



Generalized simulated method-of-moments estimators for multivariate copulas

Mohamed Belalia¹ · Jean-François Quessy² 

Received: 15 December 2023 / Revised: 1 May 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

This paper introduces a general semi-parametric method for estimating a vector of parameters in multivariate copula models. The proposed approach uses the moments of the multivariate probability integral random variable to generalize the inversion of Kendall's tau estimator. What makes the new methodology attractive is the fact that it can be performed as soon as one can simulate from the assumed parametric family of copulas. This feature is especially helpful when explicit expressions are not available for the theoretical moments. The consistency and asymptotic normality of the proposed estimators are established under mild conditions. An extensive simulation study indicates that the price to pay for the estimation of the moments is modest and that the new estimators are almost as accurate as the pseudo-maximum likelihood (PML) estimator. The usefulness of the proposed estimators is illustrated on the modelling of multivariate data with copula models where the PML estimator is hardly computable.

Keywords Pairwise models · Probability integral transformation · Pseudo-maximum likelihood estimator · Semi-parametric inference · Simulated moments · U-statistics.

1 Introduction

When the goal is to find an appropriate model for a random vector $\mathbf{X} \in \mathbb{R}^d$, a well-established strategy is to model the marginal behaviours and the dependence structure separately. This approach is possible thanks to a theorem of Sklar (1959), which states

✉ Jean-François Quessy
Jean-Francois.Quessy@uqtr.ca

Mohamed Belalia
Mohamed.Belalia@uwindsor.ca

¹ Department of Mathematics and Statistics, University of Windsor, 401 Sunset avenue, Windsor, Canada

² Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, P.B. 500, Trois-Rivières G9A 5H7, Canada

that there exists $C : [0, 1]^d \rightarrow [0, 1]$ such that for all $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$,

$$\mathbb{P}(\mathbf{X} \leq \mathbf{x}) = C \{\mathbb{P}(X_1 \leq x_1), \dots, \mathbb{P}(X_d \leq x_d)\}.$$

Whenever the marginal distributions are continuous, the function C is unique and is called the copula of \mathbf{X} . See Nelsen (2007) and Joe (2015) for details on copulas. In that context, it is customary to assume that C belongs to a parametric family $\mathcal{C} = \{C_\theta; \theta \in \Theta \subseteq \mathbb{R}^p\}$ and then estimate the unknown parameter from a sample of copies $\mathbf{X}_1, \dots, \mathbf{X}_T$ of \mathbf{X} .

A popular estimator of a copula parameter is the pseudo-maximum likelihood (PML) estimator introduced by Oakes (1994) and later investigated by Genest et al. (1995). In principle, this method is applicable regardless of the dimension of the vector and of the number of parameters. However, the PML estimator requires an explicit expression for the copula density, which is not always the case, notwithstanding the fact that the density may be explicit but intractable. The PML estimator can also be numerically unstable, especially when the family has several parameters. Some authors have proposed approaches to address these shortcomings. For example, minimum-distance estimators were considered by Biau and Wegkamp (2005) based on copula densities and by Tsukahara (2005) and Weiß (2011) relying on goodness-of-fit metrics. When there is only one parameter to estimate, a common strategy is the inversion of Kendall's tau. This estimator was considered by Genest et al. (2006) in the bivariate case and by Genest et al. (2011) for d -dimensional copulas.

This paper extends the use of the inversion of Kendall's tau to families comprising multiple parameters. The proposed estimators are based on unbiased estimations of the moments of the multivariate probability integral transformation (MPIT) random variable, thereby avoiding the need to estimate the marginal distributions. The proposed method is similar to Brahim and Necir (2012), who suggested using higher moments of the MPIT. However, their estimators are biased and their approach is limited to cases where the vector of theoretical moments is explicitly invertible. In order to circumvent the aforementioned constraint, an approach based on simulated moments à la McFadden (1989) is adopted. Hence, the proposed estimators can be performed as soon as it is possible to simulate from a given copula family.

The manuscript is organized as follows. A generic method-of-moments estimator based on a vector of U-statistics is introduced in Sect. 2; its consistency and asymptotic normality are established under mild conditions for both the standard and simulated versions. Section 3 explains how to estimate the moments of the multivariate probability integral transformation without biases and describes the new estimators of copula parameters suitably adapted to the parametric structure at hand. In Sect. 4, the performance of the proposed estimators is investigated and compared to competing procedures through an extensive simulations study. Section 5 illustrates the introduced methodologies on the modelling of multivariate data with copula models that have a complex parametric structure. Section 6 ends the paper with a brief discussion. The mathematical proofs can be found in an appendix and the Matlab code is available at www.uqtr.ca/MyMatlabWebpage.

2 A generic method-of-moments estimator of copula parameters

2.1 Statistical functionals

The estimators proposed in this paper are special cases of a generic method-of-moments estimator based on a vector of U-statistics. Specifically, let $\mathbf{X} = (X_1, \dots, X_d)$ be a random vector from a d -variate distribution function F that has continuous marginals F_1, \dots, F_d and a unique copula C . Consider the vector $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_L)$, where for each $\ell \in \{1, \dots, L\}$, $\kappa_\ell := \kappa_\ell(\mathbf{x}_1, \dots, \mathbf{x}_m)$ is a symmetric function in its m arguments. Then, for $\mathbf{X}_1, \dots, \mathbf{X}_m$ i.i.d. F , define the statistical functional

$$S_{\boldsymbol{\kappa}}(F) = \mathbb{E} \{ \boldsymbol{\kappa}(\mathbf{X}_1, \dots, \mathbf{X}_m) \}.$$

In order to develop semi-parametric estimators of the parameters of a given family of copulas, it is necessary that $S_{\boldsymbol{\kappa}}(F)$ be free of the marginal distributions. To this end, $\boldsymbol{\kappa}$ must be such that $S_{\boldsymbol{\kappa}}(F) = S_{\boldsymbol{\kappa}}(C)$, i.e.,

$$\mathbb{E} \{ \boldsymbol{\kappa}(\mathbf{X}_1, \dots, \mathbf{X}_m) \} = \mathbb{E} \{ \boldsymbol{\kappa}(\mathbf{U}_1, \dots, \mathbf{U}_m) \}, \tag{1}$$

where $\mathbf{U}_1, \dots, \mathbf{U}_m$, with $\mathbf{U}_j = (F_1(X_{j1}), \dots, F_d(X_{jd}))$, are i.i.d. C .

Assuming a random sample $\mathbf{X}_1, \dots, \mathbf{X}_T$ i.i.d. F , an unbiased estimator of the vector of means $\boldsymbol{\mu}_C^{\boldsymbol{\kappa}} = \mathbb{E} \{ \boldsymbol{\kappa}(\mathbf{U}_1, \dots, \mathbf{U}_m) \}$ is given by the L -dimensional vector of U -statistics

$$\boldsymbol{\mu}_T^{\boldsymbol{\kappa}} = \binom{T}{m}^{-1} \sum_{1 \leq t_1 < \dots < t_m \leq T} \boldsymbol{\kappa}(\mathbf{X}_{t_1}, \dots, \mathbf{X}_{t_m}), \tag{2}$$

where in the above expression, the sum of vectors is taken component by component. From Theorem 3, p. 122 of Lee (1990), $\boldsymbol{\mu}_T^{\boldsymbol{\kappa}}$ converges almost surely to $\boldsymbol{\mu}_C^{\boldsymbol{\kappa}}$ as long as $\mathbb{E}\{|\boldsymbol{\kappa}(\mathbf{X}_1, \dots, \mathbf{X}_m)|\} < \infty$.

The next result states the asymptotic normality of $\sqrt{T}(\boldsymbol{\mu}_T^{\boldsymbol{\kappa}} - \boldsymbol{\mu}_C^{\boldsymbol{\kappa}})$. The proof is a direct application of Theorem 2, p. 76 of Lee (1990) and is therefore omitted. Here and in the sequel, \rightsquigarrow means ‘‘converges in distribution to’’.

Lemma 1 For each $\ell \in \{1, \dots, L\}$, let $\mathbf{U}_1, \dots, \mathbf{U}_m$ be i.i.d. C and define $\boldsymbol{\kappa}_C^* = (\boldsymbol{\kappa}_1^*, \dots, \boldsymbol{\kappa}_L^*)$, where

$$\boldsymbol{\kappa}_\ell^*(\mathbf{u}) = m \left[\mathbb{E} \{ \kappa_\ell(\mathbf{u}, \mathbf{U}_2, \dots, \mathbf{U}_m) \} - \boldsymbol{\mu}_{C,\ell}^{\boldsymbol{\kappa}} \right].$$

Provided that $\Sigma_C^{\boldsymbol{\kappa}} = \mathbb{E}\{\boldsymbol{\kappa}_C^*(\mathbf{U}_1)^\top \boldsymbol{\kappa}_C^*(\mathbf{U}_1)\} \in \mathbb{R}^{L \times L}$ is non-singular, $\sqrt{T}(\boldsymbol{\mu}_T^{\boldsymbol{\kappa}} - \boldsymbol{\mu}_C^{\boldsymbol{\kappa}}) \rightsquigarrow \mathcal{N}(\mathbf{0}_L, \Sigma_C^{\boldsymbol{\kappa}})$.

2.2 A generalized method-of-moments estimator

Suppose that the copula C of the population belongs to a copula family

$$C = \{C_{\theta}; \theta \in \Theta \subset \mathbb{R}^L\}.$$

In that case, there exists $\theta_0 \in \Theta$ called the “true value”, so that $C = C_{\theta_0}$. As is well known, a method-of-moments estimator consists of estimating θ_0 by selecting $\theta \in \Theta$ such that the sample moments match their theoretical counterparts. In the current context, let $\mu^{\kappa}(\theta) = \mu_{C_{\theta}}^{\kappa}$ and assume that the map $\mu^{\kappa} : \Theta \rightarrow \mathbb{R}^L$ is one-to-one on an open set $\Theta \subset \mathbb{R}^L$. A method-of-moments estimator of θ_0 is the unique solution of $\mu^{\kappa}(\theta) = \mu_T^{\kappa}$, namely

$$\theta_T^{\kappa} = (\mu^{\kappa})^{-1}(\mu_T^{\kappa}).$$

The consistency and asymptotic normality of θ_T^{κ} are stated next.

Proposition 1 *Let $\mathbf{X}_1, \dots, \mathbf{X}_T$ be i.i.d. C_{θ_0} and assume that $\mu^{\kappa} : \Theta \rightarrow \mathbb{R}^L$ is one-to-one and continuously differentiable at $\theta_0 \in \Theta$ with nonsingular first order derivative $\mathbf{v}_0^{\kappa} \in \mathbb{R}^{L \times L}$ at θ_0 . Let also $A_0 = (\mathbf{v}_0^{\kappa})^{-1}$ and define Σ_0^{κ} as the covariance matrix in Lemma 1 when $C = C_{\theta_0}$. Then θ_T^{κ} exists with a probability tending to one and*

$$\sqrt{T}(\theta_T^{\kappa} - \theta_0) \rightsquigarrow \mathcal{N}\left(\mathbf{0}_L, A_0^{\kappa} \Sigma_0^{\kappa} (A_0^{\kappa})^{\top}\right).$$

2.3 Simulated version of the generalized method-of-moments estimator

It is difficult to compute the proposed method-of-moments estimator when μ^{κ} is not explicitly invertible. This is even worse when there is no explicit expression for μ^{κ} . In such cases, it is useful to express θ_T^{κ} as the minimum-distance estimator

$$\theta_T^{\kappa} = \arg \min_{\theta \in \Theta} \{\mu_T^{\kappa} - \mu^{\kappa}(\theta)\} M_T \{\mu_T^{\kappa} - \mu^{\kappa}(\theta)\}^{\top},$$

where $M_T \in \mathbb{R}^{L \times L}$ is a weight matrix that converges in probability to a positive definite matrix $M_0 \in \mathbb{R}^{L \times L}$ as $T \rightarrow \infty$. Nevertheless, this expression does not solve the problem of cases where $\mu^{\kappa}(\theta)$ admits no explicit expression.

