.70(0.17)	**0.77**(0.20)	10
	Frank	1.51(0.30)	1.21(0.30)	1.49(0.30)	6.28(0.13)	3.31(0.21)	**0.84**(0.25)	8
	Student	1.55(0.31)	1.21(0.31)	1.53(0.31)	5.40(0.18)	2.81(0.24)	**0.94**(0.28)	10
	Gumbel	1.45(0.31)	1.21(0.30)	1.53(0.30)	5.61(0.15)	2.96(0.22)	**0.96**(0.28)	10
	Joe	1.50(0.33)	1.25(0.32)	1.50(0.33)	6.40(0.11)	3.50(0.14)	**1.20**(0.33)	16
	Clayton	1.60(0.33)	1.28(0.34)	1.58(0.34)	6.50(0.11)	3.53(0.15)	**1.25**(0.33)	17
Independent		2.49(0.56)	2.11(0.53)	2.48(0.56)	0.96(0.40)	1.31(0.47)	**1.17**e-**14**(0.00)	0
Sample size n = 1000								
Copulas			Methods					$\widehat{N}_{opt}$
		Emp	Beta	Check	Berns10	Berns25	$\widehat{C}^{[\widehat{N}_{opt}]}$	
$\tau = 0.3$	Gauss	1.70(0.38)	1.51(0.37)	1.70(0.38)	1.55(0.39)	1.08(0.38)	**0.99**(0.35)	3
	Frank	1.68(0.37)	1.49(0.36)	1.68(0.37)	1.90(0.43)	1.18(0.42)	**0.75**(0.32)	2
	Student	1.69(0.38)	1.51(0.37)	1.69(0.38)	1.59(0.39)	1.09(0.39)	**1.07**(0.32)	2
	Gumbel	1.70(0.37)	1.51(0.36)	1.70(0.37)	1.82(0.31)	1.14(0.37)	**1.11**(0.35)	5
	Joe	1.66(0.36)	1.47(0.35)	1.65(0.36)	2.33(0.24)	1.24(0.31)	**1.18**(0.34)	7
	Clayton	1.68(0.37)	1.49(0.36)	1.68(0.37)	1.90(0.43)	1.18(0.42)	**0.75**(0.32)	12
$\tau = 0.55$	Gauss	1.49(0.32)	1.30(0.31)	1.49(0.33)	2.92(0.31)	1.53(0.34)	**0.93**(0.30)	6
	Frank	1.49(0.32)	1.31(0.31)	1.49(0.32)	3.67(0.31)	1.85(0.36)	**0.79**(0.24)	3
	Student	1.49(0.33)	1.29(0.32)	1.48(0.33)	2.95(0.31)	1.56(0.34)	**0.87**(0.29)	5
	Gumbel	1.50(0.31)	1.31(0.31)	1.49(0.32)	3.29(0.26)	1.71(0.31)	**1.17**(0.30)	12
	Joe	1.47(0.32)	1.28(0.31)	1.47(0.32)	4.06(0.19)	2.03(0.23)	**1.23**(0.31)	17
	Clayton	1.52(0.33)	1.33(0.32)	1.51(0.33)	4.00(0.19)	2.00(0.24)	**1.21**(0.32)	13
$\tau = 0.8$	Gauss	1.09(0.22)	0.91(0.22)	1.08(0.22)	5.35(0.12)	2.69(0.17)	**0.77**(0.20)	14
	Frank	1.06(0.20)	0.91(0.20)	1.00(0.20)	6.27(0.093)	3.27(0.15)	**0.64**(0.18)	9
	Student	1.09(0.22)	0.90(0.21)	1.08(0.22)	5.38(0.13)	2.72(0.17)	**0.74**(0.20)	13
	Gumbel	1.10(0.23)	0.93(0.23)	1.10(0.23)	5.60(0.11)	2.90(0.16)	**0.83**(0.22)	16
	Joe	1.09(0.22)	0.93(0.22)	1.08(0.22)	6.38(0.07)	3.43(0.10)	**0.89**(0.22)	20
	Clayton	1.13(0.24)	0.96(0.24)	1.12(0.25)	6.49(0.08)	3.49(0.10)	**0.92**(0.24)	20
Independent		1.77(0.41)	1.58(0.40)	1.76(0.41)	0.68(0.29)	0.94(0.34)	**1.17**e-**14**(0.00)	0

We consider three levels of dependence with kendall's τ's equal to $\tau = 0.3$ (low dependence), $\tau = 0.55$ (middle dependence) and $\tau = 0.8$ (high dependence) for each copula model. We generate in both case iid data using each copula model of sizes $n = 500$ and $n = 1000$.

In order to evaluate the quality of an estimator $\widehat{C}$ for a given copula C, we consider three performances measures: the first one is the mean integrated absolute error (MIAE), defined by

$$\mathrm{MIAE}(\widehat{C}) = \mathbb{E}\left(\int_{[0,1]^d} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) - C(\mathbf{u})\right| d\mathbf{u}\right),$$

the second one is the mean integrated squared error (MISE), defined by

$$\mathrm{MISE}(\widehat{C}) = \mathbb{E}\left(\int_{[0,1]^d} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) - C(\mathbf{u})\right|^2 d\mathbf{u}\right),$$

Table 4
Trivariate copulas: relative MIAE $\times 10^3$, MISE $\times 10^5$ and MKSE $\times 10^3$ with sample size $n = 200$. Values in bold indicate the minimum of relative MIAE, MISE or MKSE and underlined values indicate the minimum of standard deviations (in brackets).

		Sample size n = 200					
Copulas	Methods	$\tau = 0.3$			$\tau = 0.8$		
		MKSE $\times 10^3$	MIAE $\times 10^3$	MISE $\times 10^5$	MKSE $\times 10^3$	MIAE $\times 10^3$	MISE $\times 10^5$
Clayton	Emp	48.4(9.37)	8.94(2.47)	16.1(9.47)	30.5(6.43)	2.87(0.39)	2.71(0.88)
	Beta	**39.9**(9.38)	**8.13**(2.67)	**13.5**(9.28)	**23.9**(7.13)	**1.91**(0.62)	**1.72**(1.17)
	Check	47.9(9.37)	8.87(2.45)	15.8(9.30)	29.6(7.04)	1.94(0.39)	1.87(0.89)
	CN	45.9(8.16)	14.3(1.73)	33.8(9.82)	61.7(0.43)	0.012(0.12)	25.8(0.93)
Joe	Emp	49.5(9.39)	9.06(1.86)	16.0(7.08)	31.4(8.61)	3.81(0.81)	3.63(1.94)
	Beta	40.2(8.84)	8.10(2.06)	13.1(7.10)	24.8(8.32)	**2.82**(0.95)	2.65(1.20)
	Check	48.8(9.44)	**9.01**(1.90)	15.8(7.19)	30.7(8.41)	3.23(0.68)	3.23(1.62)
	CN	**38.0**(8.62)	8.04(2.10)	**13.0**(7.23)	**20.5**(6.99)	3.33(0.90)	**2.45**(1.70)
Gumbel	Emp	47.2(8.67)	8.63(1.38)	14.1(4.92)	30.4(6.23)	3.42(0.56)	3.04(1.18)
	Beta	38.4(8.30)	7.66(1.57)	11.3(4.89)	23.6(7.29)	**2.40**(0.69)	2.06(1.26)
	Check	46.7(8.58)	8.62(1.40)	14.0(4.90)	29.8(5.81)	2.88(0.45)	2.71(9.61)
	CN	**34.6**(7.67)	**7.41** (1.66)	**11.07**(4.95)	**18.4**(5.64)	2.99(0.59)	**1.85**(0.97)
Frank	Emp	47.0(8.97)	8.91(2.35)	15.3(9.53)	29.3(6.85)	3.25(0.55)	3.04(1.18)
	Beta	38.1(8.65)	7.94(2.54)	12.5(9.55)	**24.8**(6.61)	**2.24**(0.65)	**2.06**(1.26)
	Check	46.3(8.92)	8.84(2.32)	15.0(9.42)	28.5(6.50)	2.54(0.46)	2.71(0.96)
	CN	**19.8**(7.20)	**5.98**(2.85)	**7.29**(8.49)	56.9(1.92)	10.6(0.19)	20.3(0.97)
Student	Emp	48.6(8.79)	8.76(1.66)	14.7(6.13)	29.1(6.13)	3.15(0.53)	2.60(0.98)
	Beta	39.3(8.23)	7.76(1.84)	11.7(6.12)	22.7(7.25)	**2.04**(0.63)	1.70(1.20)
	Check	47.9(8.90)	8.73(1.68)	14.5(6.20)	28.8(6.49)	2.44(0.40)	2.25(0.91)
	CN	**27.8**(6.96)	**6.78**(2.12)	**9.02**(6.05)	**16.8**(5.18)	2.74(0.52)	**1.49**(0.87)
Normal	Emp	47.4(7.94)	8.65(1.51)	14.4(5.31)	28.6(5.92)	3.13(0.43)	2.58(0.95)
	Beta	38.4(7.98)	7.65(1.71)	11.5(5.51)	22.5(6.59)	**1.99**(0.60)	1.67(1.08)
	Check	46.8(8.19)	7.65(2.54)	14.2(5.35)	28.1(5.92)	2.35(0.40)	2.21(8.80)
	CN	**24.5**(6.28)	**6.54**(1.96)	**8.08**(5.19)	**16.3**(4.90)	2.71(0.50)	**1.43**(0.77)
Independent		MKSE $\times 10^3$		MIAE $\times 10^3$		MISE $\times 10^5$	
	Emp	49.7(8.09)		8.77(1.50)		14.9(5.40)	
	Beta	41.4(7.23)		7.82(1.63)		12.0(5.24)	
	Check	49.1(8.03)		8.73(1.53)		14.7(5.47)	
	CN	**2.22e-11**(0.00)		**7.99e-13**(0.00)		**4.38e-13**(0.00)	

$\widehat{N}_{opt}$							
τ	Clayton	Joe	Gumbel	Frank	Student	Gauss	Independent
0.3	1	7	6	1	3	2	0
0.8	2	7	7	2	7	7	

and the last one is the mean Kolmogorov-Smirnov error (MKSE) defined by

$$\text{MKSE}(\widehat{C}) = \mathbb{E}\left(\sup_{\mathbf{u}\in[0,1]^d} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) - C(\mathbf{u})\right|\right).$$

We use the following approximations

$$\|\widehat{C}^{[\mathbf{N}]} - C\|_p^2 \approx \frac{1}{T^d} \sum_{j_1,\dots,j_d=1}^{T-1} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{j}/T) - C(\mathbf{j}/T)\right|^p,\ p = 1, 2,$$

$$\sup_{\mathbf{u}\in[0,1]^d} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) - C(\mathbf{u})\right| \approx \sup_{j_1,\dots,j_d=1,\dots,T-1} \left|\widehat{C}^{[\mathbf{N}]}(\mathbf{j}/T) - C(\mathbf{j}/T)\right|,$$

where $\mathbf{j}/T = (j_1/T, \dots, j_d/T)$. We fix a grid such that $j_1, j_2 \in \{1, 2, \dots, 35\}$ and $T = 36$ in the bivariate case, and $j_i/T \in \{0.01, 0.0712, \dots, 0.99\}$ for $i = 1, 2, 3$ in three dimensional case. Finally our numerical results are the average over 1,000 Monte-Carlo replications. All results are normalized, that is: we consider relative MIAE, MISE, and MKSE defined by $\text{MIAE}(\widehat{C})/\|C\|_1$, $\text{MISE}(\widehat{C})/\|C\|_2^2$, and $\text{MKSE}(\widehat{C})/\sup|C|$.

Two-dimensional case. In the case where $d = 2$, Tables 1-3 display the relatives MIAE, MISE and MKSE for the considered two sample sizes and three levels of dependence. It can be observed that the copula estimator $\widehat{C}^{[\mathbf{N}]}$ performs highly than the rest of the estimators overall on these error criteria and for different sample sizes and levels of dependence. It is important to precise some comments on the individual comparisons:

Table 5
Trivariate copulas: relative MIAE$\times 10^3$, MISE$\times 10^5$ and MKSE$\times 10^3$ with sample size $n = 500$. Values in bold indicate the minimum of relative MIAE, MISE or MKSE and underlined values indicate the minimum of standard deviations (in brackets).

Copulas	Methods	**Sample size n = 500** $\tau = 0.3$ MKSE $\times 10^3$	MIAE $\times 10^3$	MISE $\times 10^5$	$\tau = 0.8$ MKSE $\times 10^3$	MIAE $\times 10^5$	MISE $\times 10^5$
Clayton	Emp	30.4(5.86)	5.61(1.39)	6.28(3.21)	19.8(4.16)	1.60(0.27)	0.85(0.43)
	Beta	**27.1**(5.42)	**5.23**(1.48)	**5.57**(3.26)	**17.0**(4.23)	**1.20**(0.35)	**0.70**(0.49)
	Check	30.4(5.90)	5.58(1.38)	6.22(3.20)	19.6(4.16)	1.29(0.27)	0.80(0.44)
	CN	29.7(7.65)	6.73(1.40)	8.35(3.94)	60.7(0.25)	12.0(0.06)	25.3(0.63)
Joe	Emp	31.9(6.27)	5.86(1.3)	6.76(3.26)	19.8(4.10)	2.24(0.31)	1.33(0.46)
	Beta	28.3(5.83)	5.50(1.34)	6.02(3.11)	17.0(3.94)	**1.83**(0.37)	**1.05**(0.49)
	Check	31.7(6.20)	5.85(1.31)	6.74(3.26)	19.6(4.06)	2.07(0.31)	4.88(0.46)
	CN	**25.7**(7.20)	**5.20**(1.53)	**5.57**(3.80)	**15.5**(4.12)	2.10(0. 460)	1.10(0.62)
Gumbel	Emp	30.5(6.28)	5.80(1.35)	6.48(3.31)	18.8(3.74)	2.00(0.27)	1.09(0.36)
	Beta	27.1(5.76)	5.44(1.43)	5.76(3.32)	15.8(3.45)	**1.60**(0.33)	0.83(0.39)
	Check	30.2(6.18)	5.79(1.36)	6.45(3.14)	18.6(3.76)	1.82(0.27)	1.04(0.36)
	CN	**24.3**(5.37)	**5.06**(1.48)	**5.07**(3.66)	**14.3**(3.49)	1.85(0.33)	**0.82**(0.37)
Frank	Emp	29.1(5.15)	5.5(0.98)	5.81(2.15)	17.8(3.32)	1.82(0.22)	**1.09**(0.360)
	Beta	25.9(4.68)	5.16(1.03)	5.07(2.09)	**15.5**(3.15)	1.42(0.26)	0.83(0.39)
	Check	28.9(5.07)	5.55(0.98)	5.77(2.14)	17.7(3.31)	1.61(0.21)	1.04(0.36)
	CN	**15.8**(3.34)	**4.52**(1.12)	**3.70**(2.04)	18.3(2.12)	**0.34**(0.17)	2.13(0.24)
Student	Emp	30.0(6.27)	5.51(1.034)	5.81(2.40)	18.7(3.97)	1.84(0.27)	0.99(0.36)
	Beta	26.9(5.71)	5.10(1.08)	5.06(2.35)	15.6(3.97)	**1.40**(0.31)	0.75(0.36)
	Check	29.9(6.19)	5.48(1.033)	5.77(2.40)	18.7(3.97)	1.61(0.26)	0.95(0.35)
	CN	**18.6**(4.89)	**4.40**(1.25)	**3.76**(2.24)	**12.4**(3.45)	1.63(0.34)	**0.64**(0.38)
Normal	Emp	30.0(6.01)	5.46(1.06)	5.78(2.45)	18.6(4.16)	1.81(0.26)	0.99(0.37)
	Beta	26.8(5.80)	5.05(1.11)	5.03(2.42)	15.6(4.01)	**1.34**(0.32)	0.74(0.39)
	Check	29.9(6.00)	5.44(1.07)	5.75(2.47)	18.5(4.13)	1.55(0.26)	0.94(0.36)
	CN	**17.9**(5.03)	**4.44**(1.52)	**3.31**(3.31)	**13.8**(3.65)	1.66(0.33)	**0.68**(0.41)
Independent		MKSE $\times 10^3$		MIAE $\times 10^3$		MISE $\times 10^5$	
	Emp	31.3(5.51)		5.59(1.03)		6.09(2.43)	
	Beta	27.9(5.48)		5.20(1.06)		5.33(2.31)	
	Check	31.1(5.42)		5.57(1.03)		6.05(2.43)	
	CN	**2.22**e-**13**(0.00)		**7.34**e-**15**(0.00)		**3.67**e-**29**(0.00)	
N_{opt}							
τ	Clayton	Joe	Gumbel	Frank	Student	Gauss	Independent
0.3	2	9	8	1	3	3	0
0.8	2	12	12	4	10	13	

- For copulas with high dependence (Kendall's $\tau = 0.8$), the empirical beta copula can give better results than $\widehat{C}^{[\mathbf{N}]}$, but with large standard deviation in term of relative MIAE regardless of the sample size. In that case, Bernstein copulas with smoothing parameter $k = 10$ and $k = 25$ are really worse.
  Finally an automatic selection method is required to be able to detect the best Bernstein copula but the computational cost was too heavy here.
- For the independent copulas, the degree parameter $N = 0$ is chosen all the time. In that case $\widehat{C}^{[\mathbf{N}]} = \widehat{C}^{[\mathbf{0}]}$ coincides with the true copula and it largely dominates its competitors in all scenarios.
- Moreover, we can see that our approach gives much better results compared to other estimators with minimal standard deviation in terms of relative MISE and MKSE.
- We also remark that our method requires only small order of shifted Legendre polynomial $\mathbf{N}$. This smoothing parameter increases with both the sample size (n) and the level of dependence (τ).

Three dimensional case. In the case where $d = 3$, Tables 4 and 5 report the relative MIAE, MISE and MKSE for $n = 200$ and $n = 500$ and for Kendall's $\tau = 0.30$ and 0.8, respectively. Our results are based on 100 Monte-Carlo replications. With the exception of the Clayton copula where the beta estimator gave the best result, the copula estimator $\widehat{C}^{[\mathbf{N}]}$ dominates notably all the performances with an extremely good results for the independence case since the choice $\mathbf{N} = 0$ is almost always chosen and then the estimator fits exactly the density.

Table 6
Relative MISE $\times 100$ for bivariate copula densities. Values in bold indicate the minimum of relative MISE and underlined values indicate the minimum of standard deviations (in brackets).

Sample size n = 500										
Copula		Methods								
		$\widehat{c}_{Mr}$	$\widehat{c}_{Bk}$	$\widehat{c}_{wa}$	$\widehat{N}_{opt}$	$\widehat{c}^{[\widehat{N}_{opt}]}$	$\widehat{c}^1_{Pt}$	$\widehat{c}^2_{Pt}$	$\widehat{c}_{Be10}$	$\widehat{c}_{Be25}$
$\tau = 0.3$	Clayton	48.3(1.4)	31.3(3.8)	68.5($\underline{0.67}$)	5	15.0(4.3)	23.0(5.4)	**7.65**(3.35)	38.7(2.1)	22.9(4.8)
	Joe	63.3($\underline{1.2}$)	43.7(3.7)	61.7(1.7)	7	**9.1**(5.0)	35.9(5.7)	15.2(4.70)	51.7(2.0)	34.2(4.9)
	Gumbel	42.6($\underline{1.4}$)	26.9(3.1)	49.5(1.8)	3	25.2(1.7)	19.4(5.0)	**7.81**(4.05)	32.3(2.1)	19.3(3.9)
	Frank	3.0(1.1)	2.1(0.75)	30.4(1.9)	1	**1.9**($\underline{0.23}$)	2.8(1.2)	2.82(2.33)	1.9(0.78)	2.8(1.0)
	Student(df=17)	17.0(1.6)	9.1(2.1)	38.5(1.5)	2	9.6($\underline{0.87}$)	4.5(2.5)	**2.03**(1.35)	10.6(1.6)	6.2(2.6)
	Gauss	12.4(1.4)	6.2(1.8)	35.9(1.3)	2	**6.2**($\underline{0.82}$)	3.2(2.0)	**1.80**(1.58)	7.1(1.5)	4.2(2.0)
$\tau = 0.55$	Clayton	75.6(0.99)	53.5(3.1)	92.1(0.13)	9	26.6(3.0)	55.6(5.7)	**18.1**(5.52)	74.6($\underline{0.90}$)	58.0(2.0)
	Joe	78.6(0.97)	56.8(2.8)	79.6($\underline{0.78}$)	10	30.0(3.4)	58.3(4.7)	**20.8**(5.06)	76.6(0.84)	61.2(2.1)
	Gumbel	63.9(1.4)	39.1(3.9)	70.1($\underline{1.1}$)	6	24.2(3.6)	39.9(6.7)	**13.65**(4.32)	61.4(1.3)	43.4(3.1)
	Frank	4.2(1.1)	2.9(0.99)	32.4(1.2)	3	**1.6**($\underline{0.57}$)	2.5(1.2)	6.28(3.37)	6.3(1.1)	2.9(1.1)
	Student	37.6(1.8)	16.5(3.6)	58.3($\underline{1.1}$)	5	**9.8**(2.6)	14.5(5.2)	**1.40**(1.12)	35.7(1.8)	18.1(3.3)
	Gauss	31.6(1.8)	12.3(3.4)	54.0($\underline{1.2}$)	5	7.0(2.4)	10.5(5.1)	**0.87**(0.78)	30.0 (1.7)	13.7(3.4)
$\tau = 0.8$	Clayton	83.6(1.1)	70.3(2.1)	98.1($\underline{0.02}$)	17	57.7(1.0)	80.7(3.0)	**21.3**(5.48)	91.2(0.18)	84.0(0.40)
	Joe	84.5(0.92)	70.6(1.8)	91.5($\underline{0.19}$)	16	61.5(0.87)	81.4(2.98)	**22.4**(5.01)	91.4(0.18)	84.3(0.38)
	Gumbel	71.1(1.8)	49.5(3.2)	85.8($\underline{0.36}$)	10	45.2(1.1)	67.3(4.6)	**14.5**(3.61)	83.6(0.36)	71.1(0.79)
	Frank	6.2(1.7)	5.4(2.18)	38.3($\underline{0.65}$)	8	5.3(3.0)	**4.9**(2.1)	15.1(7.66)	28.7(0.91)	12.9(1.5)
	Student	50.5(2.5)	22.4(3.9)	80.9($\underline{0.46}$)	10	15.4(2.5)	43.3(7.5)	**0.94**(0.75)	70.4(0.68)	50.5(1.6)
	Gauss	45.4(2.7)	17.3(3.7)	78.3($\underline{0.48}$)	10	11.0(2.2)	38.3(8.8)	**0.58**(0.55)	67.3(0.74)	46.1(1.7)
Sample size n = 1000										
Copula		Methods								
		$\widehat{c}_{Mr}$	$\widehat{c}_{Bk}$	$\widehat{c}_{wa}$	$\widehat{N}_{opt}$	$\widehat{c}^{[\widehat{N}_{opt}]}$	$\widehat{c}^1_{Pt}$	$\widehat{c}^2_{Pt}$	$\widehat{c}_{Be10}$	$\widehat{c}_{Be25}$
$\tau = 0.3$	Clayton	47.0(1.00)	27.4(3.20)	67.1($\underline{0.38}$)	12	**6.40**(2.70)	19.8(4.60)	6.93(2.82)	38.5(1.50)	21.8(3.60)
	Joe	62.0($\underline{1.00}$)	39.0(3.10)	60.6(1.20)	7	**5.70**(2.90)	31.4(5.40)	13.4(3.54)	51.5(1.50)	32.9(3.10)
	Gumbel	41.3($\underline{0.99}$)	23.1(2.40)	47.1(1.20)	5	11.2(2.40)	16.5(4.00)	**6.43**(2.43)	31.9(1.40)	17.7(3.00)
	Frank	2.1(0.61)	1.30(0.45)	27.2(0.86)	2	**0.46**($\underline{0.29}$)	1.90(0.66)	1.64(0.93)	1.40(0.54)	1.50(0.54)
	Student	15.7(1.20)	7.20(1.60)	35.8(0.88)	2	9.20($\underline{0.54}$)	3.50(1.90)	**1.07**(0.73)	10.2(1.30)	4.70(1.70)
	Gauss	11.3(0.88)	5.00(1.20)	33.3(0.86)	3	3.80($\underline{0.82}$)	2.30(1.40)	**0.93**(0.67)	6.90(1.00)	3.30(1.30)
$\tau = 0.55$	Clayton	73.8(0.78)	49.1(2.5)	91.9($\underline{0.07}$)	13	23.3(2.10)	50.5(5.70)	**15.6**(5.08)	74.5(0.63)	56.0(1.6)
	Joe	77.1(0.89)	52.3(2.20)	79.3($\underline{0.52}$)	17	21.8(4.20)	53.5(5.90)	**17.6**(4.56)	76.5(0.5603)	60.7(1.20)
	Gumbel	61.8(1.00)	33.8(2.80)	69.5($\underline{0.73}$)	12	**10.3**(2.7)	35.5(5.80)	12.5(2.69)	61.1(0.93)	42.0(1.90)
	Frank	3.30(0.68)	2.00(0.63)	30.4(0.62)	3	**1.28**($\underline{0.30}$)	1.90(0.87)	4.2(2.48)	6.00(0.82)	2.30(0.71)
	Student(df=17)	34.9(1.60)	12.9(2.80)	56.9($\underline{0.76}$)	5	8.10(1.60)	12.5(4.50)	**1.00**(0.65)	35.5(1.40)	17.6(2.30)
	Gauss	29.20(1.40)	9.70(2.50)	52.7($\underline{0.80}$)	6	**3.50**(1.40)	8.10(3.70)	4.72(0.37)	29.9(1.20)	13.7(2.30)
$\tau = 0.8$	Clayton	81.9(1.00)	67.1(1.80)	98.1($\underline{0.02}$)	20	46.6(1.40)	78.4(2.80)	**19.3**(4.49)	9.2(0.12)	84.0(0.28)
	Joe	83.0(1.00)	67.9(1.80)	91.4($\underline{0.14}$)	20	47.5(1.2)	79.1(3.00)	**19.1**(5.72)	91.4(0.12)	84.3(0.31)
	Gumbel	68.2(1.50)	45.7(2.40)	85.6($\underline{0.23}$)	16	38.2(1.60)	63.5(4.90)	**13.3**(2.66)	83.5(0.23)	70.8(0.53)
	Frank	4.80(1.10)	3.90(1.80)	37.7($\underline{0.42}$)	9	**3.6**(1.7)	3.97(1.60)	11.1(4.86)	28.7(0.65)	12.9(1.10)
	Student	45.8(2.30)	18.6(3.00)	80.5($\underline{0.30}$)	13	12.6(1.40)	39.6(7.10)	**4.88**(10.62)	70.2(0.46)	50.4(1.10)
	Gauss	41.1(2.20)	14.4(2.70)	78.1($\underline{0.34}$)	14	8.90(1.60)	32.5(6.70)	**0.26**(0.26)	67.2(0.54)	46.2(1.20)
Independent		0.15(0.10)	0.62(0.22)	25.90(0.88)	0	**0.00**($\underline{0.00}$)	1.90(0.64)	0.79(0.46)	0.44(0.19)	1.70(0.51)

4.3. Finite-sample performance of copula density estimators

Analogously to Section 4.2, we run Monte-Carlo simulations to evaluate the performance of the projection estimator of copula density $\widehat{c}^{[\mathbf{N}]}$ given by (7). A finite-sample comparison with the following recent developments in copula density estimators are considered.