To avoid the above-mentioned drawbacks, a simulated version of the generic method-of-moments estimator θ_T^{κ} is proposed. The idea is in the same spirit as that investigated by Oh and Patton (2013) for copula-based time series models, which is itself inspired by the simulated method-of-moments estimators studied by McFadden (1989), Pakes and Pollard (1989) and Newey and McFadden (1994). To describe the idea in the current context, consider a version of $\mu^{\kappa}(\theta)$ based on a simulated sample $\mathbf{U}_1^{\theta}, \dots, \mathbf{U}_S^{\theta}$ i.i.d. C_{θ} , namely

$$\mu_S^k(\theta) = \binom{S}{m}^{-1} \sum_{1 \leq s_1 < \dots < s_m \leq S} \kappa(\mathbf{U}_{s_1}^\theta, \dots, \mathbf{U}_{s_m}^\theta).$$

The proposed simulated method-of-moments estimator of $\theta_0 \in \Theta$ is then

$$\theta_{T,S}^k = \arg \min_{\theta \in \Theta} \{\mu_T^k - \mu_S^k(\theta)\} M_T \{\mu_T^k - \mu_S^k(\theta)\}^\top. \tag{3}$$

Proposition 2 states the consistency of $\theta_{T,S}^k$ as $T, S \rightarrow \infty$, i.e., $\theta_{T,S}^k$ converges in probability to the true value $\theta_0 \in \Theta$. The conditions under which it happens are mild. In particular, it is no longer required that μ^k is one-to-one.

Proposition 2 *Let $\mathbf{X}_1, \dots, \mathbf{X}_T$ be i.i.d. from a distribution with continuous marginals and unique copula that belongs to a family $\{C_\theta; \theta \in \Theta \subset \mathbb{R}^L\}$, where Θ is compact. For $\theta_0 \in \Theta$ being the true value, assume that $\mu^k(\theta_0) \neq \mu^k(\theta)$ as soon as $\theta \neq \theta_0$ and that μ^k is Lipschitz continuous on Θ . Then, as $T, S \rightarrow \infty$, $\theta_{T,S}^k$ is consistent for θ_0 .*

Remark 1 Unlike standard results on the consistency of simulated method-of-moments estimators (see Pakes and Pollard 1989; McFadden 1989, for instance), Proposition 2 allows S, T to go to infinity at different rates. In other words, it is assumed more generally that $T/S \rightarrow \zeta \in [0, \infty)$ as $T, S \rightarrow \infty$, so that the case when $\mu(\theta)$ is explicit is recovered at the limit when $T/S \rightarrow 0$.

Lemma 2 *A sufficient condition for the Lipschitz continuity of μ^k is that the density c_θ of C_θ be uniformly Lipschitz, i.e., there exists $K \in (0, \infty)$ such that*

$$\sup_{\mathbf{u} \in [0,1]^d} |c_{\theta_1}(\mathbf{u}) - c_{\theta_2}(\mathbf{u})| \leq K \|\theta_1 - \theta_2\|,$$

where $\|\cdot\|$ is the euclidean norm.

The uniform Lipschitz property as stated in Lemma 2 holds for the Farlie–Gumbel–Morgenstern family of bivariate copulas whose members have density $c_\theta(u_1, u_2) = 1 + \theta(1 - 2u_1)(1 - 2u_2)$ for $\theta \in [-1, 1]$, since one readily obtains

$$|c_{\theta_1}(u_1, u_2) - c_{\theta_2}(u_1, u_2)| \leq |\theta_1 - \theta_2|.$$

However, this uniform Lipschitz condition is rather strong and not verified for many commonly used copula families.

It is possible to show the almost sure convergence (strong consistency) of $\theta_{T,S}^k$, but its weak consistency is enough to establish its asymptotic normality. This result is precisely the subject of the next proposition.

Proposition 3 *Let $\mathbf{X}_1, \dots, \mathbf{X}_T$ be i.i.d. from a distribution with continuous marginals and unique copula that belongs to a family $\{C_\theta; \theta \in \Theta \subset \mathbb{R}^L\}$. Assume that the true value θ_0 is an interior point of Θ and that $g_0^k(\theta) = \mu^k(\theta_0) - \mu^k(\theta)$ possesses a*

derivative G_0^κ at θ_0 such that $B_0^\kappa = G_0^\kappa M_0 (G_0^\kappa)^\top$ is nonsingular. Also, suppose that for $g_{T,S}^\kappa(\theta) = \mu_T^\kappa - \mu_S^\kappa(\theta)$,

$$g_{T,S}^\kappa(\theta_{T,S}^\kappa) M_T g_{T,S}^\kappa(\theta_{T,S}^\kappa)^\top - \inf_{\theta \in \Theta} g_{T,S}^\kappa(\theta) M_T g_{T,S}^\kappa(\theta)^\top \leq o_{\mathbb{P}}(T^{-1}).$$

Then for $\Omega_0^\kappa = (B_0^\kappa)^{-1} G_0^\kappa M_0 \Sigma_0^\kappa M_0 (G_0^\kappa)^\top (B_0^\kappa)^{-1}$,

$$\sqrt{\frac{TS}{T+S}} (\theta_{T,S}^\kappa - \theta_0) \rightsquigarrow \mathcal{N}(\mathbf{0}_L, \Omega_0^\kappa) \text{ as } T, S \rightarrow \infty.$$

3 Moments of the multivariate probability integral transformation

3.1 Population versions

The estimators that will be developed in the sequel are based on the multivariate probability integral transformation (MPIT). The MPIT was first introduced by Genest and Rivest (1993) as a tool for testing the fit to bivariate copulas. Specifically, the MPIT of a random vector $\mathbf{X} = (X_1, \dots, X_d) \sim F$ is the random variable $W = F(\mathbf{X})$. When the marginals F_1, \dots, F_d of F are continuous, one can invoke Sklar’s Theorem and write

$$W = C \{F_1(X_1), \dots, F_d(X_d)\} = C(\mathbf{U}),$$

where $\mathbf{U} = (F_1(X_1), \dots, F_d(X_d)) \sim C$. In other words, the stochastic behaviour of W depends only on the unique copula C of F .

For certain parametric structures, it is advantageous to work with the pair-by-pair probability integral transformations related to \mathbf{X} . Specifically, let $C_{jj'}$ be the copula of the pair $(X_j, X_{j'})$ and consider $W_{jj'} = C_{jj'}(U_j, U_{j'})$, where $U_j = F_j(X_j)$ and $U_{j'} = F_{j'}(X_{j'})$. The first moment of $W_{jj'}$ is related to Kendall’s measure of association of $(U_j, U_{j'})$ through the relationship

$$\tau_{jj'} = 4 \mathbb{E}(W_{jj'}) - 1.$$

A similar claim can be made about a d -variate extension due to Kendall and Smith (1940) defined as the mean of the pair-by-pair Kendall’s tau. Another extension of Kendall’s tau proposed by Joe (1990) is related to the first moment of W . These two d -variate Kendall measures are given respectively by

$$\tau_d = \binom{d}{2}^{-1} \sum_{j < j'} \{4 \mathbb{E}(W_{jj'}) - 1\} \quad \text{and} \quad \tilde{\tau}_d = \frac{2^d \mathbb{E}(W) - 1}{2^{d-1} - 1}.$$

3.2 Unbiased estimation of the moments of the MPIT

The next proposition provides an unbiased estimator of the moments of W based on a sample $\mathbf{X}_1, \dots, \mathbf{X}_T$ i.i.d. from a d -variate distribution F with continuous marginals F_1, \dots, F_d and unique copula C .

Proposition 4 For $a \in \mathbb{N}$, let $\mu_a = \mathbb{E}(W^a)$, where $W = C(\mathbf{U})$ and $\mathbf{U} \sim C$. Then an unbiased estimator of μ_a is provided by the U -statistic $\widehat{\mu}_a$ with symmetric kernel of order $a + 1$ defined for $\mathbf{x}_1, \dots, \mathbf{x}_{a+1} \in \mathbb{R}^d$ by

$$\mathcal{K}_a(\mathbf{x}_1, \dots, \mathbf{x}_{a+1}) = \frac{1}{a + 1} \sum_{j=1}^{a+1} \mathbb{I} \left\{ \left(\max_{\ell \in \{1, \dots, a+1\}, \ell \neq j} \mathbf{x}_\ell \right) < \mathbf{x}_j \right\}.$$

This kernel satisfies (1), i.e., is marginal-free.

Note that when $a = 1$, $\mathcal{K}_1(\mathbf{x}_1, \mathbf{x}_2) = \{\mathbb{I}(\mathbf{x}_1 < \mathbf{x}_2) + \mathbb{I}(\mathbf{x}_2 < \mathbf{x}_1)\}/2$. This is, up to a constant, the kernel of the empirical Kendall’s measure of association.

One is now in a position to establish the convergence in distribution of the vector of U -statistics $\widehat{\boldsymbol{\mu}} = (\widehat{\mu}_1, \dots, \widehat{\mu}_L)$.

Proposition 5 Let $\mathbf{X}_1, \dots, \mathbf{X}_T$ be i.i.d. from a d -variate distribution with continuous marginals and unique copula C . For $\widehat{\boldsymbol{\mu}} = (\widehat{\mu}_1, \dots, \widehat{\mu}_L)$ and $\boldsymbol{\mu} = (\mu_1, \dots, \mu_L)$, then $\sqrt{T}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})$ converges to the L -variate normal distribution with zero means and variance-covariance matrix $\Sigma \in \mathbb{R}^{L \times L}$ such that for each $a, a' \in \{1, \dots, L\}$, $\Sigma_{aa'} = \mathbb{E}\{\mathcal{K}_a^*(\mathbf{U}) \mathcal{K}_{a'}^*(\mathbf{U})\}$, $\mathbf{U} \sim C$, where

$$\mathcal{K}_a^*(\mathbf{u}) = \{C(\mathbf{u})\}^a + a \mathbb{E} \left\{ \mathbb{I}(\mathbf{U} > \mathbf{u}) (C(\mathbf{U}))^{a-1} \right\} - (a + 1) \mu_a.$$

3.3 Estimators of copula parameters

Let $\mathbf{X}_1, \dots, \mathbf{X}_T$ be a random sample of independent and identically distributed \mathbb{R}^d -valued vectors whose joint distribution F has continuous marginal distributions. It is assumed that the unique copula C of F belongs to $\{C_\theta; \theta \in \Theta \subseteq \mathbb{R}^L\}$ and the goal is to estimate the unknown parameter $\theta_0 \in \Theta \subset \mathbb{R}^L$.

One possibility to estimate θ_0 is to use the first L moments of $W = C_\theta(\mathbf{U})$, i.e., let $\boldsymbol{\mu}(\theta) = (\mathbb{E}(W), \dots, \mathbb{E}(W^L))$, where it is understood that the expectation is taken with respect to C_θ . An empirical version of $\boldsymbol{\mu}(\theta)$ is the vector $\boldsymbol{\mu}_T$ of the first L empirical moments. The corresponding vector of kernels is then $\boldsymbol{\kappa} = (\mathcal{K}_1, \dots, \mathcal{K}_L)$, where \mathcal{K}_a is defined in Proposition 4. The simulated method-of-moments estimator is thus of the form given in (3), i.e.,

$$\widehat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \{\boldsymbol{\mu}_T - \boldsymbol{\mu}_S(\boldsymbol{\theta})\} M_T \{\boldsymbol{\mu}_T - \boldsymbol{\mu}_S(\boldsymbol{\theta})\}^\top, \tag{4}$$

where $\boldsymbol{\mu}_S(\boldsymbol{\theta})$ estimates $\boldsymbol{\mu}(\boldsymbol{\theta})$ based on a sample of size S from C_θ .