- The probit-transformation estimator studied in Geenens et al. (2017), denoted by $\widehat{c}^1_{Pt}$ for the local log-linear estimator and by $\widehat{c}^2_{Pt}$ for the local log-quadratic estimator;
- The Bernstein copula density estimator studied in (Bouezmarni et al., 2013; 2010; Janssen et al., 2014), denoted by $\widehat{c}_{Be10}$ and $\widehat{c}_{Be25}$ with smoothing parameter $k = 10$ and $k = 25$, respectively;
- The wavelet estimator studied in Genest et al. (2009), denoted by $\widehat{c}_{wa}$, with Haar wavelets;
- The beta kernel estimator studied in Charpentier et al. (2007), denoted by $\widehat{c}_{Bk}$;
- The Mirror reflection kernel estimator studied in Gijbels and Mielniczuk (1990), denoted by $\widehat{c}_{Mr}$.

The functions $\widehat{c}^1_{Pt}$, $\widehat{c}^2_{Pt}$, $\widehat{c}_{Ph}$, $\widehat{c}_{Be}$ and $\widehat{c}_{Bk}$ are provided in the R package **kdecopula** (Nagler, 2018).

Table 7
Relative MKSE ×100 for bivariate copula densities. Values in bold indicate the minimum of relative MKSE and underlined values indicate the minimum of standard deviations (in brackets).

Sample size n = 500										
Copula		Methods								
		$\widehat{c}_{Mr}$	$\widehat{c}_{Bk}$	$\widehat{c}_{wa}$	$\widehat{N}_{opt}$	$\widehat{c}^{[\widehat{N}_{opt}]}$	$\widehat{c}^1_{Pt}$	$\widehat{c}^2_{Pt}$	$\widehat{c}_{Be10}$	$\widehat{c}_{Be25}$
$\tau = 0.3$	Clayton	88.9(1.10)	71.5(4.30)	100(0.00)	5	43.4(9.20)	60.6(7.60)	**33.41**(8.50)	79.8(2.20)	60.2(6.80)
	Joe	92.5(0.82)	76.8(3.30)	85.6(1.60)	7	**24.8**(13.3)	69.4(5.60)	44.2(7.54)	83.6(1.70)	67.2(4.90)
	Gumbel	89.5(1.10)	71.3(4.10)	81.4(2.30)	3	67.1(3.50)	59.9(8.00)	**35.6**(10.0)	78.5(2.60)	59.2(6.50)
	Frank	46.5(5.20)	29.7(9.60)	100(0.00)	1	**23.4** (3.30)	43.4(20.0)	64.9(34.1)	29.2(10.9)	44.3(18.3)
	Student	77.5(2.40)	58.6(6.90)	100(0.00)	2	60.3(4.20)	38.0(12.1)	**30.0**(14.9)	64.4(4.60)	48.6(14.6)
	Gauss	73.2(2.70)	52.6(8.50)	100(0.00)	2	54.4(5.2)	30.5(14.7)	**27.11**(14.9)	59.0(6.00)	42.1(15.4)
$\tau = 0.55$	Clayton	90.7(0.61)	76.3(2.10)	100(0.00)	9	53.6(3.10)	78.0(4.00)	**44.0**(7.85)	89.9(0.57)	79.4(1.40)
	Joe	91.9(0.58)	78.2(1.90)	91.4(0.5)	10	56.6(3.20)	79.4(3.20)	**47.1**(6.59)	90.5(0.52)	81.1(1.30)
	Gumbel	88.3(0.96)	69.3(3.60)	87.7(0.88)	6	53.2(4.40)	70.5(5.90)	**40.8**(7.15)	86.2(0.97)	72.9(2.70)
	Frank	46.3(5.60)	29.7(11.4)	100(0.00)	3	**14.4**(6.40)	26.1(12.0)	73.2(24.4)	39.5(6.80)	30.6(13.0)
	Student	77.9(2.00)	53.4(6.60)	100(0.00)	5	39.0(8.90)	49.5(8.40)	**16.0**(7.58)	74.8(2.10)	56.2(5.90)
	Gauss	75.3(2.40)	48.6(7.50)	100(0.00)	5	33.8(10.3)	44.0(11.4)	**12.1**(6.82)	72.2(2.30)	52.0(7.40)
$\tau = 0.8$	Clayton	92.6(0.58)	85.0(1.20)	100(0.00)	17	77.1(0.69)	91.0(1.60)	**46.1**(7.05)	96.2(0.10)	92.5(0.22)
	Joe	93.1(0.47)	85.2(1.10)	96.5(0.09)	16	79.6(0.58)	91.3(1.60)	**47.5**(6.37)	96.3(0.09)	92.6(0.21)
	Gumbel	87.8(1.00)	74.2(2.40)	94.1(0.21)	10	71.3(0.87)	85.6(2.60)	**39.9**(5.47)	93.6(0.21)	87.3(0.53)
	Frank	0.39.6(6.90)	31.9(13.6)	100(0.00)	8	32.2(15.1)	**23.0**(9.70)	72.9(23.1)	56.7(2.40)	37.2(7.30)
	Student	75.5(2.00)	52.7(4.60)	100(0.00)	10	43.8(4.40)	70.2(6.00)	**11.5**(5.61)	86.8(0.51)	74.5(1.40)
	Gauss	72.0(2.20)	47.3(5.30)	100(0.00)	10	38.2(5.10)	66.5(7.50)	**8.99**(4.99)	85.3(0.58)	71.7(1.80)
Independent		17.0(5.80)	38.5(11.6)	1.00.7(2.70)	0	**0.00**(0.00)	133(43.9)	74.9(38.5)	47.3(16.9)	102.7(43.5)
Sample size n = 1000										
Copula		Methods								
		$\widehat{c}_{Mr}$	$\widehat{c}_{Bk}$	$\widehat{c}_{wa}$	$\widehat{N}_{opt}$	$\widehat{c}^{[\widehat{N}_{opt}]}$	$\widehat{c}^1_{Pt}$	$\widehat{c}^2_{Pt}$	$\widehat{c}_{Be10}$	$\widehat{c}_{Be25}$
$\tau = 0.3$	Clayton	87.9(0.84)	67.1(3.90)	100(000)	12	**14.4**(8.80)	56.4(7.00)	32.3(7.62)	79.7(1.50)	59.3(5.10)
	Joe	91.7(0.69)	72.6(2.90)	85.4(1.00)	7	**19.7**(10.5)	64.5(5.80)	41.9(5.94)	83.5(1.20)	66.4(3.20)
	Gumbel	88.5(0.85)	66.3(3.60)	80.4(1.50)	5	40.8(7.3)	55.8(7.0)	**33.4**(7.59)	78.2(1.70)	57.6(5.20)
	Frank	43.8(3.67)	26.3(8.10)	100(000)	2	**12.4**(5.40)	37.0(16.5)	54.2(23.6)	27.7(9.10)	33.5(11.6)
	Student	75.6(2.10)	52.5(6.50)	100(00)	2	59.0(2.6)	34.7(11.4)	**20.5**(10.5)	62.6(4.10)	44.2(9.70)
	Gauss	71.4(2.10)	48.4(7.40)	100(000)	3	42.1(7.90)	26.9(12.8)	**19.6**(9.52)	58.2(4.40)	40.8(11.7)
$\tau = 0.55$	Clayton	89.6(0.49)	73.2(1.80)	100(000)	13	50.3(2.30)	74.4(4.20)	**40.9**(7.67)	89.8(0.39)	79.4(1.10)
	Joe	91.0(0.54)	75.2(1.50)	91.3(0.34)	17	47.4(4.90)	76.2(4.20)	**43.5**(6.34)	90.5(0.34)	80.8(0.79)
	Gumbel	86.9(0.69)	64.6(2.70)	87.6(0.63)	12	**33.3**(5.00)	66.7(5.30)	39.6(4.42)	86.0(0.67)	71.9(1.70)
	Frank	43.9(4.20)	26.7(10.2)	100(000)	3	**11.1**(4.50)	25.2(11.4)	60.4(20.5)	39.1(4.90)	29.5(10.5)
	Student	75.3(1.96)	47.5(5.80)	100(000)	5	34.2 (6.50)	46.3(8.80)	**13.8**(5.84)	74.3(1.70)	54.9(4.2)
	Gauss	72.7(2.00)	43.2(6.60)	100(0.00)	6	21.2(9.10)	39.0(9.50)	**9.10**(4.13)	71.7(1.70)	50.8(5.00)
$\tau = 0.8$	Clayton	91.7(0.53)	83.1(1.10)	100(0.00)	20	69.4(1.10)	89.7(1.60)	**44.1**(5.61)	96.2(0.07)	92.5(0.16)
	Joe	92.3(0.54)	83.6(1.10)	96.5(0.07)	20	70.1(0.96)	90.1(1.70)	**43.5**(7.77)	96.3(0.06)	92.6(0.17)
	Gumbel	86.3(0.83)	71.4(1.90)	94.1(0.14)	16	65.4(1.40)	83.4(2.96)	**38.4**(4.21)	93.5(0.13)	87.1(0.31)
	Frank	36.2(6.00)	27.5(12.4)	100(0.00)	9	**27.4**(12.9)	23.4(9.90)	62.0(16.86)	56.5(1.70)	37.5(5.20)
	Student	72.2(1.70)	48.0(4.30)	100(0.00)	13	39.5(3.20)	67.1(6.20)	**10.62**(4.26)	86.6(0.40)	74.2(0.97)
	Gauss	69.0(1.70)	43.5(4.70)	78.1(0.34)	14	33.4(4.20)	61.6(6.20)	**5.88**(3.43)	85.2(0.44)	71.5(1.10)
Independent		18.1(4.90)	34.4(11.3)	100(0.00)	0	**0.00**(0.00)	115.8(33.5)	52.32(23.9)	34.2(11.8)	72.6(25.7)

Note that we are not looking at the thresholding (Autin et al. (2010)) and penalised hierarchical B-splines (Kauermann et al. (2013)) approaches due to the slowness of their calculation which makes them less competitive. We consider the same classic copulas and performance measures described above but narrowing to 100 Monte-Carlo replications.

Table 6 reports the relative MISE and Table 7 reports the relatives MKSE for the considered two sample sizes and three levels of dependence. They show that two approaches clearly outperform all the competitors considered for all sample sizes and levels of dependence: the approach based on *copula coefficients* and the local log-quadratic $\widehat{c}^2_{Pt}$ proposed by Geenens et al. (2017). It appears that the degree parameter (truncation order) of $\widehat{c}^{[N]}$ has an influence on his performance. It increases when the level of dependence approaches one. Note that the maximum selected degree, in all scenarios of the simulation study is $N = 20$. But in the majority of cases this is between $N = 0$ (for the independent case) and $N = 10$. Globally the estimator proposed by Geenens et al. (2017) has high performances when τ is greater. If condition (3) is not satisfied it should be preferable to use this estimator when (u_1, u_2) is close to the corner of (0,0) or (1,1). However Geenens et al. (2017) only treated the bivariate case both in their paper and in the package proposed by Nagler (2018).

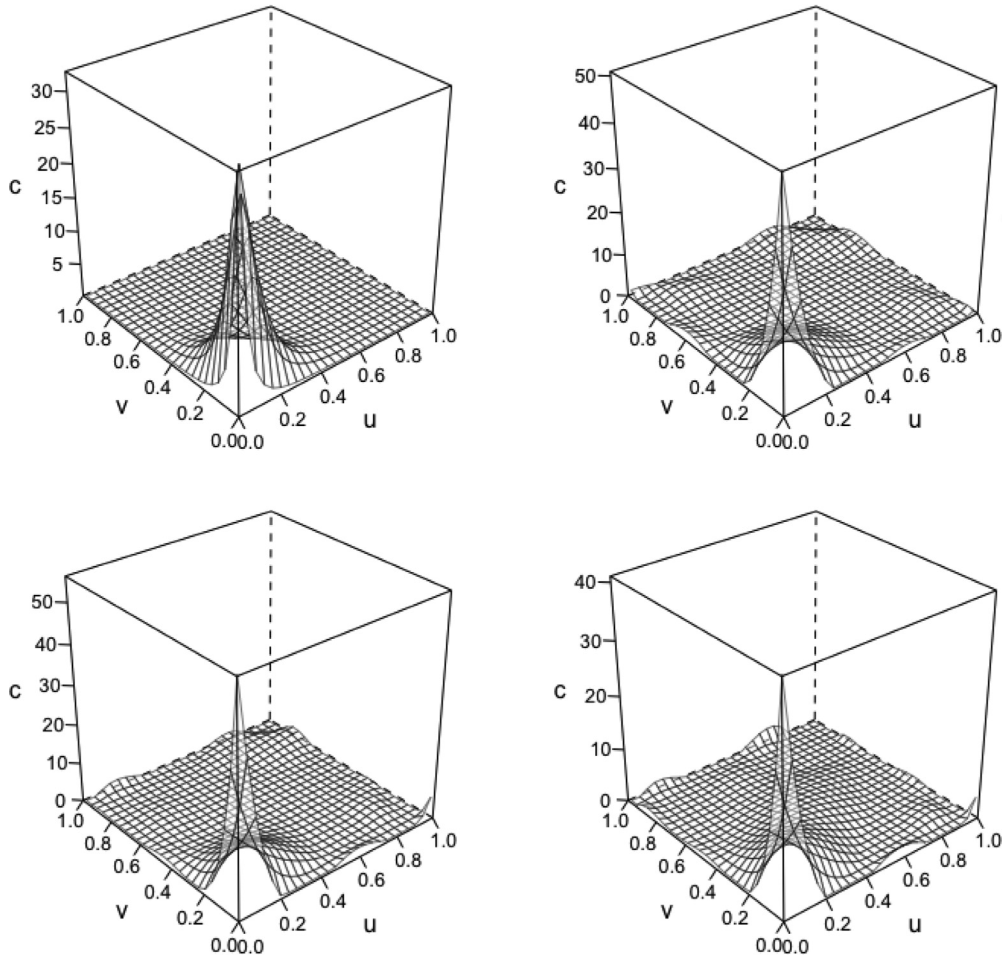

Fig. 1. Three dimensional gumbel density $c(u, v, w)$ with parameter $\theta = 4$ on a $26 \times 26 \times 1$ grid where $w = 0.1$ and its potential approximations with sample size $n = 1000$. Upper: true density function (left) and $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 4, 4)$ (right). Bottom: $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 5, 4)$ (left) and $\mathbf{N} = (4, 5, 1)$(right).

4.4. Three-dimensional illustration

We propose here to illustrate graphically the three-dimensional density estimation by fixing one coordinate, that is we represent $\widehat{c}^{[\mathbf{N}]}(x_1, x_2, x_3)$ when x_3 is fixed. We also let free the three components of the approximation degree $\mathbf{N} = (N_1, N_2, N_3)$ to show the potential of our method. We consider the Gumbel copula density of parameter $\theta = 4$. This copula is very challenging since the Gumbel copula density is heavy-tailed. It is not defined at points (x_1, x_2, x_3) when one component is equal to zero, or when $(x_1, x_2, x_3) = (1, 1, 1)$, and it does not satisfy condition (3). We represent our estimators $\widehat{c}^{[\mathbf{N}]}(x_1, x_2, x_3)$ when x_3 is fixed and equal to 0.1 (close to zero), 0.5, and 0.9 (close to one). We also let vary the components of $\mathbf{N}$, considering three cases: i) the automatic choice when $N_1 = N_2 = N_3$ which gives $\mathbf{N} = (4, 4, 4)$, ii) the automatic choice with different components which gives here $\mathbf{N} = (4, 5, 4)$, iii) and another arbitrary value $\mathbf{N} = (4, 5, 1)$.

Figures 1 -3 represent the copula densities estimations obtained. The estimation procedure can be improved by letting the components free, but with a cost of calculus, passing to a number K of possibilities at K^d, where K is the maximum degree approximation. We can observe the estimations are deteriorated at the edges of the domain $[0, 1]^2$.

4.5. Extreme value copulas

Here we chose less regular bivariate copulas families such as Marshall-Olkin, Tawn, Galambos and Hüsler-Reiss extreme value copulas (see McNeil et al. (2015)) with a sample size $n = 100$ and Kendall's $\tau = 0.3$. Table 8 contains relative MIAE, MISE and MKSE. The empirical copula seems to be the least efficient. On the other hand, $C^{[N]}$ has a very good performance.

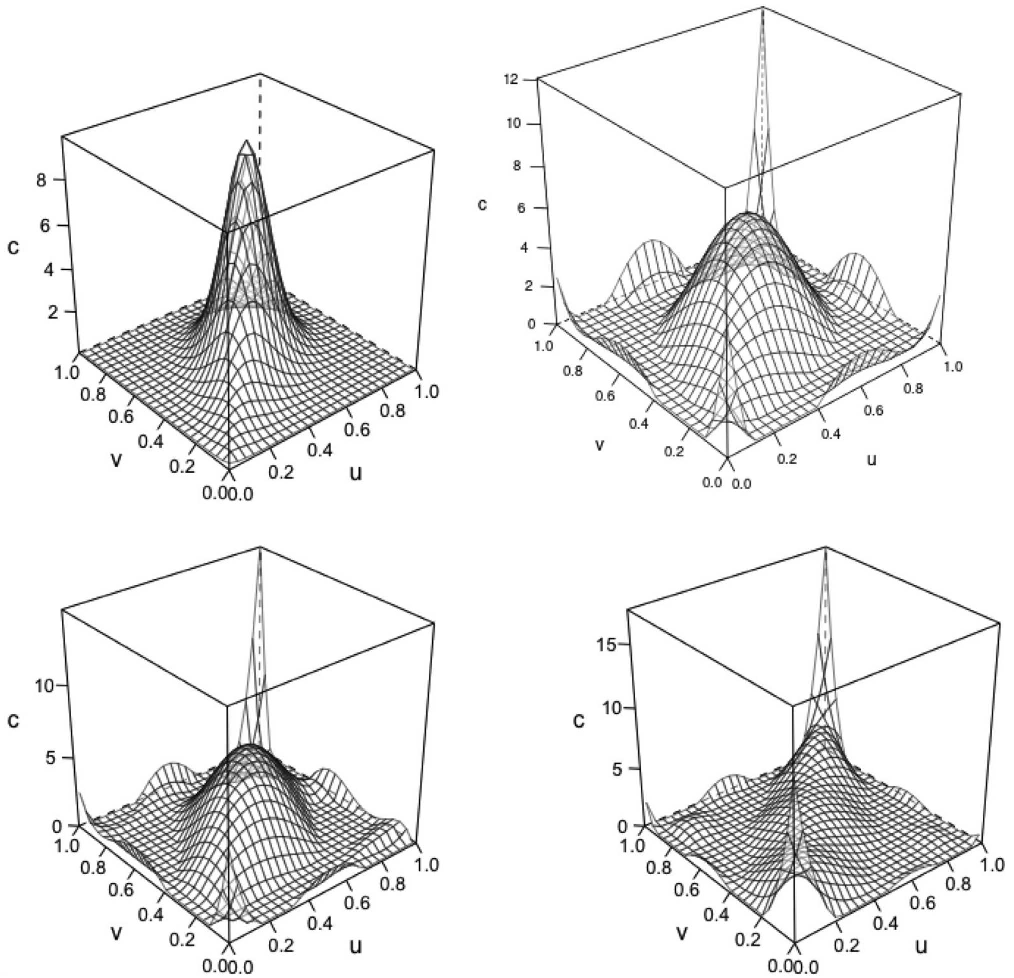

Fig. 2. Three dimensional gumbel density $c(u, v, w)$ with parameter $\theta = 4$ on a $26 \times 26 \times 1$ grid where $w = 0.5$ and its potential approximations with sample size $n = 1000$. Upper: true density function (left) and $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 4, 4)$ (right). Bottom: $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 5, 4)$ (left) and $\mathbf{N} = (4, 5, 1)$(right).

4.6. A 6-dimensional copula

Finally, we consider five families of copulas of dimension 6: Joe, Student, Gumbel, Clayton and Frank copulas. We fix a level of dependence $\tau = 0.3$ and we consider Monte-Carlo simulations with $n = 100$ observations.

Table 9 presents the results obtained by four copula estimators in terms of relative MISE, MIAE and MKSE. Our estimator performs best for Joe, Gumbel and Franck copulas, while the empirical beta estimator performs best for Student and Clayton copulas.

4.7. Additional comments

From our simulation study we observe that

- For a fixed level of dependence, $\widehat{N}_{opt}$ increases with n. This can be explained by the fact that we can then better estimate the higher order moments and they can therefore bring precision in the estimation.
- For a fixed sample size n, $\widehat{N}_{opt}$ increases with the level of dependence. This can be explained by the fact that the structure of the copulas is more complex and that higher order moments provide more information. For example the first copula coefficients are simple correlations and are not sufficient to summarize a high level dependence.
- It appears that $\widehat{N}_{opt}$ is larger for both Joe and Clayton copulas (for a fixed level of dependence). It seems that more copula coefficients are needed to approximate such copulas (or copula densities). This is probably due to the higher asymmetry of these two copulas.
- Tables 10 and 11 show the computation time for both estimators in the bivariate case. As expected, the copula density estimator $\widehat{c}^{[N]}$ is faster than the copula estimator $\widehat{C}^{[N]}$. When the order of approximation N is fixed, the calculation time of

Table 8
Bivariate extreme-value copulas with sample size $n = 100$. Values in bold indicate the minimum of relative MIAE, MISE or MKSE and underlined values indicate the minimum of standard deviations (in brackets).

Copulas	Methods	MIAE $\times 10^2$	MISE $\times 10^2$	MKSE $\times 10^2$
Marshall-Olkin	Emp	4.52(<u>1.14</u>)	0.22(<u>0.12</u>)	5.54(1.25)
	Beta	3.91(1.28)	0.17(<u>0.12</u>)	4.04(1.72)
	Check	4.42(1.18)	0.22(0.13)	5.36(1.25)
	Berns10	3.68(1.52)	0.17(0.14)	4.08(1.26)
	Berns25	**3.51**(1.35)	**0.15**(<u>0.12</u>)	3.65(1.20)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 3$	3.76(1.33)	0.16(<u>0.12</u>)	**3.63**(<u>1.14</u>)
Tawn	Emp	3.53(1.19)	0.15(0.09)	5.24(<u>0.88</u>)
	Beta	2.83(1.11)	0.11(0.09)	3.52(1.01)
	Check	3.32(1.18)	0.14(0.09)	5.05(0.89)
	Berns10	3.63(1.02)	0.15(0.13)	3.07(1.63)
	Berns25	2.63(1.09)	0.09(0.10)	2.78(1.24)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 2$	**2.61**(<u>0.81</u>)	**0.08**(<u>0.07</u>)	**2.32**(1.09)
Galambos	Emp	5.66(1.30)	0.24(0.14)	4.64(<u>1.22</u>)
	Beta	3.96(1.21)	0.18(0.14)	3.93(1.37)
	Check	5.48(1.30)	0.23(0.15)	4.51(1.26)
	Berns10	2.37(1.02)	**0.11**(<u>0.12</u>)	3.14(1.61)
	Berns25	2.99(1.11)	0.13(0.13)	3.38(1.44)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 1$	**2.36**(<u>0.83</u>)	**0.11**(<u>0.12</u>)	**3.03**(1.65)
Hüsler-Reiss	Emp	1.40(<u>0.15</u>)	0.026(0.009)	3.43(0.74)
	Beta	**0.66**(0.25)	0.015(0.011)	2.29(0.70)
	Check	0.68(<u>0.15</u>)	0.017(0.008)	3.22(0.74)
	Berns10	4.86(0.25)	0.353(0.033)	5.95(0.32)
	Berns25	2.06(0.27)	0.081(0.020)	3.43(<u>0.47</u>)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 6$	1.14(0.19)	**0.014**(<u>0.007</u>)	**1.42**(0.48)

Table 9
Relative MIAE, MISE and MKSE based on $n = 100$ observations of six-dimensional copulas. Values in bold indicate the minimum of relative MIAE, MISE or MKSE and underlined values indicate the minimum of standard deviations (in brackets).

Copulas	Methods	MIAE $\times 10^2$	MISE $\times 10^2$	MKSE $\times 10^2$
Joe	Emp	26.19(9.32)	6.49(4.85)	23.43(10.60)
	Beta	23.80(<u>8.31</u>)	5.23(<u>3.64</u>)	20.20(<u>8.09</u>)
	Check	26.17(8.83)	6.42(5.01)	23.13(10.31)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 2$	**19.55**(9.95)	**4.31**(4.95)	**17.83**(9.80)
Student	Emp	20.70(8.54)	4.75(4.47)	19.59(12.38)
	Beta	**17.17**(<u>6.56</u>)	**3.28**(<u>2.99</u>)	**17.04**(<u>10.69</u>)
	Check	21.28(9.14)	4.87(4.61)	19.91(12.05)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 3$	20.54(14.10)	8.02(4.89)	25.19(19.64)
Gumbel	Emp	25.17(10.19)	6.45(5.44)	22.48(12.76)
	Beta	21.29(8.40)	4.81(4.15)	19.60(11.49)
	Check	25.58(10.11)	6.52(5.29)	22.69(12.76)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 2$	**16.31**(<u>8.13</u>)	**3.28**(<u>3.47</u>)	**15.86**(<u>10.99</u>)
Clayton	Emp	22.70(20.93)	9.01(14.44)	25.34(15.29)
	Beta	**20.22**(22.56)	**7.98**(**15.12**)	**18.16**(17.50)
	Check	23.25(19.86)	8.69(13.51)	24.51(15.55)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 3$	25.69(<u>19.44</u>)	8.81(<u>9.53</u>)	20.54(<u>15.07</u>)
Frank	Emp	26.75 (<u>9.56</u>)	8.23 (8.59)	26.80 (21.82)
	Beta	22.14 (9.92)	6.41 (<u>7.52</u>)	24.91 (<u>21.00</u>)
	Check	28.45 (11.89)	9.14 (10.21)	27.95 (22.72)
	$\widehat{C}^{[\widehat{N}_{opt}]}$ & $\widehat{N}_{opt} = 3$	**20.40** (11.38)	**6.11** (10.22)	**22.03** (23.31)

Table 10
Computation time required to estimate the copula density and copula on a 5×5 point grid when N is fixed.