Many parametrization schemes have a pair-by-pair structure of the form $\theta = (\Sigma, \boldsymbol{\gamma})$, where $\Sigma \in \mathbb{R}^{d \times d}$ is a correlation matrix whose off-diagonal entry $\Sigma_{jj'}$ appears only in the distribution of $(X_j, X_{j'})$ for each $j \neq j' \in \{1, \dots, d\}$. This pattern occurs, for example, for models derived from the multivariate Normal. If $\boldsymbol{\gamma} \in \mathbb{R}^q$ is a parameter that appears in the distribution of every sub-vector of \mathbf{X} , one can estimate $\Sigma_{jj'}$ and $\boldsymbol{\gamma}$ from the first $q + 1$ moments of $W_{jj'} = C_{\Sigma_{jj'}, \boldsymbol{\gamma}}(U_j, U_{j'})$, yielding $\widehat{\Sigma}_{jj'}$ and $\widehat{\boldsymbol{\gamma}}_{jj'}$. The global parameter can then be estimated with

$$\widehat{\boldsymbol{\gamma}} = \binom{d}{2}^{-1} \sum_{j < j' \in \{1, \dots, d\}} \widehat{\boldsymbol{\gamma}}_{jj'}. \quad (5)$$

Remark 2 It is worth mentioning the work by Brahimi and Necir (2012), who also suggested using the first L moments of W . Their methodology is however limited to cases where the vector of theoretical moments is explicitly invertible and is based on a biased estimation of the moments using the empirical copula.

3.4 On the use of alternative probability integral random variables

Other moments of a copula could be used for the purpose of parameter estimation. For instance, as pointed out by Quessy (2009), Spearman's measure of association and some of its multivariate extensions can be expressed as the expectation of a symmetric kernel of a U-statistic. Indeed, Spearman's rho is an affine transformation of the expectation of $W_{\text{Sp}} = F_1(X_1) \times \dots \times F_d(X_d)$, where $\mathbf{X} = (X_1, \dots, X_d)$ follows a distribution F with continuous marginals F_1, \dots, F_d . One could then consider $\mathbb{E}(W_{\text{Sp}}^a)$, but its estimation with a U-statistic involves a kernel of order $a \times (d + 1)$. To see it, note that

$$\mathbb{E}(W_{\text{Sp}}^a) = \mathbb{E}_{\mathbf{X}} \left[\left\{ \mathbb{E} \mathbb{I}(\mathbf{Y}^{\Pi} \leq \mathbf{X}) \right\}^a \right] = \mathbb{E}_{\mathbf{X}, \mathbf{Y}_1^{\Pi}, \dots, \mathbf{Y}_a^{\Pi}} \left\{ \prod_{j=1}^a \mathbb{I}(\mathbf{Y}_j^{\Pi} \leq \mathbf{X}) \right\},$$

where $\mathbf{Y}_1^{\Pi}, \dots, \mathbf{Y}_a^{\Pi}$ are i.i.d. $F_1 \times \dots \times F_d$ and $\mathbf{X} \sim F$. The expression inside the brackets involves $a \times (d + 1)$ independent random variables. This makes the use of W_{Sp} less attractive than the use of W , especially as the number of parameters increases.

4 Sampling properties of the estimators

4.1 Preliminaries

This section investigates the performance of the estimators defined in Eqs. (4) and (5). Comparisons with competing estimators are also made. In the sequel, the accuracy of an estimator $\widehat{\theta}$ of a given parameter $\theta \in \mathbb{R}$ is measured by its relative bias (RB) and

relative root mean-squared error (RRMSE), namely

$$RB_{\theta}(\hat{\theta}) = \frac{1}{\theta} \mathbb{E}(\hat{\theta} - \theta) \quad \text{and} \quad RRMSE_{\theta}(\hat{\theta}) = \frac{1}{\theta} \sqrt{\mathbb{E}\{(\hat{\theta} - \theta)^2\}}.$$

As explained by Oh and Patton (2013) (see also Gouriéroux et al. 1996), it is crucial that the random number generator seed involved in their computation be fixed across the generation of the simulated datasets of size S . Otherwise, the evaluation of the estimated function will be unstable and the optimization will not converge. The minimization in Eq. (3) is performed using the MATLAB routine FMINSEARCHBND written by John D’Errico. Unlike for example the Newton algorithm, the latter does not require the existence of derivatives. The maximum number of iterations has been set to 40. Many choices are possible for the weight matrix, including the inverse of the efficient weight matrix. However, in line with a recommendation by Oh and Patton (2013), one only considers $M_T = I_L$ throughout in order to simplify the analyses.

4.2 Calibration of the simulated estimator

A popular estimator in the case of a one-parameter bivariate copula families $\mathcal{C} = \{C_{\theta}; \theta \in \Theta \subseteq \mathbb{R}\}$ is the inversion of Kendall’s tau (IKT). Specifically, if $(X_{11}, X_{12}), \dots, (X_{T1}, X_{T2})$ are random pairs from a population with continuous marginals and a copula $C \in \mathcal{C}$, the IKT estimator is defined by

$$\hat{\theta}^{\tau} = \tau_C^{-1}(\tau_T).$$

In that expression, τ_C is the population value of Kendall’s tau, i.e.,

$$\tau_C(\theta) = 4 \int_0^1 \int_0^1 C_{\theta}(u_1, u_2) dC_{\theta}(u_1, u_2) - 1,$$

and τ_T is its empirical counterpart, i.e.,

$$\tau_T = \frac{4}{T(T-1)} \sum_{i < j} \mathbb{I}\{(X_{i1} - X_{j1})(X_{i2} - X_{j2}) > 0\} - 1.$$

Whereas the simulated method-of-moments estimator is designed for cases where an explicit expression for the vector of moments is unavailable, it can be performed in the case of copula families with explicit expressions for their moments. This is the case for the one-parameter Clayton, Gumbel, Normal and Frank bivariate copula families described in Table 1.

An investigation on the accuracy of $\hat{\theta}^{\tau}$ and its estimated version $\hat{\theta}$ has been performed in the light of their estimated RRMSE as estimated from 1000 replicates. This study can provide not only information on the loss of accuracy due to the use of the estimated version instead of the mathematical inversion, but also on the role of the number of simulated samples S on the performance of $\hat{\theta}$. To this end, the values

Table 1 Copula C_θ , Kendall's tau $\tau_C(\theta)$ and parameter space Θ for parametric families of one-parameter bivariate copulas

Family	$C_\theta(u_1, u_2)$	$\tau_C(\theta)$	Θ
Clayton	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$	$\frac{\theta}{\theta + 2}$	$(0, \infty)$
Gumbel	$\exp \left\{ - \left(\ln u_1 ^\theta + \ln u_2 ^\theta \right)^{1/\theta} \right\}$	$1 - \frac{1}{\theta}$	$[1, \infty)$
Normal*	$\Phi_\theta \left\{ \Phi^{-1}(u_1), \Phi^{-1}(u_2) \right\}$	$\frac{2}{\pi} \sin^{-1} \theta$	$[0, 1]$
Frank	$-\frac{1}{\theta} \ln \left\{ 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right\}$	$1 - \frac{4}{\theta} + \frac{4}{\theta^2} \int_0^\theta \frac{x}{e^x - 1} dx$	$\mathbb{R} \setminus \{0\}$

* In the expression of the Normal copula, Φ_θ is the cdf of the bivariate normal with correlation θ and Φ is the cdf of the standard univariate normal

Table 2 Estimation, based on 1000 replicates, of the relative root mean-squared error (RRMSE) of the inversion of Kendall’s tau estimator $\widehat{\theta}^\tau$ and of the simulated estimator $\widehat{\theta}$ when $S \in \{100, 250, 500\}$ for sample sizes $T \in \{50, 100\}$ under four one-parameter families of bivariate copulas

C	τ_C	$T = 50$			$T = 100$				
		$\widehat{\theta}^\tau$	$\widehat{\theta}_{100}$	$\widehat{\theta}_{250}$	$\widehat{\theta}_{500}$	$\widehat{\theta}^\tau$	$\widehat{\theta}_{100}$	$\widehat{\theta}_{250}$	$\widehat{\theta}_{500}$
Clayton	1/4	0.548	0.652	0.587	0.575	0.366	0.500	0.419	0.390
	1/2	0.345	0.393	0.358	0.350	0.220	0.285	0.254	0.237
	3/4	0.277	0.333	0.300	0.285	0.182	0.240	0.203	0.192
Gumbel	1/4	0.384	0.454	0.418	0.397	0.260	0.351	0.300	0.283
	1/2	0.145	0.168	0.154	0.150	0.109	0.144	0.123	0.116
	3/4	0.061	0.070	0.065	0.063	0.042	0.057	0.049	0.045
Normal	1/4	0.346	0.411	0.379	0.361	0.239	0.344	0.283	0.255
	1/2	0.114	0.138	0.121	0.118	0.075	0.104	0.084	0.079
	3/4	0.028	0.034	0.030	0.029	0.019	0.027	0.022	0.021
Frank	1/4	0.402	0.482	0.439	0.422	0.279	0.364	0.328	0.303
	1/2	0.223	0.252	0.240	0.230	0.149	0.196	0.166	0.157
	3/4	0.181	0.203	0.191	0.187	0.111	0.151	0.127	0.119

$S \in \{100, 250, 500\}$ have been considered; the corresponding estimators are referred respectively to $\widehat{\theta}_{100}$, $\widehat{\theta}_{250}$ and $\widehat{\theta}_{500}$. The results for sample sizes $T \in \{50, 100\}$ for a Kendall’s tau that belongs to $\{1/4, 1/2, 3/4\}$ are found in Table 2. Note that the expression of Kendall’s tau for Frank’s copula has been inverted numerically.

Looking at Table 2, one can see that the RRMSE are smaller when $T = 100$ compared to $T = 50$, as expected. It is also not a surprise that $\widehat{\theta}^\tau$ is more accurate than its simulated versions. However, the loss of efficiency for using $\mu_S(\theta)$ instead of $\mu(\theta)$ is rather mild as soon as $S = 250$. In fact, when $T = 100$, the average of the relative efficiency $\text{RRMSE}(\widehat{\theta}^\tau)/\text{RRMSE}(\widehat{\theta})$ over the twelve models is 74,0%, 87,2% and 93,3% when $S = 100, 250, 500$, respectively. Note finally that the RRMSE decreases as the level of dependence increases, i.e. as τ_C increases, for all the models.

4.3 On the use of higher moments of the MPIT

One could ask if considering higher moments of W can lead to better estimators. To answer this question, at least in part, let $\widehat{\theta}_{\{2\}}$ and $\widehat{\theta}_{\{1,2\}}$ be the simulated method-of-moments estimators based respectively on the second moment of W and on its first two moments. The estimated relative bias and relative root mean-squared error of these two estimators, as well as of $\widehat{\theta}$, are found in Table 3 for the one-parameter families of Table 1. The sample size is $T = 100$ and the number of simulated samples has been set to $S = 100$. It is reasonable to think that the use of higher values of S would not influence much the relative efficiency of the estimators. The conclusion from Table 3 is that using higher moments does not result in more efficient estimators. Even though the RRMSE of $\widehat{\theta}_{\{1,2\}}$ is slightly smaller than that of $\widehat{\theta}$ in some cases, the gain in accuracy is rather small. As a general rule, one can advocate the use of the

Table 3 Estimation, based on 1000 replicates, of the relative bias (RB) and relative root mean-squared error (RRMSE) of the simulated method-of-moments estimators based on the first, second, and first two moments of W when $T = 100$ and $S = 100$ under four one-parameter families of bivariate copulas

C	τ_C	RB			RRMSE		
		$\hat{\theta}$	$\hat{\theta}_{\{2\}}$	$\hat{\theta}_{\{1,2\}}$	$\hat{\theta}$	$\hat{\theta}_{\{2\}}$	$\hat{\theta}_{\{1,2\}}$
Clayton	1/4	0.0247	0.0432	0.0048	0.5057	0.6192	0.4963
	1/2	0.0166	0.0347	0.0303	0.2903	0.3437	0.3146
	3/4	0.0188	0.0132	0.0097	0.2391	0.2715	0.2518
Gumbel	1/4	-0.0350	-0.0415	-0.0502	0.3687	0.3692	0.3715
	1/2	-0.0152	-0.0130	-0.0178	0.1414	0.1443	0.1410
	3/4	-0.0119	-0.0092	-0.0091	0.0613	0.0615	0.0571
Normal	1/4	-0.0308	-0.0454	-0.0362	0.3325	0.3519	0.3287
	1/2	-0.0146	-0.0168	-0.0135	0.1031	0.1129	0.1093
	3/4	-0.0033	-0.0060	-0.0038	0.0255	0.0303	0.0260
Frank	1/4	-0.0267	-0.0106	-0.0248	0.3776	0.4311	0.3755
	1/2	0.0127	0.0154	0.0151	0.2028	0.2161	0.2025
	3/4	0.0052	0.0101	0.0114	0.1568	0.1821	0.1536

simulated method-of-moments estimator with the same number of moments of W as the number of parameters to estimate.

4.4 Comparison with competing procedures

In this section, the performance of $\hat{\theta}$ will be compared to two other semi-parametric procedures, namely the pseudo-maximum likelihood (PML) estimator and the simulated method-of-moments estimator of Oh and Patton (2013). Other possibilities exist, for instance minimum distance (MD) estimators derived from goodness-of-fit criteria. Based on an extensive simulation study, Weiß (2011) concluded that no MD estimator stands out and are worse than the PML estimator in terms of bias and mean-squared error. For that reason, no MD estimator will be considered in the upcoming simulations study.

The PML estimator is the rank-based version of the maximum-likelihood estimator and is sometimes referred to the *canonical maximum likelihood* (see Cherubini et al. 2004, for instance). If C_θ admits a density c_θ , the PML estimator of θ as defined by Shih and Louis (1995) and Genest et al. (1995) is

$$\hat{\theta}_{\text{ML}} = \arg \max_{\theta \in \Theta} \sum_{i=1}^T \ln c_\theta \left(\frac{R_{i1}}{T+1}, \dots, \frac{R_{id}}{T+1} \right),$$

where for each $j \in \{1, \dots, d\}$, R_{ij} is the rank of X_{ij} among X_{1j}, \dots, X_{Tj} . The estimator of Oh and Patton (2013) is based on the vector of statistics $\mathbf{v}_T = (\rho_T^{\text{Sp}}, \lambda_T^{.05}, \lambda_T^{.10}, \lambda_T^{.90}, \lambda_T^{.95})$, where ρ_T^{Sp} is Spearman's rank correlation and

$$\lambda_T^q = \begin{cases} \frac{C_T(q, q)}{q}, & \text{when } q \leq 1/2; \\ \frac{1 - 2q + C_T(q, q)}{1 - q}, & \text{when } q > 1/2, \end{cases}$$

where C_T is the bivariate empirical copula. Recall that

$$C_T(u_1, u_2) = \frac{1}{T} \sum_{i=1}^T \mathbb{I} \left(\frac{R_{i1}}{T+1} \leq u_1, \frac{R_{i2}}{T+1} \leq u_2 \right).$$

Letting $\mathbf{v}_S(\boldsymbol{\theta})$ be a version of \mathbf{v}_T based on a sample of size S from $C_{\boldsymbol{\theta}}$, the estimator of Oh and Patton (2013) is

$$\hat{\boldsymbol{\theta}}_{\text{OP}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \{\mathbf{v}_T - \mathbf{v}_S(\boldsymbol{\theta})\}^T M_T \{\mathbf{v}_T - \mathbf{v}_S(\boldsymbol{\theta})\}.$$

As suggested by Oh and Patton (2013), the weight matrix is set to $M_T = \mathbf{I}_5$.

4.4.1 One-parameter families

The performances of $\hat{\boldsymbol{\theta}}_{\text{ML}}$, $\hat{\boldsymbol{\theta}}_{\text{OP}}$ and $\hat{\boldsymbol{\theta}}$ have been compared under the four families of Tables 1. The maximization of the pseudo-likelihood for the computation of $\hat{\boldsymbol{\theta}}_{\text{ML}}$ uses the MATLAB procedure `FMINSEARCHBND`. To increase the numerical stability, the computation of Gumbel’s density uses the finite-difference approximation $c_{\boldsymbol{\theta}}^{\text{Gu}}(u_1, u_2) \approx C_{\boldsymbol{\theta}}^{\text{Gu}}(u_1 + \epsilon, u_2 + \epsilon) + C_{\boldsymbol{\theta}}^{\text{Gu}}(u_1, u_2) - C_{\boldsymbol{\theta}}^{\text{Gu}}(u_1 + \epsilon, u_2) - C_{\boldsymbol{\theta}}^{\text{Gu}}(u_1, u_2 + \epsilon)$, where $\epsilon = 1 \times 10^{-8}$. The four copula models are parametrized in term of Kendall’s tau $\tau_C \in [0, 1]$ and the optimum is searched inside $[.01, .99]$ with the initial value $x_0 = 1/2$. The results on the relative bias and relative root mean-squared error are found in Table 4 when $T = 100$. The number of simulated samples has been set to $S = 500$.

First note that the estimator $\hat{\boldsymbol{\theta}}_{\text{OP}}$ of Oh and Patton (2013) is significantly more biased than $\hat{\boldsymbol{\theta}}$, except when $\tau_C = 1/4$ for the Normal and Frank copulas. This difference in accuracy is also reflected in their respective RRMSE. As expected, the pseudo-maximum likelihood (PML) estimator performs generally better than the simulated method-of-moments (SMM) estimator in terms of RRMSE. However, the average of the relative efficiency over the twelve models is 87,8%, so the loss of efficiency for using $\hat{\boldsymbol{\theta}}$ instead of $\hat{\boldsymbol{\theta}}_{\text{ML}}$ is small. Interestingly, the relative bias of the SMM estimator is often smaller than that of the PML estimator when the level of dependence is small or moderate, i.e., when $\tau \in \{1/4, 1/2\}$. Overall, the SMM estimator can be safely recommended when the use of the PML estimator is problematic, e.g., when the density is not tractable and/or when there is a high number of parameters to estimate.

4.4.2 The two-parameters chi-square copula

As defined by Bárdossy (2006) and later investigated by Quessy et al. (2016), the d -variate chi-square copula is the dependence structure of $\mathbf{X} = ((Z_1 + \gamma_1)^2, \dots, (Z_d +$

Table 4 Estimation, based on 1000 replicates, of the relative bias (RB) and relative root mean-squared error (RRMSE) of $\hat{\theta}_{ML}$, $\hat{\theta}_{OP}$ and $\hat{\theta}$ when $T = 100$ and $S = 500$ under four one-parameter families of bivariate copulas

C	τ_C	RB $\hat{\theta}_{ML}$	$\hat{\theta}_{OP}$	$\hat{\theta}$	RRMSE $\hat{\theta}_{ML}$	$\hat{\theta}_{OP}$	$\hat{\theta}$
Clayton	1/4	0.0811	0.0057	-0.0017	0.3481	0.4732	0.4077
	1/2	0.0330	0.0913	0.0238	0.2029	0.5858	0.2331
	3/4	-0.0239	0.1146	0.0199	0.1574	0.6543	0.1938
Gumbel	1/4	0.0641	-0.0311	0.0082	0.2628	0.3610	0.2827
	1/2	0.0089	-0.0237	-0.0077	0.1014	0.1792	0.1139
	3/4	-0.0029	-0.0152	-0.0010	0.0385	0.1014	0.0436
Normal	1/4	0.0462	-0.0199	-0.0181	0.2305	0.3216	0.2590
	1/2	0.0130	-0.0204	-0.0049	0.0744	0.1367	0.0867
	3/4	-0.0014	-0.0172	-0.0015	0.0164	0.0539	0.0198
Frank	1/4	0.0094	-0.0058	-0.0111	0.2775	0.4068	0.3004
	1/2	0.0181	0.0755	0.0149	0.1429	0.3562	0.1557
	3/4	-0.0062	0.0659	0.0131	0.1086	0.4247	0.1238

γ_d)²), where (Z_1, \dots, Z_d) is d -variate normal with zero means, unit variances and correlation Σ , and $\gamma_1, \dots, \gamma_d \in [0, \infty)$ are non-centrality parameters. One recovers the Normal copula at the limit when $\gamma_1 = \dots = \gamma_d \rightarrow \infty$. Unlike the Normal, the chi-square copula is radially asymmetric.

The results in Table 5 concern the estimation of the parameters of the bivariate chi-square copula when $\gamma_1 = \gamma_2 = \gamma$, where one can find the estimated relative bias and relative root mean-squared error of $(\hat{\gamma}_{ML}, \hat{\theta}_{ML})$, $(\hat{\gamma}_{OP}, \hat{\theta}_{OP})$ and $(\hat{\gamma}, \hat{\theta})$ when $T = 100$ and $S = 100$. In that case, the density is

$$c_{\gamma, \theta}(u_1, u_2) = \frac{\phi_{\theta} \{h_{\gamma}(u_1) - \gamma, h_{\gamma}(u_2) - \gamma\} + \phi_{\theta} \{h_{\gamma}(u_1) + \gamma, -h_{\gamma}(u_2) + \gamma\}}{D_{\gamma}(u_1) D_{\gamma}(u_2)} + \frac{\phi_{\theta} \{h_{\gamma}(u_1) - \gamma, -h_{\gamma}(u_2) - \gamma\} + \phi_{\theta} \{h_{\gamma}(u_1) + \gamma, h_{\gamma}(u_2) + \gamma\}}{D_{\gamma}(u_1) D_{\gamma}(u_2)},$$

where ϕ_{θ} is the density of the bivariate normal distribution with zero means, unit variances and correlation $\theta \in [-1, 1]$, and for $G_{\gamma}(x) = \Phi(\sqrt{x} - \gamma) + \Phi(\sqrt{x} + \gamma) - 1$, $h_{\gamma}(u) = \{G_{\gamma}^{-1}(u)\}^{1/2}$ and $D_{\gamma}(u) = \phi\{h_{\gamma}(u) - \gamma\} + \phi\{h_{\gamma}(u) + \gamma\}$.