Estimators		computing time in seconds			
		$n = 100$	$n = 200$	$n = 500$	$n = 1000$
$\widehat{c}^{[N]}$	$N = 5$	0.341	0.343	0.344	0.368
$\widehat{C}^{[N]}$	$N = 5$	0.368	0.372	0.377	0.382

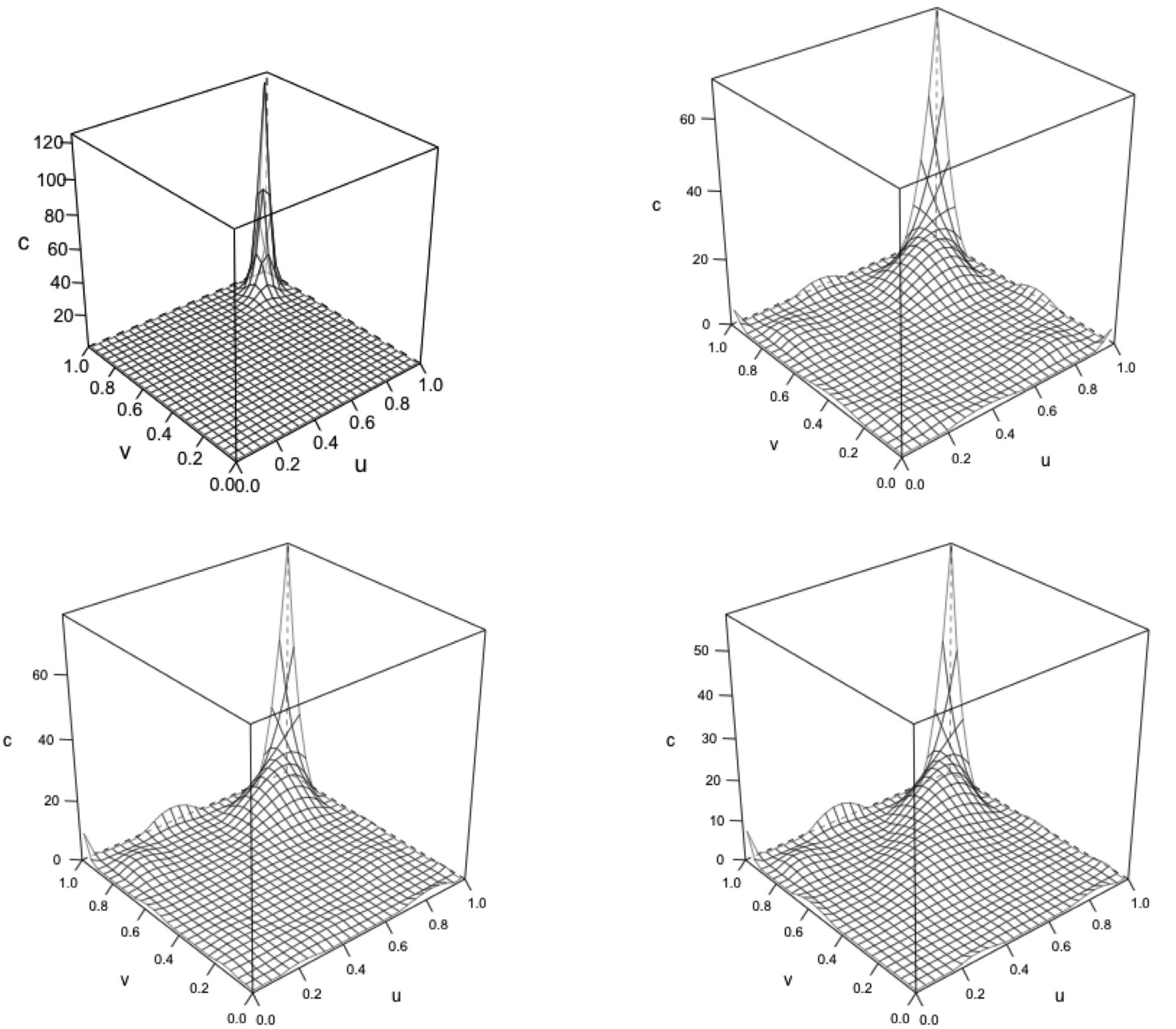

Fig. 3. Three dimensional gumbel density $c(u, v, w)$ with parameter $\theta = 4$ on a $26 \times 26 \times 1$ grid where $w = 0.9$ and its potential approximations with sample size $n = 1000$. Upper: true density function (left) and $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 4, 4)$ (right). Bottom: $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = (4, 5, 4)$ (left) and $\mathbf{N} = (4, 5, 1)$(right).

Table 11
Computation time required to estimate the copula density and copula on a 5×5 point grid when n is fixed.

Estimators		computing time in seconds			
		$N = 1$	$N = 5$	$N = 10$	$N = 15$
$\widehat{c}^{[N]}$	$n = 500$	0.07	0.35	0.99	1.98
$\widehat{C}^{[N]}$	$n = 500$	0.07	0.37	1.05	2.08

both estimators increases very slowly with the sample size n (Table 10). When n is fixed the computation time increases linearly with N (Tables 11).

5. Real data applications

5.1. Insurance data

We consider a classical dataset in the copula literature, namely the Loss-ALAE dataset, that was collected by the US Insurance Services Office. It contains 1,500 general liability claims each consiting of the indemnity payment (Loss) and the allocated loss adjustment expense (ALAE). We excluded 34 censored observations of the dataset. Many authors have used

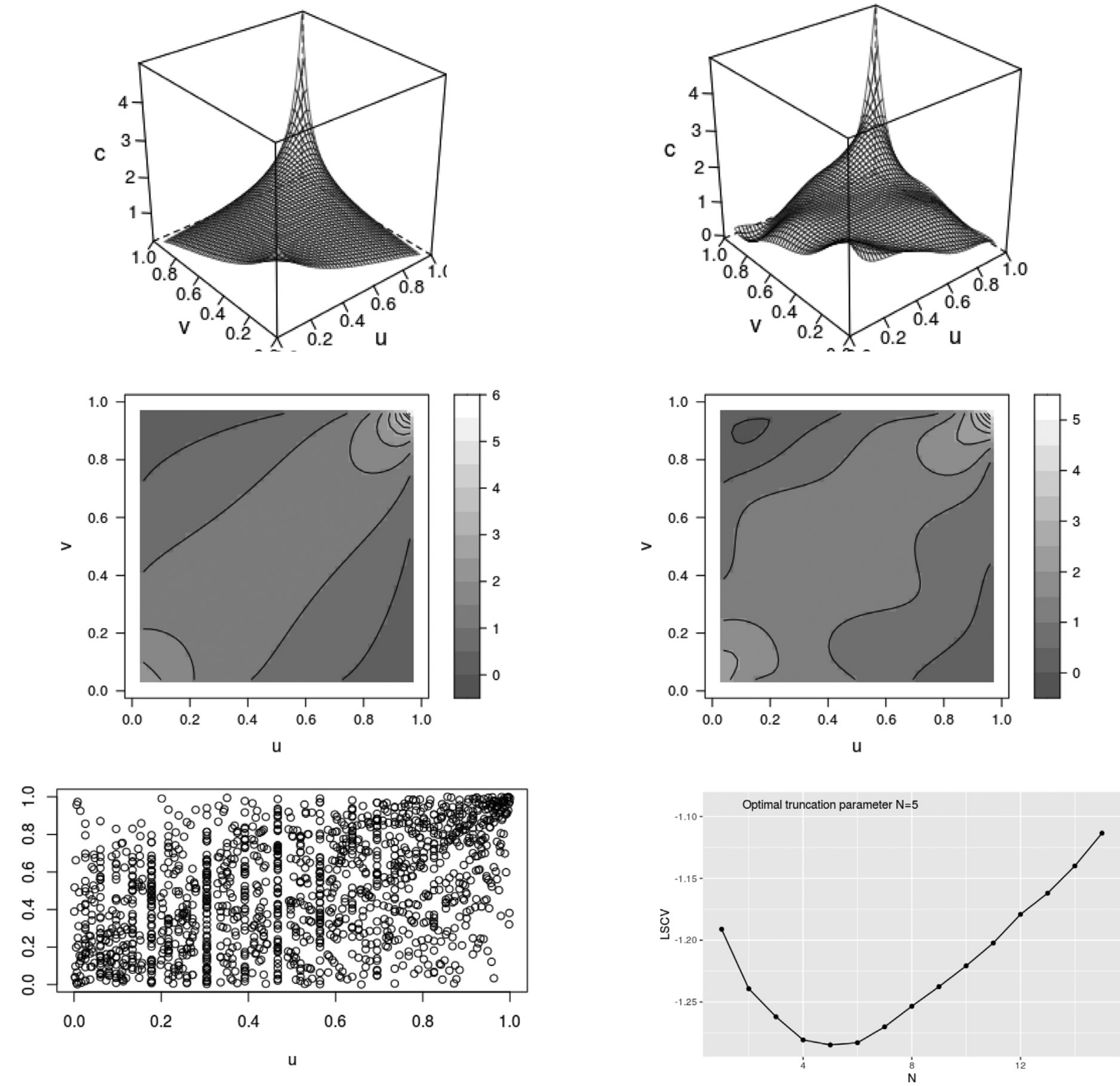

Fig. 4. Loss-ALAE data. Upper: Gumbel copula density with parameter 1.45 (left); $\widehat{c}^{[\mathbf{N}]}$ with $\mathbf{N} = 5$ (right). Middle: Gumbell contours lines (left); $\widehat{c}^{[\mathbf{5}]}$ contours lines (right). Bottom: panel the rank-rank plot (left); the LSCV function (right).

copulas to model the dependence between the variables Loss and ALEA, including Frees and Valdez (1998), Klugman and Parsa (1999), Chen and Fan (2005), Genest et al. (2006), Denuit et al. (2006) and Chen et al. (2010). The general conclusion is that the Gumbel copula provides an adequate fit for these data. Our purpose here is not to take back the analyses, but to confront our nonparametric estimator to the adjusted Gumbel copula. Figure 4 (right in the last row) shows the graph of the function LSCV using the selection rule prescribed in Section 4.1. The selected truncation parameter is $N = 5$. This figure also shows the proximity between our estimator $\widehat{C}^{[5]}$ and the Gumbell copula. To confirm the proximity of the copula as well as its density, we propose to compare them point by point on a grid $\mathcal{G} = \{(u, v)/u, v \in \{1/36, 2/36, \ldots, 35/36\}\}$ with 35×35 points uniformly chosen on $[0, 1]^2$. The corresponding 1225 values associated to the copula density estimator and to the theoretical Gumbel copula density are represented in Figure 5. The similarities between the two densities are clearly demonstrated, even at the boundaries of $[0, 1]^2$. The same analysis of the copula is displayed in Figure 6. It appears that the Gumbel copula and the estimator copula merge very well. Note that the grid numbers on the x-axis of these figures represent the row number sequence in the data frame $\mathcal{G}$. So each number corresponds to a couple (u_j, v_j).

Note that Geenens et al. (2017) also studied the Loss-ALEA dataset and simlarly concluded that the probit-transformation estimator was very close to the Gumbel density as shown in Figure 4 of their paper.

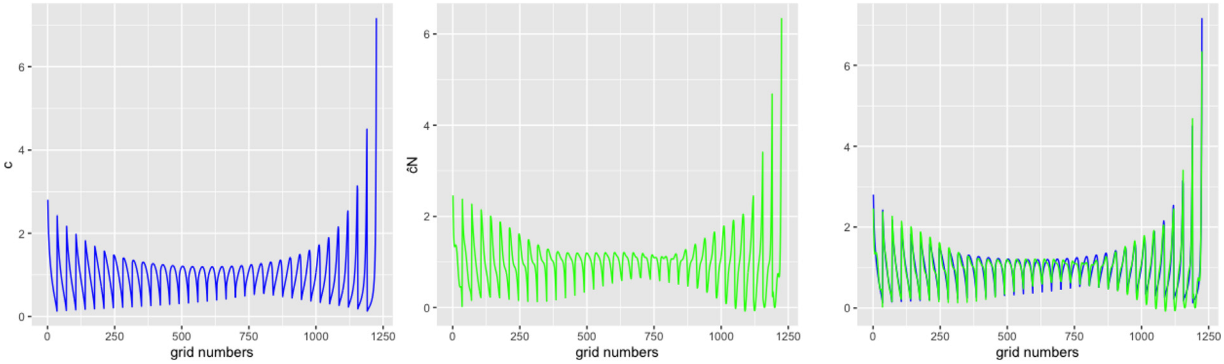

Fig. 5. Loss-ALAE. 2-D line plot: Gumbel copula density with parameter 1.45 (left); $\widehat{c}^{[5]}$ (middle); and their superposition (right).

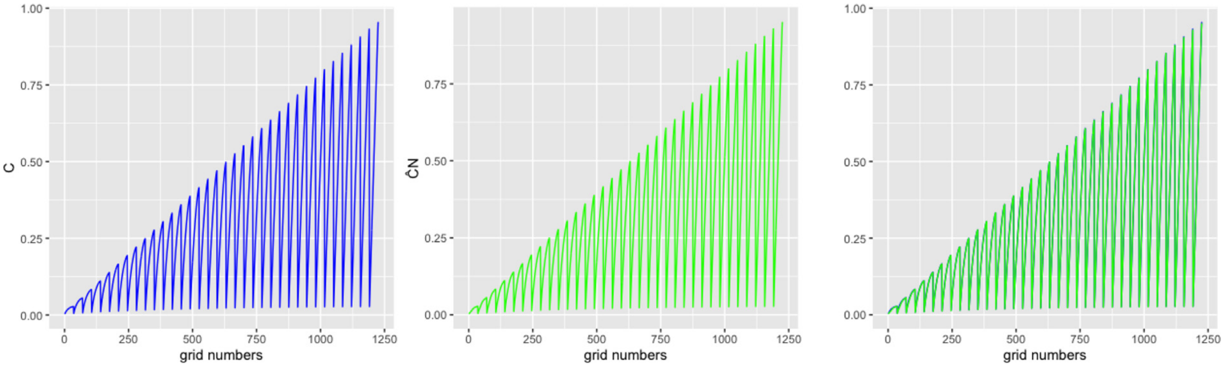

Fig. 6. Loss-ALAE. 2-D line plot: Gumbel copula with parameter 1.45 (left); $\widehat{C}^{[5]}$ (middle); and their superposition (right).

5.2. Financial data

In this data analysis, we study the dependence structure between the series of the most used exchange rates in foreign exchange markets. Data are available from IMF (International Monetary Fund) rates database and consists in daily currency exchange rates from January 1^{st}, 1994 to July 31^{th}, 2020 for a total of 5,551 business days (any day except weekends and some holidays). We exclude from our analysis 1,049 observations with at least one missing value. By the standard continuously compounded return formula, we consider the log-returns of six exchange rates: the Euro (EUR), the Great British pound (GBP), the Japanese yen (YEN), the Canadian dollar (CAD), the Swiss franc (CHF) and the Australian dollar (AUD). The time plots of these returns are shown in Figure 9 (see Supplementary material). These time plots of log returns show the stylized fact of clustering volatility. We also notice the appearance of extreme values. Summary statistics for these returns are displayed in Table 12 (see Supplementary Material). It reveals that the mean and median of log-returns of these six exchange rates are very close to zero. Their distribution are positively skewed (right tailed) and leptokurtic (kurtosis value higher than the kurtosis of normal distribution whose value is three). The CHF has the highlest kurtosis and skewness, the EUR the lowest kurtosis and the CAD the lowest skewness. As it was expected due to more frequent occurrence of extreme values in Figure 9, none of daily log returns passed the Jarque-Bera (JB) test ($P_{value} < 0.001$) where the null hypothesis is that the log-return series are normally distributed. As a result, using standard models is not appropriate in this context because one of the most important financial time series models assumption is the normality of the log-return series.

Tables 13 and 14 (see Supplementary Material) give respectively the optimal degree parameters and Kendall's rank correlations between the six daily log-returns. It is shown that the proposed estimators do not need a high optimal truncation order. The largest degree parameter is $N = 16$ (EUR/CHF) and the smallest one is $N = 3$ (EUR/YEN). Futher, various Kendall correlations are small values, positive for some pairs of daily log returns and negative for others. The pair EUR/CHF has the highest Kendall correlation in absolute value ($\tau \approx 0.6309$).

As it was expected from the simulation studies in Section 4, we notice a relation between the Kendall coefficients and the optimal degree parameters of the estimators. Both indicators increase or decrease simultaneously. This information can be useful for the development of risk diversification strategies since the investment in a portfolio of various assets reduces risk, particularly in exchange rate management. The study could be extended to the calculations of the tail value at risk and risk measure expected shortfall. The different pairs of daily log returns are presented in Figure 7 and 8. We see that the proposed approach captures very well the asymmetry in the dependence structure between these pairs.

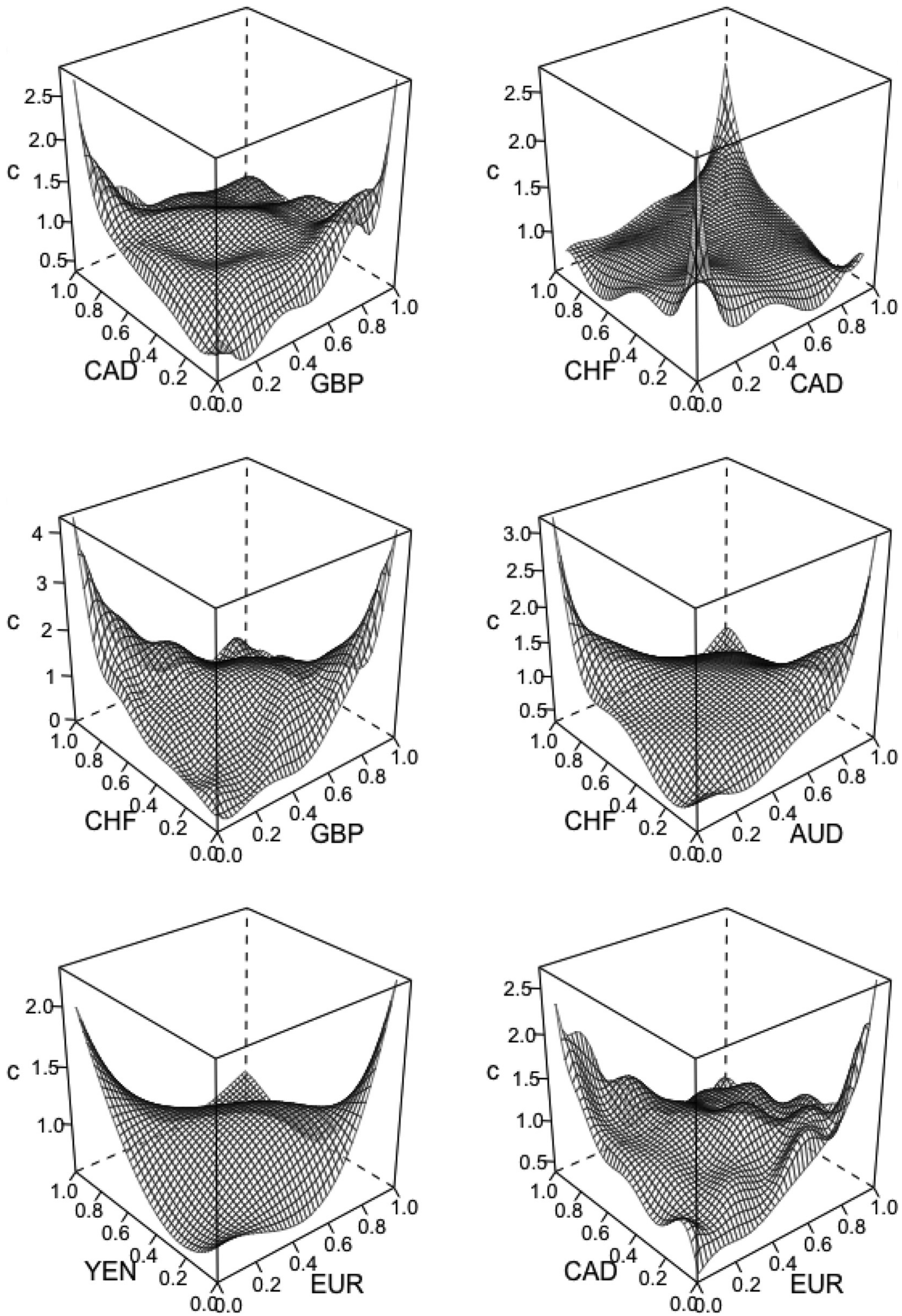

Fig. 7. Estimators of the copula density for Canadian Dollar/Great British pound, Swiss franc/Canadian Dollar, Swiss franc/Great British pound, Swiss franc/Australian Dollar, Japanese yen/Euro and Canadian Dollar/Euro.

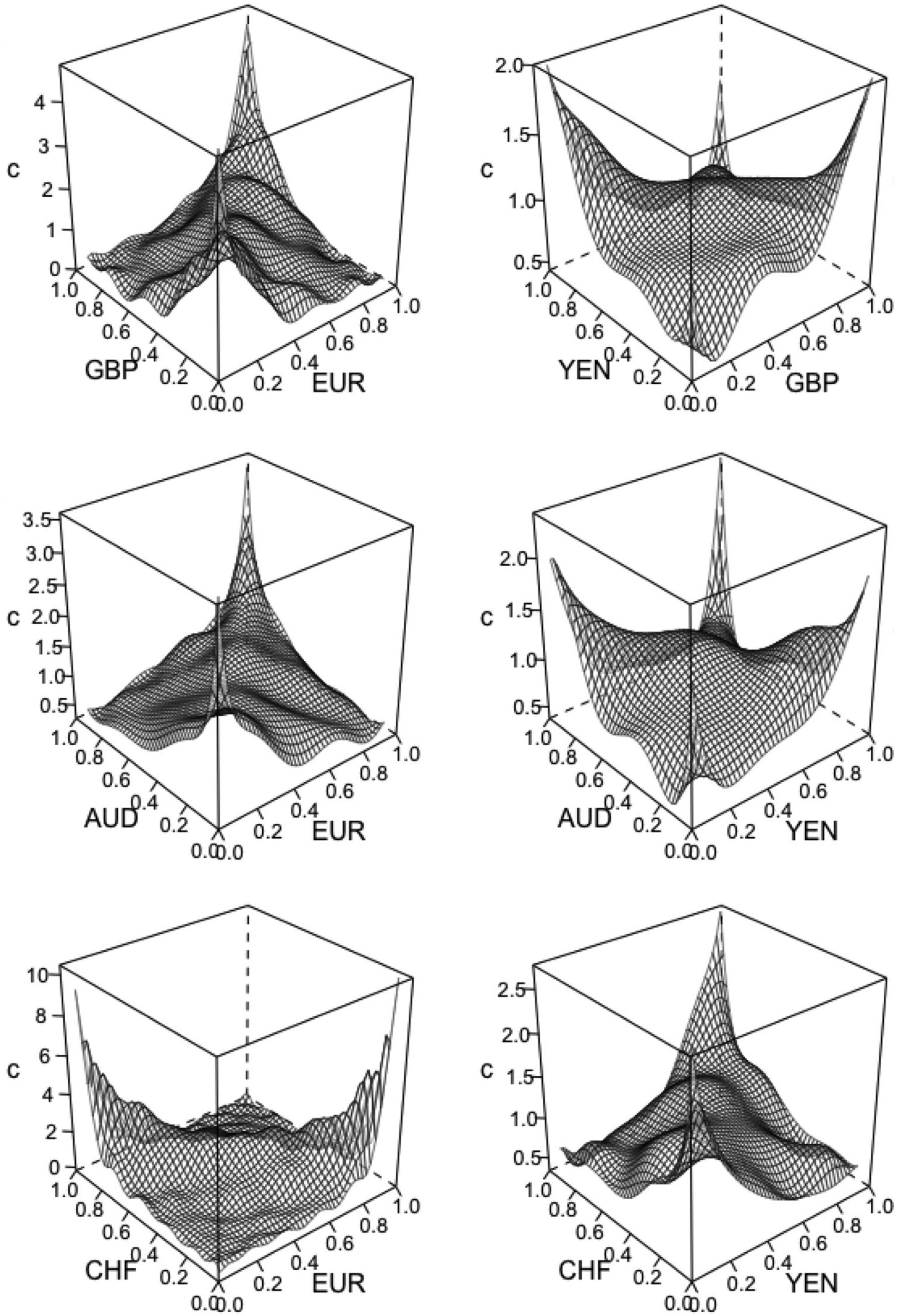

Fig. 8. Estimators of the copula density for Euro/Great British pound, Japanese yen/Great British pound, Australian Dollar/Euro, Australian Dollar/Japanese yen, Swiss franc/Euro and Swiss franc/Japanese yen.

6. Conclusion

In this paper, a nonparametric estimator of the copula density based on shifted Legendre polynomials proposed by Gui (2009) is adapted to the multivariate case. A new copula estimator is then deduced based on copula coefficients under the square integrability assumption (3). Both estimators are easy to implement with an automatic selection of their degrees of approximation. A R code is available on Github-yvesngounou. Various theoretical properties are demonstrated. Experimental results clearly demonstrated the good performance of the proposed estimators. The proposed copula estimators seem to outperform all its recent competitors found in the literature according to the relative MIAE, MISE and MKSE criteria in various scenarios. These experiment results also showed the superiority of the proposed copula density estimator with respect to recent estimators in the literature, excepted for the local log-quadratic estimator of Geenens et al. (2017) which seems to be often better but which is only available in the bivariate case.

Real situations in actuarial and financial frameworks have demonstrated the adaptability of the proposed method. The performances could be improved when considering different orders for $N_1, N_2, \ldots, N_d$ instead of equal values $N_1 = \cdots = N_d$. This seems possible at a reasonable computationally cost when d is not too large.

In the independent case, the estimator is extremely accurate because the correct copula is generally selected. This result may lead to thinking about the construction of a test of independence. Moreover, the correspondence that we have shown between the copula coefficients and the Spearman's rho could be an interesting direction for future research. Finally, this approach by orthogonal projections should also allow us to construct a test statistic to compare copulas.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the many helpful suggestions of two anonymous referees and the editors. The second author acknowledges support by Research Chair DIALog under the aegis of the Risk Foundation, an initiative by CNP Assurances.