For the estimation of the non-centrality parameter γ , one observes that the accuracy of the three estimators, both in terms of RB and RRMSE, increase as γ increases. For a given value of γ , their accuracy also increases as τ_C increases, except when $\gamma = 2$. Looking at the relative performance of the three estimators, the PML estimator is clearly the best. Overall, $\hat{\gamma}_{OP}$ is systematically, although slightly, more accurate than $\hat{\gamma}$. Turning to the estimation of the dependence parameter θ , one can say that $\hat{\theta}$ stands out positively from its two competitors when $\gamma \in \{3/2, 2\}$. When $\gamma \in \{1/2, 1\}$, there

Table 5 Estimation, based on 1000 replicates, of the relative bias (RB) and relative root mean-squared error (RRMSE) of the estimators $\hat{\theta}_{ML}$, $\hat{\theta}_{OP}$ and $\hat{\theta}$ of θ and the estimators $\hat{\gamma}_{ML}$, $\hat{\gamma}_{OP}$ and $\hat{\gamma}$ of the non-centrality parameter γ under the two-parameter chi-square copula when $T = 100$ and $S = 100$

γ	θ	RB			RRMSE		
		$\hat{\gamma}_{ML}$	$\hat{\gamma}_{OP}$	$\hat{\gamma}$	$\hat{\gamma}_{ML}$	$\hat{\gamma}_{OP}$	$\hat{\gamma}$
1/2	1/4	1.7125	2.9937	3.5184	2.6608	3.2004	3.6904
	1/2	0.3959	2.5665	3.0818	1.5394	2.7305	3.2189
	3/4	-0.0401	2.0579	2.3270	0.9758	2.2267	2.4378
1	1/4	0.3113	0.9042	1.1449	0.9833	1.0491	1.2460
	1/2	-0.0568	0.7229	0.8575	0.6934	0.8308	0.9352
	3/4	-0.1678	0.5384	0.5762	0.5726	0.6917	0.7121
3/2	1/4	0.0524	0.2670	0.3949	0.5725	0.4199	0.4990
	1/2	-0.0229	0.1346	0.1900	0.4823	0.2811	0.2914
	3/4	-0.0360	0.0612	0.0370	0.3902	0.2992	0.3247
2	1/4	-0.1330	-0.0384	0.0414	0.4478	0.2417	0.2404
	1/2	-0.1176	-0.1288	-0.1136	0.3802	0.2148	0.2008
	3/4	-0.0672	-0.1508	-0.2460	0.3051	0.2605	0.3415

γ	θ	RB			RRMSE		
		$\hat{\theta}_{ML}$	$\hat{\theta}_{OP}$	$\hat{\theta}$	$\hat{\theta}_{ML}$	$\hat{\theta}_{OP}$	$\hat{\theta}$
1/2	1/4	0.1788	0.1720	-0.3090	0.6529	0.7123	0.5854
	1/2	0.0246	-0.1667	-0.4033	0.2787	0.3288	0.4776
	3/4	0.0039	-0.2272	-0.2569	0.1077	0.2700	0.2872
1	1/4	0.5125	0.3673	-0.0716	0.7634	0.7437	0.5291
	1/2	0.1566	-0.0312	-0.1574	0.2774	0.2375	0.2742
	3/4	0.0515	-0.1613	-0.0912	0.1050	0.2178	0.1445
3/2	1/4	0.6101	0.4414	0.0671	0.7925	0.7545	0.5029
	1/2	0.2059	0.0257	-0.0221	0.2772	0.2140	0.2080
	3/4	0.0613	-0.1234	-0.0156	0.0891	0.1910	0.1072
2	1/4	0.6764	0.4375	0.0955	0.8520	0.7324	0.5362
	1/2	0.2603	0.0524	0.0109	0.3119	0.2139	0.2065
	3/4	0.0831	-0.1108	0.0027	0.1004	0.1873	0.1095

is no clear trend as to which estimator performs better overall. Generally speaking, $\hat{\theta}$ is the best when $\tau_C = 1/4$, $\hat{\theta}^{OP}$ when $\tau_C = 1/2$ and $\hat{\theta}^{ML}$ when $\tau_C = 3/4$.

4.5 Multivariate models

4.5.1 The Archimedean family

A d -dimensional copula is a member of the Archimedean family if it can be expressed in the form $C_\Psi(\mathbf{u}) = \Psi \{ \Psi^{-1}(u_1) + \dots + \Psi^{-1}(u_d) \}$, where $\Psi : [0, \infty) \rightarrow [0, 1]$ is called the generator and satisfies $(-1)^j \Psi^{[j]} \geq 0$ for each $j \in \{1, \dots, d\}$, where

$\Psi^{[j]}(t) = \partial^j \Psi(t) / \partial^j t$. See McNeil and Nešlehová (2009) for more details. The Clayton, Gumbel and Frank copulas whose bivariate versions are detailed in Table 1 are particular cases of this class, where the generators are respectively $\Psi_\theta^{Cl}(t) = (\theta t + 1)^{-1/\theta}$, $\Psi_\theta^{Gu}(t) = e^{-t^{1-\theta}}$ and

$$\Psi_\theta^{Fr}(t) = -\frac{1}{\theta} \ln \left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right).$$

If it is assumed that the d -variate copula of a population belongs to a given parametric Archimedean family with $\theta \in \Theta \subset \mathbb{R}^L$, then θ_0 can be estimated by the simulated method-of-moments estimator in (4) based on the first L moments of W . Alternatively, since θ_0 appears in the distribution of any possible pair of variables, another estimator similar to the one in (5) is

$$\hat{\theta}^* = \binom{d}{2}^{-1} \sum_{j < j' \in \{1, \dots, d\}} \hat{\theta}_{jj'}.$$

Table 6 reports the performance of $\hat{\theta}$ and $\hat{\theta}^*$ in terms of relative bias and RRMSE when $T = 100$ and $S = 250$ for the multivariate one-parameter Clayton, Gumbel and Frank copulas in dimensions $d \in \{3, 4, 5\}$. First observe that the RRMSE decreases as the dimension d increases. Note also the drastic decrease in RRMSE as Kendall's tau passes from 1/4 to 1/2, then a slightly smaller decrease for the passage from 1/2 to 3/4. Note also that in most cases, $\hat{\theta}$ is less biased than $\hat{\theta}^*$. However, the opposite is true when looking at the RRMSEs, although the differences are minimal. In short, it can be concluded that both estimation strategies work well.

4.5.2 Elliptical copulas and their squared versions

Elliptical distributions are parametrized in a pair-by-pair fashion. Specifically, as initially defined by Cambanis et al. (1981), a vector $\mathbf{X} \in \mathbb{R}^d$ is said to follow an elliptically contoured distribution if it admits the stochastic representation $\mathbf{X} = R\mathbf{A}\mathcal{U}$, where $R > 0$ is the radial random variable, $\Sigma = \mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{d \times d}$ is symmetric and positive definite and \mathcal{U} is uniformly distributed on the unit sphere in \mathbb{R}^d . An elliptical copula is simply the copula extracted from an elliptical distribution, as first investigated by Fang et al. (2002). Elliptical copulas then inherit from the pairwise parametrization of elliptical distributions.

Apart from the Normal copula, a popular elliptical model is Student's copula with $\gamma \in (0, \infty)$ degrees of freedom and parameter $\theta \in (-1, 1)$, which can be expressed implicitly by $C_{\gamma, \theta}(u_1, u_2) = \Omega_{\gamma, \theta} \left\{ \Omega_\gamma^{-1}(u_1), \Omega_\gamma^{-1}(u_2) \right\}$, where Ω_γ is the cumulative distribution function of the univariate Student and $\Omega_{\gamma, \theta}$ is the cdf of Student's bivariate distribution. Another model is the generalized Laplace copula (see Kozubowski et al.

Table 6 Estimation, based on 1000 replicates, of the relative bias (upper panel) and root mean-squared error (lower panel) of the simulated method-of-moments estimators $\hat{\theta}$ and $\hat{\theta}^*$ when $T = 100$ and $S = 250$ under one-parameter multivariate Archimedean copulas

C	τ_C	$d = 3$		$d = 4$		$d = 5$	
		$\hat{\theta}$	$\hat{\theta}^*$	$\hat{\theta}$	$\hat{\theta}^*$	$\hat{\theta}$	$\hat{\theta}^*$
Clayton	1/4	0.0255	0.0196	0.0230	0.0212	0.0201	0.0239
	1/2	0.0110	0.0232	0.0118	0.0200	0.0084	0.0206
	3/4	0.0090	0.0185	0.0133	0.0190	0.0130	0.0231
Gumbel	1/4	0.0080	-0.0088	0.0181	-0.0041	0.0117	-0.0046
	1/2	0.0010	-0.0065	-0.0017	-0.0041	-0.0029	-0.0080
	3/4	-0.0015	-0.0022	-0.0040	-0.0048	-0.0013	-0.0025
Frank	1/4	0.0073	0.0002	0.0057	-0.0031	-0.0037	-0.0101
	1/2	0.0103	0.0071	0.0100	0.0076	0.0057	0.0063
	3/4	-0.0011	0.0004	0.0046	0.0079	0.0024	0.0073
Clayton	1/4	0.3002	0.2882	0.2662	0.2581	0.2382	0.2152
	1/2	0.2099	0.2001	0.1903	0.1771	0.1714	0.1606
	3/4	0.1723	0.1658	0.1671	0.1520	0.1563	0.1433
Gumbel	1/4	0.2333	0.2245	0.2048	0.1981	0.1880	0.1729
	1/2	0.1009	0.0933	0.0948	0.0861	0.0880	0.0822
	3/4	0.0416	0.0395	0.0388	0.0355	0.0348	0.0317
Frank	1/4	0.2330	0.2336	0.1942	0.1853	0.1730	0.1638
	1/2	0.1277	0.1244	0.1139	0.1076	0.1064	0.0979
	3/4	0.0936	0.0911	0.0894	0.0803	0.0816	0.0744

2013, for details) extracted from the multivariate distribution whose density is

$$f_{\Sigma, \gamma}(\mathbf{x}) = \frac{2 |\Sigma|^{-1/2}}{(2\pi)^{d/2} \Gamma(\gamma)} \left(\sqrt{\frac{\mathbf{x}\Sigma^{-1}\mathbf{x}^\top}{2}} \right)^{\gamma-d/2} K_{\gamma-d/2} \left(\sqrt{2\mathbf{x}\Sigma^{-1}\mathbf{x}^\top} \right),$$

with K_λ the modified Bessel function of index λ .

As is well known, Kendall’s tau of a bivariate elliptical copula is given by the simple formula $\tau_C(\theta) = (2/\pi) \sin^{-1} \theta$, whatever the form of the radial random variable (see Fang et al. 2002, for instance). Hence, $\mathbb{E}(W) = \{\tau_C(\theta) + 1\}/4$ does not depend on R . As is stated in the next proposition, it is indeed the case for any moment of order $a \in \mathbb{N}$ of W . This theoretical result is illustrated in the two top panels of Fig. 1, where one can find the estimated curves of $\mathbb{E}(W^2)$ as a function of $\theta \in (0, 1)$ for various values of γ in the case of the Student and Laplace copulas.

Proposition 6 *The moment of order $a \in \mathbb{N}$ of an elliptical copula characterized by some radial variable R does not depend on R .*

In the light of Proposition 6, it is not possible to estimate (γ, θ) with the SMM estimator based on the MPIT random variable. It is however possible to estimate the

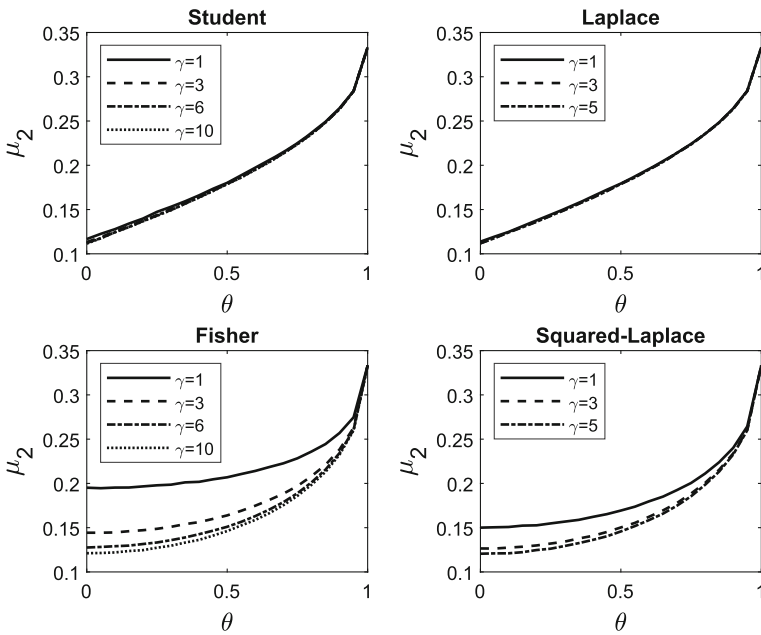


Fig. 1 Moment of order two of the probability integral random variable as a function of $\theta \in [0, 1]$ for the bivariate Student, Laplace, Fisher and Squared-Laplace copulas

parameters of the so-called squared version of an elliptical copula. As defined by Quesy and Durocher (2019), the squared copula associated to a d -variate copula C is the distribution of $(|1 - 2U_1|, \dots, |1 - 2U_d|)$ when $(U_1, \dots, U_d) \sim C$. When C is an elliptical copula, the resulting squared construction has a pair-by-pair parametric structure $\Sigma \in \mathbb{R}^{d \times d}$ and a global parameter $\boldsymbol{\gamma} \in \mathbb{R}^d$. See the bottom panels of Fig. 1 for $\mathbb{E}(W^2)$ as a function of θ for the squared-Student (introduced by Favre et al. (2018) as the Fisher copula) and squared-Laplace copulas.

As is explained in Sect. 3.3, the entries of Σ can be estimated using the first $q + 1$ moments of $W_{jj'}$ for each $j < j' \in \{1, \dots, d\}$, and then estimate $\boldsymbol{\gamma}$ with the mean of $\widehat{\boldsymbol{\gamma}}_{jj'}, j < j' \in \{1, \dots, d\}$, i.e., with the estimator in (5). The results in Table 7 concern the Fisher and squared-Laplace copulas in dimension $d \in \{3, 4, 5\}$ when $T = 100$ and $S = 100$. To simplify the presentation of results, the correlation matrix $\Sigma \in \mathbb{R}^{d \times d}$ has been taken equicorrelated in such a way that $\Sigma_{jj'} = \sin(\pi \tau_C / 2)$ for each $j \neq j' \in \{1, \dots, d\}$, where $\tau_C \in \{1/4, 1/2, 3/4\}$. From experiences not presented here, considering more general correlation matrices has a negligible influence on the performance of the estimator.

Looking at the results in Table 7, it can be seen that the performance of $\widehat{\boldsymbol{\gamma}}$ is, more or less, equivalent for the three dimensions considered. For the Fisher copula, the accuracy significantly increases as τ_C increases when $\gamma \in \{3, 6\}$; when $\gamma = 10$, the estimator is the most accurate when $\tau = 1/2$. Turning to the squared-Laplace copula, the accuracy of the estimator in relation to the level of Kendall's tau depends on γ .

Table 7 Estimation, based on 1000 replicates, of the relative bias (RB) and relative root mean-squared error (RRMSE) of the simulated estimator $\hat{\gamma}$ when $T = 100$ and $S = 100$ under multivariate two-parameters pairwise copulas

Copula	τ_C	$d = 3$		$d = 4$		$d = 5$	
		RB	RRMSE	RB	RRMSE	RB	RRMSE
Fisher $\gamma = 3$	1/4	2.5443	2.5977	2.4679	2.4973	2.7798	2.7955
	1/2	1.6341	1.7970	1.5873	1.6788	2.1137	2.1452
	3/4	1.0611	1.4622	1.0293	1.2318	1.2151	1.3112
Fisher $\gamma = 6$	1/4	1.0403	1.0708	1.0529	1.0689	1.0637	1.0735
	1/2	0.6407	0.7087	0.6384	0.6754	0.6417	0.6651
	3/4	0.1372	0.4840	0.1522	0.3458	0.1291	0.2844
Fisher $\gamma = 10$	1/4	0.2854	0.3222	0.2728	0.2929	0.2699	0.2837
	1/2	0.0018	0.1702	0.0112	0.1208	0.0005	0.1028
	3/4	-0.3341	0.4222	-0.3271	0.3798	-0.3295	0.3605
Sq.-Laplace $\gamma = 1$	1/4	3.4904	3.6190	3.4298	3.5077	4.1996	4.2262
	1/2	2.0727	2.4733	2.5566	2.6986	3.2914	3.3478
	3/4	1.7074	2.2759	1.6146	1.9372	2.0355	2.1776
Sq.-Laplace $\gamma = 3$	1/4	0.7452	0.7904	0.8818	0.9018	1.0008	1.0117
	1/2	0.1156	0.4185	0.1771	0.3253	0.5867	0.6135
	3/4	-0.1160	0.5461	-0.1399	0.3956	0.0506	0.2526
Sq.-Laplace $\gamma = 5$	1/4	0.2341	0.2763	0.2359	0.2613	0.2398	0.2563
	1/2	-0.0290	0.1852	-0.0258	0.1420	-0.0407	0.1133
	3/4	-0.3540	0.4456	-0.3680	0.4194	-0.3683	0.3971

Hence, whereas the accuracy increases as τ_C increases when $\gamma = 1$, there is no clear trend when $\gamma \in \{3, 5\}$.

5 Data analysis of hockey data

Ice hockey is a fast-paced team game that is played continuously and for which the measure of a player’s quality with appropriate indicators is a real challenge. For the illustration that follows, the $T = 410$ forwards that played at least 16,000 s in the 2019–2020 season of the National Hockey League (NHL) have been considered. Five variables, namely X_1 : *Points*, X_2 : *Expected goals (xg) with a player on*, X_3 : *Playing in attack*, X_4 : *Scoring chances* and X_5 : *Number of shots*, have been selected to characterize their offensive skills. Each variable have been rescaled to a block of 60 min. The pairwise scatterplots of the raw data and of the standardized ranks is found in Fig. 2.

Looking at Fig. 2, a radially asymmetric dependence structure featuring more weights in the upper tails seems to emerge. This is confirmed by the test of radial symmetry of Bahraoui and Quessy (2017) based on the copula characteristic function with the normal weight and smoothing parameter $\sigma = 1$. Indeed, the p-value of the

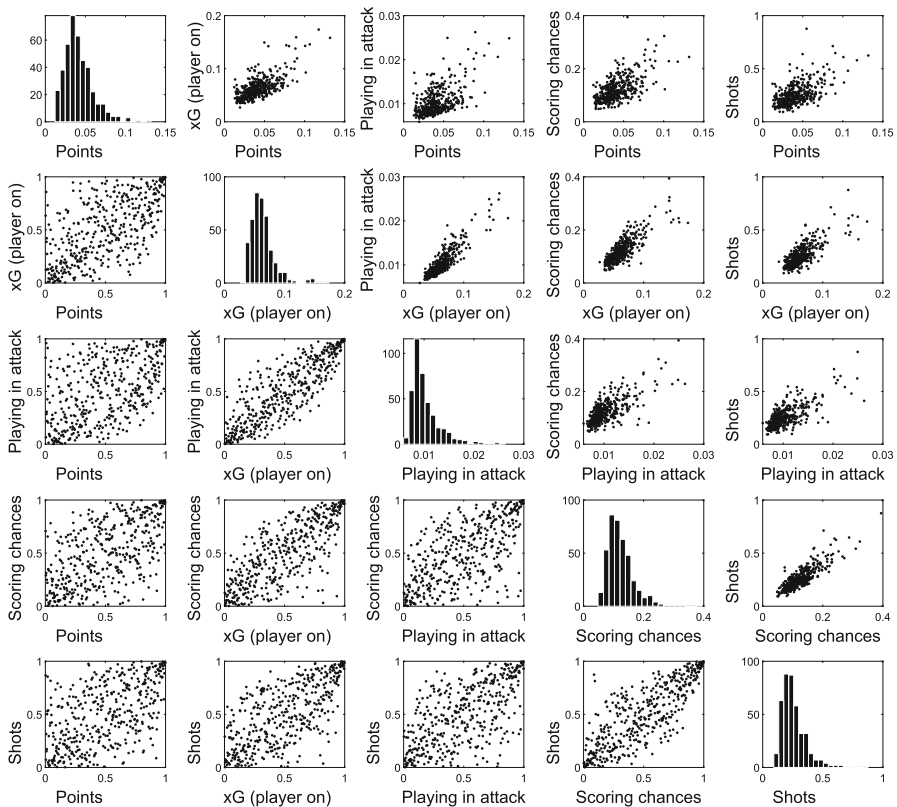


Fig. 2 Five-dimensional Hockey data set: histograms (on the diagonal), pairwise scatterplots of the raw data (above the diagonal) and of standardized ranks (below the diagonal)

test, as estimated from 10,000 multiplier bootstrap samples, is 4.61%. For this reason, parameter estimation has been performed for the following seven radially asymmetric copula families in order to capture this radially asymmetric behaviour:

- (i) The one-parameter survival-Clayton and Gumbel Archimedean copulas;
- (ii) The chi-square copula with non-centrality parameter $\gamma \in [0, \infty)$;
- (iii) The squared versions of the Student, Laplace and Pearson type II copulas;
- (iv) A special case of the skew-Student copula as defined by Demarta and McNeil (2005), i.e., the dependence structure of $\mathbf{X} = \mathbf{Z}/\sqrt{Y} + \gamma_2 \mathbf{1}_d/Y$, where \mathbf{Z} is standard normal with correlation $\Sigma \in \mathbb{R}^{d \times d}$, $\gamma_1 Y$ is chi-square with $\gamma_1 \geq 1$ degrees of freedom, $\gamma_2 \in \mathbb{R}$ is an asymmetry parameter and $\mathbf{1}_d = (1, \dots, 1)$.

The results of the estimation based on the simulated method-of-moments estimator performed with $S = 250$ are in Table 8. As a criterion for choosing an appropriate model among the seven copula families, the ability of the model to reproduce Kendall's matrix has been considered. Specifically, a sample of size $T = 2,500$ have been simulated from each estimated model and the Frobenius matrix distance between the sample Kendall matrix K_T of the data and that of the simulated sample have been

Table 8 Results of the parameter estimation for the five-dimensional Hockey data set

Copula family	Estimated parameters	Frobenius distance from the empirical Kendall matrix
Survival Clayton	$\hat{\theta} = 1.7310$	0.4463
Gumbel	$\hat{\theta} = 0.5384$	0.5004
Chi-square	$\hat{\gamma} = 1.4139$	0.1212
Fisher	$\hat{\gamma} = 6.4215$	0.7377
Squared–Laplace	$\hat{\gamma} = 2.2789$	0.7409
Squared–Pearson type II	$\hat{\gamma} = 3.2736$	0.8567
Skew–Student	$\hat{\gamma}_1 = 10.6707, \hat{\gamma}_2 = 2.5951$	0.2845

computed. For the Hockey data,

$$K_T = \begin{pmatrix} 1 & 0.460 & 0.359 & 0.411 & 0.378 \\ 0.460 & 1 & 0.671 & 0.583 & 0.502 \\ 0.359 & 0.671 & 1 & 0.476 & 0.470 \\ 0.411 & 0.583 & 0.476 & 1 & 0.623 \\ 0.378 & 0.502 & 0.470 & 0.623 & 1 \end{pmatrix}.$$

Looking at the results in the third column of Table 8, the Chi-square (χ^2) and skew–Student (Sk) copulas stand out among the seven models considered. In order to see if these models are good to reproduce the observed data, artificial samples of size $T = 410$ from both copula models have been simulated at the estimated values. The corresponding estimations of Σ are

$$\hat{\Sigma}^{\chi^2} = \begin{pmatrix} 1 & .720 & .536 & .639 & .575 \\ .720 & 1 & .838 & .827 & .762 \\ .536 & .838 & 1 & .731 & .735 \\ .639 & .827 & .731 & 1 & .864 \\ .575 & .761 & .735 & .864 & 1 \end{pmatrix}, \quad \hat{\Sigma}^{\text{Sk}} = \begin{pmatrix} 1 & .015 & .016 & .022 & .000 \\ .015 & 1 & .396 & .334 & .410 \\ .016 & .396 & 1 & .225 & .102 \\ .022 & .334 & .225 & 1 & .332 \\ .000 & .410 & .102 & .332 & 1 \end{pmatrix}.$$

Note that $\hat{\Sigma}^{\text{Sk}}$ results from a transformation due to Higham (2002) to make it positive definite. The resulting samples have then been put on the same scales as the raw data by taking the empirical percentiles. Their scatterplots (raw data and standardized ranks) are found in Fig. 3 and Fig. 4, respectively. Whereas the chi-square copula is better at reproducing Kendall’s matrix, the skew–Student copula seems better at reproducing the upper tail behaviour of the dependence structure.