Proofs

We will denote

$$\|\widehat{F}_j - F_j\|_\infty := \sup|\widehat{F}_j - F_j|, \quad \pi(\boldsymbol{m}) := \prod_{j=1}^{d} m_j, \quad N_{max} := \max\{N_j;\, j := 1, \ldots, d\}.$$

Let begin with the following useful lemmas

Lemma 2.

$$\|\widehat{F}_j - F_j\|_\infty = o_{\mathbb{P}}(1) \text{ and } \|\widehat{F}_j - F_j\|_\infty = \mathcal{O}_{\mathbb{P}}(\sqrt{n^{-1}}).$$

Proof. According to the Massart inequality (Massart, 1990), we get

$$\forall \epsilon > 0, \quad \mathbb{P}\Big(\|\widehat{F}_j - F_j\|_\infty > \epsilon\Big) \le 2e^{-2n\epsilon^2}, \tag{10}$$

and the results follow. □

Lemma 3.

$$\mathbb{E}\|\widehat{F}_j - F_j\|_\infty^2 \le n^{-1} \quad \text{and} \quad \mathbb{E}\|\widehat{F}_j - F_j\|_\infty \le \sqrt{n^{-1}}$$

Proof. Denoting $Y := \|\widehat{F}_j - F_j\|_\infty$ and using (10), we see that

$$\mathbb{E}(Y) \le \sqrt{\mathbb{E}(Y^2)} = \sqrt{\int_0^{+\infty} \mathbb{P}(Y > \sqrt{s})ds} \le \sqrt{\int_0^{+\infty} 2e^{-2ns}ds} = \sqrt{n^{-1}},$$

and the results follow. □

Lemma 4. *For all* $\alpha \ge 1$,

$$\sum_{\mathbf{m} \le \mathbf{N}} \big(\pi(\boldsymbol{m})\big)^\alpha \le N_{max}^{d(\alpha+1)}.$$

Proof. We have

$$\begin{aligned}\sum_{\mathbf{m} \le \mathbf{N}} \big(\pi(\boldsymbol{m})\big)^\alpha &= \prod_{j=1}^{d} \sum_{m_j=0}^{N_j} m_j^{\alpha-1} m_j \;\le\; N_{max}^{d(\alpha-1)} \prod_{j=1}^{d} \sum_{m_j=0}^{N_j} m_j \\ &\le N_{max}^{d(\alpha-1)} \prod_{j=1}^{d} \frac{N_j(N_j+1)}{2} \;\le\; N_{max}^{d(\alpha-1)} \prod_{j=1}^{d} N_j^2 \le N_{max}^{d(\alpha+1)}.\end{aligned}$$

□

Lemma 5. *For all $u \in [0,1]$ we have*

$$|Q_m(u)| \le \eta_1 m^{1/2} \text{ and } |Q'_m(u)| \le \eta_2 m^{5/2}, \qquad \forall m > 0, \tag{11}$$

where $\eta_1 = \sqrt{3}$ and $\eta_2 = 2\sqrt{3}$.

Proof. See for instance inequalities 22.14 on page 791 of Abramowitz and Stegun (1970). □

Proof of Proposition 1

Proof. From (1) the density of $\mathbf{U}$ is $h_{\mathbf{U}} = c$. Since $(\prod_{j=1}^d Q_{m_j})_{\mathbf{m}\in\mathbb{N}^d}$ forms a dense orthogonal basis with respect to the uniform measure on $[0,1]^d$, by (3) we have

$$h_{\mathbf{U}}(u_1,\dots,u_d) = \sum_{\mathbf{m}\in\mathbb{N}^d}\left(\int c(v_1,\dots,v_d)\prod_{j=1}^d Q_{m_j}(v_j)dv_1\cdots dv_d\right)\prod_{j=1}^d Q_{m_j}(u_j),$$

which gives (4), and (5) follows by integration. □

Proof of Proposition 2

Proof. Assume that $u_i = 0$. Then we have

$$\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) = \sum_{\mathbf{m}\le\mathbf{N}}\widehat{\rho}_{\mathbf{m}}\left(\int_0^0 Q_{m_i}(x_i)\mu(dx_i)\right)\prod_{\substack{j=1\\ j\neq i}}^d\int_0^{u_j} Q_{m_j}(x_j)\mu(dx_j) = 0,$$

and *i*) is proved. To prove *ii*) assume that $\mathbf{u} = (1,\dots,1,u_i,1,\dots,1)$. Then

$$\begin{aligned}
\widehat{C}^{[\mathbf{N}]}(\mathbf{u}) &= \sum_{\mathbf{m}\le\mathbf{N}}\widehat{\rho}_{\mathbf{m}}\left(\int_0^{u_i} Q_{m_i}(x_i)\mu(dx_i)\right)\prod_{\substack{j=1\\ j\neq i}}^d\int_0^1 Q_{m_j}(x_j)\mu(dx_j)\\
&= \sum_{\mathbf{m}\le\mathbf{N}}\widehat{\rho}_{\mathbf{m}}\left(\int_0^{u_i} Q_{m_i}(x_i)\mu(dx_i)\right)\prod_{\substack{j=1\\ j\neq i}}^d\delta_{0,m_j}\\
&= \widehat{\rho}_{\mathbf{0}}\int_0^{u_i} Q_0(x_i)\mu(dx_i) = \int_0^{u_i} 1\ \mu(dx_i) = u_i.
\end{aligned}$$

iii) is immediate from $\widehat{\rho}_{\mathbf{0}} = 1$ and $Q_0 = 1$. □

Proof of Proposition 3

Proof. By orthogonality of the polynomials we have

$$\begin{aligned}
\int_{[0,1]^d}\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\mu(d\mathbf{u}) &= \sum_{\mathbf{m}\le\mathbf{N}}\widehat{\rho}_{\mathbf{m}}\prod_{j=1}^d\int_0^1 Q_{m_j}(u_j)\mu(du_j)\\
&= \sum_{\mathbf{m}\le\mathbf{N}}\widehat{\rho}_{\mathbf{m}}\prod_{j=1}^d\delta_{0,m_j}\\
&= 1,
\end{aligned}$$

which yields *i*). The proof of *ii*) is very similar since $\widehat{\rho}_{\mathbf{m}} = 0$ if exactly $d-1$ components of $\mathbf{m}$ are null. *iii*) is immediate. □

Proof of Proposition 4

Proof. We have

$$\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) - \tilde{c}^{[\mathbf{N}]}(\mathbf{u}) = \sum_{\mathbf{m}\le\mathbf{N}}(\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}}\prod_{j=1}^d Q_{m_j}(u_j)$$

$$= \sum_{\mathbf{m}\le\mathbf{N}} \frac{1}{n}\sum_{i=1}^{n}\left(\prod_{j=1}^{d} Q_{m_j}(\widehat{F}_j(X_{ij})) - \prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij}))\right)\prod_{j=1}^{d} Q_{m_j}(u_j).$$

Combining (11) with Taylor expansion we obtain

$$\begin{aligned} |\prod_{j=1}^{d} Q_{m_j}(\widehat{F}_j(X_{ij})) - \prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij}))| &\le \sum_{j=1}^{d} \eta_2 m_j^{5/2} \prod_{i\neq j} \eta_1 m_i^{1/2} \|\widehat{F}_j - F_j\|_\infty \\ &\le M(\mathbf{m}) \sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty, \end{aligned} \tag{12}$$

where $M(\mathbf{m}) = \eta_2\eta_1^{d-1}\pi(\mathbf{m})^{1/2}\max(m_j)^2$. It follows that

$$\begin{aligned} |\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) - \tilde{c}^{[\mathbf{N}]}(\mathbf{u})| &\le \sum_{\mathbf{m}\le\mathbf{N}} M(\mathbf{m}) \max|\prod_{j=1}^{d} Q_{m_j}| \sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty \\ &\le \eta_2\eta_1^{2d-1} N_{max}^2 \sum_{\mathbf{m}\le\mathbf{N}} \pi(\mathbf{m}) \sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty \\ &\le \eta_2\eta_1^{2d-1} N_{max}^{2d+2} \sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty, \end{aligned} \tag{13}$$

from Lemma 4, which yields *i)* since $\|\widehat{F}_j - F_j\|_\infty = \mathcal{O}_{\mathbb{P}}\left(\sqrt{n^{-1}}\right)$ according to Lemma 2.

We obtain *ii)* analogously. □

Proof of Proposition 5.

Proof. By orthogonality of the Legendre polynomials we have

$$\|\widehat{c}^{[\mathbf{N}]} - c\|_2^2 = \sum_{\mathbf{m}\le\mathbf{N}} (\widehat{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2 + \sum_{\mathbf{m}\not\le\mathbf{N}} \rho_{\mathbf{m}}^2,$$

which gives result. □

Proof of Corollary 1

Proof. By orthogonality, we have

$$\begin{aligned} MISE(\widehat{c}^{[\mathbf{N}]}) - MISE(\tilde{c}^{[\mathbf{N}]}) &= \mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}} (\widehat{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2\right) - \mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}} (\tilde{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2\right) \\ &= \mathbb{E}\sum_{\mathbf{m}\le\mathbf{N}} (\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}})((\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}}) + 2(\tilde{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})). \end{aligned} \tag{14}$$

From (12), we first have

$$|\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}}| \le M(\mathbf{m}) \sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty,$$

and since

$$\|\widehat{F}_j - F_j\|_\infty = \mathcal{O}_{\mathbb{P}}\left(\sqrt{n^{-1}}\right),$$

it follows that

$$\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}} = \mathcal{O}_{\mathbb{P}}(\sqrt{n^{-1}}).$$

Furthermore, by the Central Limit Theorem we have

$$\sqrt{n}(\tilde{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}}) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbb{V}\big(\prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij})\big)\right),$$

where $\mathbb{V}\left(\prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij}))\right) < \infty$ since $\boldsymbol{N}$ is fixed and independent of n. Then

$$\tilde{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}} = \mathcal{O}_{\mathbb{P}}\left(\sqrt{n^{-1}}\right). \tag{15}$$

Combining (14), (12) and (15) we get

$$MISE(\widehat{c}^{[\mathbf{N}]}) = MISE(\tilde{c}^{[\mathbf{N}]}) + \mathcal{O}\left(n^{-1}\right). \qquad \square$$

Proof of Corollary 2

Proof. We want to prove that $MISE(\widehat{c}^{[\mathbf{N}]})$ tends to zero as n tends to infinity. From Proposition 5, we have

$$MISE(\widehat{c}^{[\mathbf{N}]}) := \mathbb{E}\left(\|\widehat{c}^{[\mathbf{N}]} - c\|_2^2\right) = \mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}} (\widehat{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2\right) + \sum_{\mathbf{m}\not\le\mathbf{N}} \rho_{\mathbf{m}}^2.$$

From the elementary inequality $2ab \le a^2 + b^2$ for any $a, b \in \mathbb{R}$, we get

$$\begin{aligned}\mathbb{E}\left(\|\widehat{c}^{[\mathbf{N}]} - c\|_2^2\right) &\le 2\mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}} (\widehat{\rho}_{\mathbf{m}} - \tilde{\rho}_{\mathbf{m}})^2\right) + 2\mathbb{E}\left(\sum_{\mathbf{m}\le\mathbf{N}} (\tilde{\rho}_{\mathbf{m}} - \rho_{\mathbf{m}})^2\right) + \sum_{\mathbf{m}\not\le\mathbf{N}} \rho_{\mathbf{m}}^2 \\ &:= 2A_1 + 2A_2 + A_3.\end{aligned}$$

Clearly, $A_3 \to 0$, as $N \to \infty$ as a tail of a convergent series. Moreover by (11) we have

$$\begin{aligned}A_2 &= \sum_{\mathbf{m}\le\mathbf{N}} \frac{1}{n} \mathbb{V}\left(\prod_{j=1}^{d} Q_{m_j}(F_j(X_j))\right) \\ &\le \sum_{\mathbf{m}\le\mathbf{N}} \frac{1}{n} \eta_1^{2d} \prod_{j=1}^{d} m_j \\ &\le \frac{\eta_1^{2d} N_{max}^{2d}}{n},\end{aligned}$$

and by **(H)** $A_2 \to 0$ as $n \to \infty$. To conclude, we combine (12) with (11) to obtain

$$\begin{aligned}A_1 &= \sum_{\mathbf{m}\le\mathbf{N}} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \mathbb{E}\Bigg[\left(\prod_{j=1}^{d} Q_{m_j}(\widehat{F}_j(X_{ij})) - \prod_{j=1}^{d} Q_{m_j}(F_j(X_{ij}))\right) \\ &\qquad \times \left(\prod_{j=1}^{d} Q_{m_j}(\widehat{F}_j(X_{i'j})) - \prod_{j=1}^{d} Q_{m_j}(F_j(X_{i'j}))\right)\Bigg] \\ &\le \sum_{\mathbf{m}\le\mathbf{N}} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \mathbb{E}\left(M(\mathbf{m})^2 \left(\sum_{j=1}^{d} \|\widehat{F}_j - F_j\|_\infty\right)^2\right) \\ &\le d \sum_{\mathbf{m}\le\mathbf{N}} M(\mathbf{m})^2 \sum_{j=1}^{d} \mathbb{E}\left(\|\widehat{F}_j - F_j\|_\infty^2\right),\end{aligned}$$

according to Hölder's sum inequality.