6 Conclusion

This paper developed a general parameter estimation procedure for multivariate copula models. The proposed estimators are based on the moments of the multivariate probability integral transformation (MPIT). It then generalizes the inversion of Kendall’s

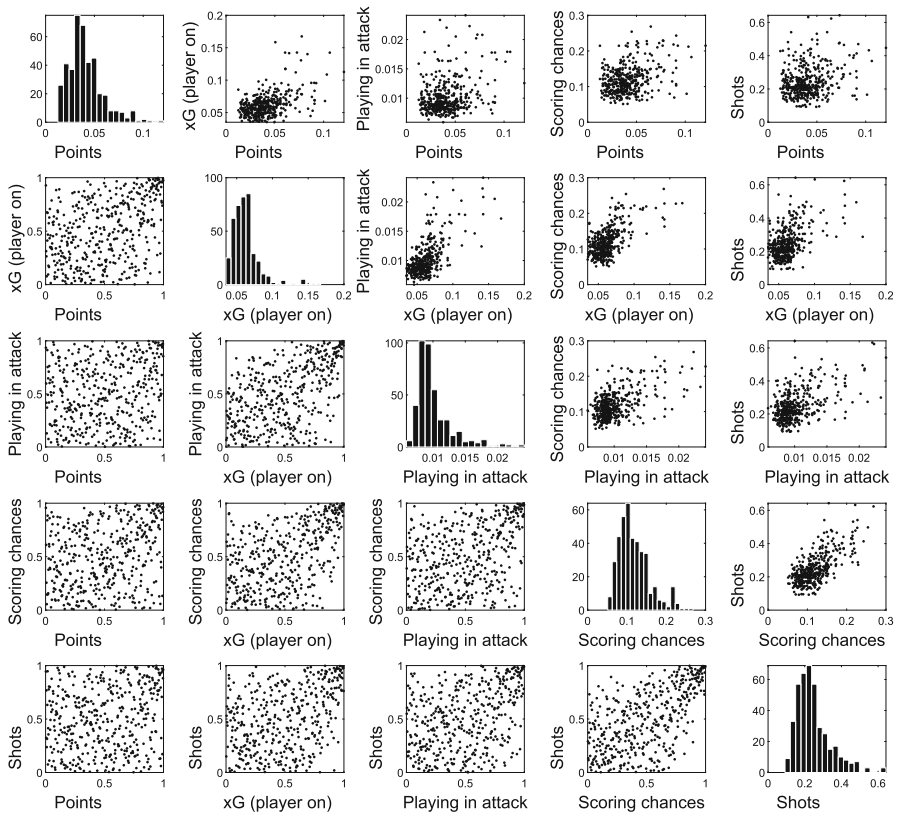


Fig. 3 Simulated data of size $T = 410$ from the estimated Chi-square copula with marginal distributions matching those of the Hockey data set

tau estimator. On one hand, moments of order greater than one are considered, making possible the estimation in multi-parameters models. On the other hand, the use of simulated moments make the methodology apply as soon as it is possible to simulate from a given parametric model. Compared to Brahimi and Necir (2012), the proposed estimators are not restricted to (the few) cases where the mapping induced by the theoretical moments is explicitly invertible. Oh and Patton (2013) also developed simulated method-of-moments estimators in copula models. What can be seen as an advantage over their method: (i) it is no longer necessary to base the estimation on pairwise dependence measures and (ii) the number of moments can match the number of parameters of the assumed parametric copula model.

Knowing how to estimate for dimensions $d > 2$ and multi-parameter models is important, especially in the context of big data, which is becoming more important. However, the applicability of the pseudo-maximum likelihood estimator (i.e., any dimension d , any number of parameters to estimate) is more theoretical than practical. Indeed, even in cases when an explicit (and numerically tractable) copula density is available, the PMLE can be computationally very intensive, especially when d is large.

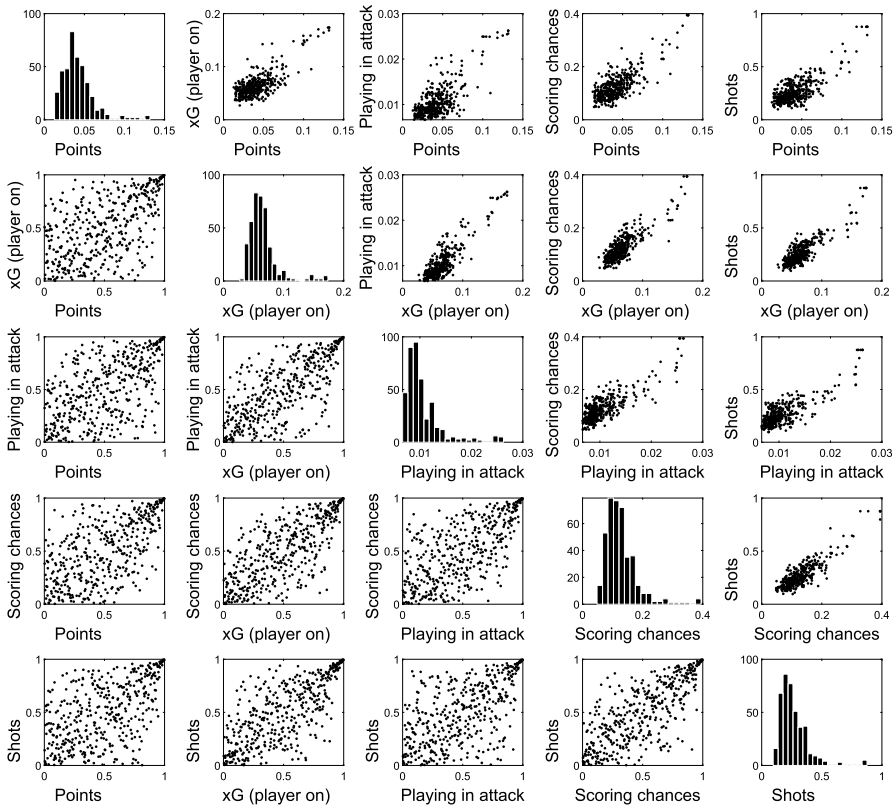


Fig. 4 Simulated data of size $T = 410$ from the estimated skew–Student copula with marginal distributions matching those of the Hockey data set

The flexibility of the method introduced in this work then appears to be of great interest for multi-parameter complex dependence models. One can mention the skew–Student copulas introduced by (Demarta and McNeil 2005), the factor copulas investigated for instance by Krupskii and Joe (2013), Krupskii and Joe (2015) and Mazo et al. (2016) as well as the vines copulas (see Czado 2019).

A Proofs

A.1 Proof of Proposition 1

The proof uses standard arguments, such as those in the statement of Theorem 4.1, p. 36, of van der Vaart (1998). By the definition of θ_T^K and a first-order Taylor expansion, $\mu_T^K = \mu^K(\theta_T^K) = \mu^K(\theta_0) + (\mu^K)'(\theta^*)(\theta_T^K - \theta_0^K)$, where θ^* lies between θ_T^K and θ_0 . One can then write $(\mu^K)'(\theta^*)\sqrt{T}(\theta_T^K - \theta_0) = \sqrt{T}(\mu_T^K - \mu^K(\theta_0))$. The stated result holds in view of Lemma 1 and since $(\mu^K)'(\theta^*)$ converges in probability to $(\mu^K)'(\theta_0) = v_0^K$. \square

A.2 Proof of Proposition 2

The difference here compared to standard method-of-moments techniques is the fact that the objective function is not continuous as a function of θ due to the estimated function $\mu_S(\theta)$. Basically, one has to appeal to the notion of stochastic equicontinuity in order to have some sort of uniform convergence for the objective function. Specifically, the proof consists in verifying the conditions of Theorem 2.1 of Newey and McFadden (1994) that ensure the consistency of minimum distance estimators of the form

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \widehat{Q}(\theta)$$

as long as Θ is compact and the objective function $\widehat{Q}(\theta)$ converges uniformly in probability to a function $Q_0(\theta)$ that is continuous and uniquely minimized at θ_0 .

In the case of the proposed simulated method-of-moments estimator, one has $\widehat{Q} := g_{T,S} M_T g_{T,S}^\top$, where $g_{T,S}(\theta) = \mu_T^k - \mu_S^k(\theta)$. Since $\mathbb{E}\{|\kappa(\mathbf{X}_1, \dots, \mathbf{X}_m)|\} < \infty$, $g_{T,S}(\theta)$ converges in probability to $g_0(\theta) = \mu^k(\theta_0) - \mu^k(\theta)$ as $T, S \rightarrow \infty$ for each $\theta \in \Theta$. Also, the assumption that μ^k is continuous entails the continuity of $Q_0 := g_0 M_0 g_0^\top$. Moreover, since $\mu^k(\theta_0) \neq \mu^k(\theta)$ as soon as $\theta \neq \theta_0$, the function Q_0 is uniquely minimized at θ_0 . It then remains to establish the uniform convergence in probability of \widehat{Q} to Q_0 . To this end, it will first be shown that $g_{T,S}$ is stochastically equicontinuous. Because $\mu_S^k(\theta)$ converges in probability to $\mu^k(\theta)$ for any fixed $\theta \in \Theta$, one has for any $\theta_1, \theta_2 \in \Theta$ that

$$\begin{aligned} g_{T,S}(\theta_1) - g_{T,S}(\theta_2) &= \mu_S^k(\theta_2) - \mu_S^k(\theta_1) \\ &= \mu^k(\theta_2) - \mu^k(\theta_1) + R_S(\theta_2) - R_S(\theta_1), \end{aligned}$$

where $(R_S(\theta_1), R_S(\theta_2))$ converges in probability to $(0, 0)$. Hence,

$$\|g_{T,S}(\theta_2) - g_{T,S}(\theta_1)\| \leq \|\mu^k(\theta_2) - \mu^k(\theta_1)\| + \|R_S(\theta_2) - R_S(\theta_1)\|.$$