Since $M(\mathbf{m}) = \eta_1^{d-1} \eta_2 \pi(\mathbf{m})^{1/2} \max(m_j)^2$ and $\mathbb{E}\left(\|\widehat{F}_j - F_j\|_\infty^2\right) \le n^{-1}$, we see that

$$\begin{aligned}A_1 &\le \frac{d^2 \eta_1^{2d-2} \eta_2^2}{n} \sum_{\mathbf{m}\le\mathbf{N}} \pi(\boldsymbol{m}) \max(m_j)^4 \\ &\le \frac{d^2 \eta_1^{2d-2} \eta_2^2 N_{max}^4}{n} \sum_{\mathbf{m}\le\mathbf{N}} \pi(\boldsymbol{m}) \\ &\le \frac{d^2 \eta_1^{2d-2} \eta_2^2 N_{max}^{2d+4}}{n},\end{aligned}$$

and then $A_1 \to 0$ as $n \to \infty$, which concludes the proof. $\square$

Proof of Proposition 6

Proof. We have

$$\mathbb{E}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right) = \mathbb{E}\left(\tilde{c}^{[\mathbf{N}]}(\mathbf{u})\right) + \mathbb{E}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) - \mathbb{E}\tilde{c}^{[\mathbf{N}]}(\mathbf{u})\right).$$

Since $\mathbb{E}(\tilde{\rho}_{\mathbf{m}}) = \rho_{\mathbf{m}}$, we have

$$\mathbb{E}\left(\tilde{c}^{[\mathbf{N}]}(\mathbf{u})\right) = \sum_{\mathbf{m}\leq\mathbf{N}} \mathbb{E}(\tilde{\rho}_{\mathbf{m}}) \prod_{j=1}^{d} Q_{m_j}(u_j) = \sum_{\mathbf{m}\leq\mathbf{N}} \rho_{\mathbf{m}} \prod_{j=1}^{d} Q_{m_j}(u_j),$$

which tends to $c(\mathbf{u})$ as n tends to ∞. From (13), according to Lemma 3, we have

$$\begin{aligned}|\mathbb{E}\widehat{c}^{[\mathbf{N}]}(\mathbf{u}) - \mathbb{E}\tilde{c}^{[\mathbf{N}]}(\mathbf{u})| &\leq \eta_2\eta_1^{2d-1}N_{max}^{2d+2}\sum_{j=1}^{d}\mathbb{E}\|\widehat{F}_j - F_j\|_\infty \\ &\leq d\eta_2\eta_1^{2d-1}\frac{N_{max}^{2d+2}}{\sqrt{n}}.\end{aligned}$$

Under assumption $(\boldsymbol{H})$, we deduce that

$$\mathbb{E}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right) \longrightarrow c(\mathbf{u}) \text{ as } n \to \infty,$$

and analogously

$$\mathbb{E}\left(\widehat{C}^{[\mathbf{N}]}(\mathbf{u})\right) \longrightarrow C(\mathbf{u}) \text{ as } n \to \infty.$$

□

Proof of Theorem 1

Proof. The result is a direct consequence of Theorem 5 (under Assumption P) or Theorem 7 (under Assumption D) of Radulović et al. (2017). □

Proof of Proposition 7

Proof. Writing $\widehat{c}^{[\mathbf{N}]}_{(-i)}$ and $\widehat{\rho}_{\mathbf{m}}^{(-i)}$ the leave-one-out estimators obtained by deleting the ith observation we have

$$\begin{aligned}\widehat{LSCV}(\mathbf{N}) &= \int_{[0,1]^d}\left(\widehat{c}^{[\mathbf{N}]}(\mathbf{u})\right)^2 d\mathbf{u} - \frac{2}{n}\sum_{i=1}^{m}\widehat{c}^{[\mathbf{N}]}_{(-i)}\left(\widehat{F}_1(X_{i1}),\ldots,\widehat{F}_d(X_{id})\right) \\
&= \sum_{\mathbf{m}\leq\mathbf{N}}\widehat{\rho}_{\mathbf{m}}^2 - \frac{2}{n}\sum_{i=1}^{n}\sum_{\mathbf{m}\leq\mathbf{N}}\widehat{\rho}_{\mathbf{m}}^{(-i)}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{ij})\right) \\
&= \sum_{\mathbf{m}\leq\mathbf{N}}\left(\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{ij})\right)\right)^2 \\
&\quad -\frac{2}{n}\sum_{i=1}^{n}\left(\sum_{\mathbf{m}\leq\mathbf{N}}\left(\frac{1}{n-1}\sum_{\substack{k=1,\\k\neq i}}^{n}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{kj})\right)\right)\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{ij})\right)\right) \\
&= \frac{1}{n^2}\sum_{\mathbf{m}\leq N}\left(\sum_{i=1}^{n}\prod_{j=1}^{d}Q_{m_j}^2\left(\widehat{F}_j(X_{ij})\right) + \sum_{k\neq i}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{kj})\right)Q_{m_j}\left(\widehat{F}_j(X_{ij})\right)\right) \\
&\quad -\frac{2}{n(n-1)}\sum_{\mathbf{m}\leq\mathbf{N}}\sum_{k\neq i}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{kj})\right)Q_{m_j}\left(\widehat{F}_j(X_{ij})\right) \\
&= \frac{1}{n^2}\sum_{\mathbf{m}\leq N}\left(\sum_{i=1}^{n}\prod_{j=1}^{d}Q_{m_j}^2\left(\widehat{F}_j(X_{ij})\right) - \frac{n+1}{n-1}\sum_{k\neq i}\prod_{j=1}^{d}Q_{m_j}\left(\widehat{F}_j(X_{ij})\right)Q_{m_j}\left(\widehat{F}_j(X_{kj})\right)\right),\end{aligned}$$

which is the desired conclusion. □

Proof of Proposition 8

Proof. With the notation of the proof of Proposition 7, if the margins are known, we have

$$\mathbb{E}(LSCV(N)) = \mathbb{E}\left(\int_{[0,1]^d} \left(\tilde{c}^{[\mathbf{N}]}(\mathbf{u})\right)^2 \mu(d\mathbf{u})\right) - 2\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n \tilde{c}^{[\mathbf{N}]}_{(-i)}(F_1(X_{i1}),\dots,F_d(X_{id}))\right).$$

In addition we have

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n \tilde{c}^{[\mathbf{N}]}_{(-i)}(F_1(X_{i1}),\dots,F_d(X_{id}))\right) = \frac{1}{n}\sum_{i=1}^n \sum_{m\le N} \mathbb{E}(\tilde{\rho}_m)\mathbb{E}\left(\prod_{j=1}^d Q_{m_j}(F_j(X_{ij}))\right)$$

$$= \sum_{m\le N} \mathbb{E}(\tilde{\rho}_m)\int_{\mathbb{R}^d} \prod_{j=1}^d Q_{m_j}(F_j(x_{ij}))h(x_1,\dots,x_d)\mu(d\mathbf{x})$$

$$= \sum_{n\le N} \mathbb{E}(\tilde{\rho}_n)\int_{[0,1]^d} \prod_{j=1}^d Q_{m_j}(u_j)c(u_1,\dots,u_d))\mu(d\mathbf{u})$$

$$= \mathbb{E}\int_{[0,1]^d} \sum_{m\le N} \tilde{\rho}_m \prod_{j=1}^d Q_{m_j}(u_j)c(u_1,\dots,u_d)\mu(d\mathbf{u})$$

$$= \mathbb{E}\int_{[0,1]^d} \tilde{c}^{[\mathbf{N}]}(\mathbf{u})c(\mathbf{u})\mu(d\mathbf{u}),$$

which implies that

$$\begin{aligned}\mathbb{E}(LSCV(N)) &= \mathbb{E}\left(\int_{[0,1]^d} \left(\tilde{c}^{[\mathbf{N}]}(\mathbf{u})\right)^2 \mu(d\mathbf{u}) - 2\int_{[0,1]^d} \tilde{c}^{[\mathbf{N}]}(\mathbf{u})c(\mathbf{u})\mu(d\mathbf{u})\right)\\ &= \mathbb{E}\left(\int_{[0,1]^d} \left(\tilde{c}^{[\mathbf{N}]}(\mathbf{u}) - c(\mathbf{u})\right)^2 \mu(d\mathbf{u})\right) - \int_{[0,1]^d} (c(\mathbf{u}))^2\mu(d\mathbf{u}),\end{aligned}$$

and we obtain

$$\mathbb{E}(LSCV(N)) = \mathbb{E}\|\tilde{c}^{[\mathbf{N}]} - c\|_2^2 - \|c\|_2^2. \tag{16}$$

If the margins are unknown, we combine Corollary 1 and equation (16) to obtain the result. □

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecosta.2023.04.002

References

Abramowitz, M., Stegun, I.A., 1970. Handbook of mathematical functions with formulas, graphs, and mathematical tables, Vol. 55. US Government printing office.
Autin, F., Le Pennec, E., Tribouley, K., 2010. Thresholding methods to estimate copula density. Journal of Multivariate Analysis 101 (1), 200–222.
Beare, B.K., 2010. Copulas and temporal dependence. Econometrica 78 (1), 395–410.
Bouezmarni, T., El Gouch, A., Taamouti, A., 2013. Bernstein estimator for unbounded copula densities. Statistics & risk modeling. 30 (4), 343–360.
Bouezmarni, T., Rombouts, J.V., Taamouti, A., 2010. Asymptotic properties of the bernstein density copula estimator for α-mixing data. Journal of Multivariate Analysis 101 (1), 1–10.
Bowman, A.W., 1984. An alternative method of cross-validation for the smoothing of density estimates. Biometrika 71 (2), 353–360.
Carley, H., Taylor, M., 2002. A new proof of sklar's theorem. In: Distributions with given marginals and statistical modelling. Springer, pp. 29–34.
Charpentier, A., Fermanian, J.-D., Scaillet, O., 2007. The estimation of copulas: Theory and practice. Copulas: From theory to application in finance 35–64.
Chatelain, S., Fougères, A.-L., Neslehová, J.G., et al., 2020. Inference for archimax copulas. Annals of Statistics 48 (2), 1025–1051.
Chatrabgoun, O., Parham, G., Chinipardaz, R., 2017. A legendre multiwavelets approach to copula density estimation. Statistical Papers 58 (3), 673–690.
Chen, X., Fan, Y., 2005. Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. Canadian Journal of Statistics 33 (3), 389–414.
Chen, X., Fan, Y., Pouzo, D., Ying, Z., 2010. Estimation and model selection of semiparametric multivariate survival functions under general censorship. Journal of Econometrics 157 (1), 129–142.
Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. Bulletins de l'Académie Royale de Belgique 65 (1), 274–292.
Denuit, M., Purcaru, O., Keilegom, I.V., 2006. Bivariate archimedean copula models for censored data in non-life insurance. Journal of Actuarial Practice 1993-2006. 5..
Fermanian, J.-D., Radulovic, D., Wegkamp, M., 2004. Weak convergence of empirical copula processes. Bernoulli 10 (5), 847–860. doi:10.3150/bj/1099579158.
Frees, E.W., Valdez, E.A., 1998. Understanding relationships using copulas. North American actuarial journal 2 (1), 1–25.
Geenens, G., Charpentier, A., Paindaveine, D., et al., 2017. Probit transformation for nonparametric kernel estimation of the copula density. Bernoulli 23 (3), 1848–1873.

Genest, C., Masiello, E., Tribouley, K., 2009. Estimating copula densities through wavelets. Insurance: Mathematics and Economics 44 (2), 170–181.
Genest, C., Quessy, J.-F., Rémillard, B., 2006. Goodness-of-fit procedures for copula models based on the probability integral transformation. Scandinavian Journal of Statistics 33 (2), 337–366.
Gijbels, I., Mielniczuk, J., 1990. Estimating the density of a copula function. Communications in Statistics-Theory and Methods 19 (2), 445–464.
Gui, W., 2009. Adaptive Series Estimators for Copula Densities. PhD thesis.
Hamdan, M., Al-Bayyati, H., 1971. Canonical expansion of the compound correlated bivariate poisson distribution. Journal of the American Statistical Association 66 (334), 390–393.
Janssen, P., Swanepoel, J., Veraverbeke, N., 2014. A note on the asymptotic behavior of the bernstein estimator of the copula density. Journal of Multivariate Analysis 124, 480–487.
Joe, H., 2014. Dependence modeling with copulas. CRC press.
Kauermann, G., Schellhase, C., Ruppert, D., 2013. Flexible copula density estimation with penalized hierarchical b-splines. Scandinavian Journal of Statistics 40 (4), 685–705.
Klugman, S.A., Parsa, R., 1999. Fitting bivariate loss distributions with copulas. Insurance: mathematics and economics 24 (1-2), 139–148.
Ko, V., Hjort, N.L., 2019. Model robust inference with two-stage maximum likelihood estimation for copulas. Journal of Multivariate Analysis 171, 362–381.
Krupskii, P., Joe, H., 2020. Flexible copula models with dynamic dependence and application to financial data. Econometrics and Statistics 16, 148–167. doi:10.1016/j.ecosta.2020.01.005.
Lancaster, H.O., 1958. The Structure of Bivariate Distributions. The Annals of Mathematical Statistics 29 (3), 719–736. doi:10.1214/aoms/1177706532.
Massart, P., 1990. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. The annals of Probability 1269–1283.
McNeil, A.J., Frey, R., Embrechts, P., 2015. Quantitative risk management: concepts, techniques and tools-revised edition. Princeton university press.
Nagler, T., 2018. kdecopula: An R package for the kernel estimation of bivariate copula densities. Journal of Statistical Software 84 (7), 1–22. doi:10.18637/jss.v084.i07.
Nelsen, R.B., 2007. An introduction to copulas. Springer Science & Business Media.
Omelka, M., Gijbels, I., Veraverbeke, N., 2009. Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. The Annals of Statistics 37 (5B), 3023–3058.
Pérez, A., Prieto-Alaiz, M., 2016. A note on nonparametric estimation of copula-based multivariate extensions of spearman's rho. Statistics & Probability Letters 112, 41–50.
Qu, L., Yin, W., 2012. Copula density estimation by total variation penalized likelihood with linear equality constraints. Computational Statistics & Data Analysis 56 (2), 384–398.
Radulović, D., Wegkamp, M., Zhao, Y., 2017. Weak convergence of empirical copula processes indexed by functions. Bernoulli 23 (4B), 3346–3384. doi:10.3150/16-BEJ849.
Rudemo, M., 1982. Empirical choice of histograms and kernel density estimators. Scandinavian Journal of Statistics 65–78.
Sancetta, A., Satchell, S., 2004. The bernstein copula and its applications to modeling and approximations of multivariate distributions. Econometric theory 535–562.
Segers, J., Sibuya, M., Tsukahara, H., 2017. The empirical beta copula. Journal of Multivariate Analysis 155, 35–51.
Sklar, A., 1959. Fonction de répartition dont les marges sont données. Inst. Stat. Univ. Paris 8, 229–231.
Taylor, C.C., 1990. Orthogonal series estimators and cross-validation. Journal of Statistical Computation and Simulation 37 (3-4), 151–158.