Since μ^k is Lipschitz continuous, there exists $\zeta > 0$ such that $\|\mu^k(\theta_2) - \mu^k(\theta_1)\| \leq \zeta \|\theta_2 - \theta_1\|$, and then one can write

$$\|g_{T,S}(\theta_2) - g_{T,S}(\theta_1)\| \leq \zeta \|\theta_2 - \theta_1\| + \|R_S(\theta_2) - R_S(\theta_1)\| = \Lambda_S \|\theta_2 - \theta_1\|,$$

where $\Lambda_S = \zeta + \|R_S(\theta_2) - R_S(\theta_1)\|/\|\theta_2 - \theta_1\|$. Because μ_S^k is bounded, it follows that one can find $\xi > 0$ such that

$$\limsup_{S \rightarrow \infty} \mathbb{E} \left(\Lambda_S^{2+\xi} \right) = \limsup_{S \rightarrow \infty} \mathbb{E} \left(\zeta + \frac{|R_S(\theta_2) - R_S(\theta_1)|}{\|\theta_2 - \theta_1\|} \right)^{2+\xi} < \infty.$$

This means that $g_{T,S}$ is asymptotically Lipschitz continuous. Hence, $g_{T,S}$ is stochastically equicontinuous, i.e., for all $\epsilon, \eta > 0$, there exists $\delta > 0$ such that

$$\limsup_{T,S \rightarrow \infty} \mathbb{P} \left\{ \sup_{\|\theta_1 - \theta_2\|} \|g_{T,S}(\theta_1) - g_{T,S}(\theta_2)\| > \eta \right\} < \epsilon.$$

Invoking Lemma 2.8 of Newey and McFadden (1994), the conditions are met in order that as $S, T \rightarrow \infty$,

$$\sup_{\theta \in \Theta} \|g_{T,S}(\theta) - g_0(\theta)\| \xrightarrow{\mathbb{P}} 0. \tag{6}$$

Since $g_0(\theta)$ is bounded, $M_T = O_p(1)$ and M_T converges in probability to M_0 , an application of the triangular and Cauchy–Schwarz inequalities yield

$$\begin{aligned} \sup_{\theta \in \Theta} |\widehat{Q}(\theta) - Q_0(\theta)| &\leq \left\{ \|g_{T,S}(\theta) - g_0(\theta)\|^2 \right. \\ &\quad \left. + 2 \|g_0(\theta)\| \|g_{T,S}(\theta) - g_0(\theta)\| \right\} \|M_T\| \\ &\quad + \|g_0(\theta)\|^2 \|M_T - M_0\| \\ &\leq \left\{ \left(\sup_{\theta \in \Theta} \|g_{T,S}(\theta) - g_0(\theta)\| \right)^2 \right. \\ &\quad \left. + K \sup_{\theta \in \Theta} \|g_{T,S}(\theta) - g_0(\theta)\| \right\} \zeta_T + K^2 \kappa_T, \end{aligned}$$

where $K \in (0, \infty)$, $\zeta_T = O_p(1)$ and $\kappa_T = o_p(1)$. In view of (6), one can conclude that \widehat{Q} converges uniformly in probability to Q_0 . □

A.3 Proof of Lemma 2

Because $\prod_{j=1}^m b_j - \prod_{j=1}^m a_j = \sum_{j=1}^m (b_j - a_j) \prod_{k<j} a_k \prod_{k>j} b_k$, one can write

$$\begin{aligned} \|\mu^\kappa(\theta_1) - \mu^\kappa(\theta_2)\| &\leq \sum_{j=1}^m \int_{[0,1]^{dm}} \|\kappa(\mathbf{u}_1, \dots, \mathbf{u}_m)\| |c_{\theta_1}(\mathbf{u}_j) - c_{\theta_2}(\mathbf{u}_j)| \\ &\quad \times \prod_{k<j} c_{\theta_1}(\mathbf{u}_k) \prod_{k>j} c_{\theta_2}(\mathbf{u}_k) \, d\mathbf{u}_1 \cdots d\mathbf{u}_m \\ &\leq K \|\theta_1 - \theta_2\| \sum_{j=1}^m \int_{[0,1]^{dm}} \|\kappa(\mathbf{u}_1, \dots, \mathbf{u}_m)\| \\ &\quad \times \prod_{k<j} c_{\theta_1}(\mathbf{u}_k) \prod_{k>j} c_{\theta_2}(\mathbf{u}_k) \, d\mathbf{u}_1 \cdots d\mathbf{u}_m \end{aligned}$$

$$= K \|\theta_1 - \theta_2\| \sum_{j=1}^m \int_{[0,1]^d} \mathbb{E}\{\|\kappa(\mathbf{u}_1, \mathbf{U}_2, \dots, \mathbf{U}_m)\|\} \, d\mathbf{u}_1.$$

Since $\mathbb{E}\{\|\kappa(\mathbf{u}_1, \mathbf{U}_2, \dots, \mathbf{U}_m)\|\} < \infty$, one can conclude that μ^κ is Lipschitz continuous. □

A.4 Proof of Proposition 3

The proof consists in verifying the conditions of Theorem 7.2 of Newey and McFadden (1994) that ensures that asymptotic normality of minimum distance estimators of the form

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \hat{g}(\theta)^\top \hat{M} \hat{g}(\theta).$$

Specifically, $\sqrt{T}(\hat{\theta} - \theta_0)$ converges in distribution to the Normal law with mean zero and variance-covariance matrix Ω_0 as long as

(C₁) \hat{M} converges in probability to a positive definite matrix M_0 and $\hat{\theta}$ converges in probability to an interior point θ_0 of Θ ;

(C₂) $\sqrt{T} \hat{g}(\theta_0)$ converges to the mean zero normal law with variance-covariance matrix Σ ;

(C₃) $\hat{g}(\hat{\theta})^\top \hat{M} \hat{g}(\hat{\theta}) \leq \inf_{\theta \in \Theta} \hat{g}(\theta)^\top \hat{M} \hat{g}(\theta) + o_{\mathbb{P}}(T^{-1})$;

(C₄) There exists a function g_0 such that $g_0(\theta_0) = 0$, $g_0(\theta)$ is differentiable at θ_0 whose derivative G_0 is such that $G_0 M_0 G_0^\top$ is nonsingular, and for any $\delta_T \rightarrow 0$ as $T \rightarrow \infty$,

$$\sup_{\|\theta - \theta_0\| \leq \delta_T} \frac{\sqrt{T} \|\hat{g}(\theta) - \hat{g}(\theta_0) - g_0(\theta)\|}{1 + \sqrt{T} \|\theta - \theta_0\|} \xrightarrow{\mathbb{P}} 0.$$

In the case of the proposed simulated method-of-moments estimator, one has $\hat{M} := M_T$ and $\hat{g}(\theta) := \mu_T^\kappa - \mu_S^\kappa(\theta)$. First, (C₁) holds by the assumption on M_T and the fact that $\theta_{T,S}^\kappa$ converges in probability to θ_0 , as ensured by Proposition 2. To establish (C₂), define $Z_{1T} = \sqrt{T}\{\mu_T^\kappa - \mu^\kappa(\theta_0)\}$ and $Z_{2S} = \sqrt{S}\{\mu_S^\kappa(\theta_0) - \mu^\kappa(\theta_0)\}$, so that $\sqrt{T} g_{T,S}^\kappa(\theta_0) = Z_{1T} - \sqrt{T/S} Z_{2S}$. From an application of Proposition 1, $Z_{1T} \rightsquigarrow Z_1$ and $Z_{2S} \rightsquigarrow Z_2$ for Z_1, Z_2 i.i.d $\mathcal{N}(0, \Sigma_0)$, and then $\sqrt{T} g_{T,S}^\kappa(\theta_0)$ converges in distribution to $\mathcal{N}(0, (1 + \zeta)\Sigma_0)$, where $\zeta = \lim_{T,S \rightarrow \infty} T/S \in [0, \infty)$. Hence, (C₂) holds. Condition (C₃) holds by assumption.

Now to establish (C₄), it will shown that the function $\Upsilon_{T,S}^\kappa(\theta) = \sqrt{T}\{g_{T,S}^\kappa(\theta) - g_0^\kappa(\theta)\}$, where $g_0^\kappa(\theta) = \mu^\kappa(\theta_0) - \mu^\kappa(\theta)$, is stochastically equicontinuous. As formalized by Andrews (1994), a function v_T is stochastically equicontinuous at τ_0 if for all $\epsilon, \eta > 0$, there exists $\delta > 0$ such that

$$\limsup_{T \rightarrow \infty} \mathbb{P} \left\{ \sup_{\|\tau - \tau_0\| < \delta} \|v_T(\tau) - v_T(\tau_0)\| > \eta \right\} < \epsilon. \tag{7}$$

To establish the stochastic equicontinuity of $\Upsilon_{T,S}^\kappa$, Assumptions A–B of Theorem 1 of Andrews (1994) will be shown to hold. Firstly, Assumption A holds true since $\{g_{T,S}^\kappa(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ is a type II class of functions and satisfies Pollard’s entropy condition with envelope

$$\max \left\{ 1, \sup_{\boldsymbol{\theta} \in \Theta} \|g_{T,S}^\kappa(\boldsymbol{\theta})\|, \Lambda_S \right\},$$

where referring to (6), $\Lambda_S = \zeta + \|R_S(\boldsymbol{\theta}_2) - R_S(\boldsymbol{\theta}_1)\|/\|\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1\|$ for $(R_S(\boldsymbol{\theta}_1), R_S(\boldsymbol{\theta}_2))$ that converges to $(0, 0)$ in probability. Assumption B holds as well because $g_{T,S}^\kappa$ is bounded and there exists $\xi > 0$ such that

$$\limsup_{S \rightarrow \infty} \mathbb{E} \left(\Lambda_S^{2+\xi} \right) < \infty.$$

As a consequence, $\Upsilon_{T,S}^\kappa$ satisfies (7) at $\tau_0 = \boldsymbol{\theta}_0$. Then, since

$$\frac{\sqrt{T} \left\| g_{T,S}^\kappa(\boldsymbol{\theta}) - g_{T,S}^\kappa(\boldsymbol{\theta}_0) - g_0^\kappa(\boldsymbol{\theta}) \right\|}{1 + \sqrt{T} \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|} \leq \sqrt{T} \left\| g_{T,S}^\kappa(\boldsymbol{\theta}) - g_{T,S}^\kappa(\boldsymbol{\theta}_0) - g_0^\kappa(\boldsymbol{\theta}) \right\|,$$

one can conclude that for any $\epsilon > 0$,

$$\begin{aligned} & \mathbb{P} \left\{ \frac{\sqrt{T} \left\| g_{T,S}^\kappa(\boldsymbol{\theta}) - g_{T,S}^\kappa(\boldsymbol{\theta}_0) - g_0^\kappa(\boldsymbol{\theta}) \right\|}{1 + \sqrt{T} \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|} > \eta \right\} \\ & \leq \mathbb{P} \left\{ \sup_{\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \delta} \sqrt{T} \left\| g_{T,S}^\kappa(\boldsymbol{\theta}) - g_{T,S}^\kappa(\boldsymbol{\theta}_0) - g_0^\kappa(\boldsymbol{\theta}_0) \right\| > \eta \right\} < \epsilon. \end{aligned}$$

This establishes (\mathcal{C}_4) and concludes the proof. □

A.5 Proof of Proposition 4

Letting $F(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x})$,

$$\begin{aligned}
 \mathbb{E} \{ \mathcal{K}_a(\mathbf{X}_1, \dots, \mathbf{X}_{a+1}) \} &= \frac{1}{a+1} \sum_{j=1}^{a+1} \mathbb{P} \left\{ \left(\max_{\ell \in \{1, \dots, a+1\}, \ell \neq j} \mathbf{X}_\ell \right) < \mathbf{X}_j \right\} \\
 &= \mathbb{P} \{ \mathbf{X}_{a+1} > \max(\mathbf{X}_1, \dots, \mathbf{X}_a) \} \\
 &= \mathbb{P} (\mathbf{X}_1 < \mathbf{X}_{a+1}, \dots, \mathbf{X}_a < \mathbf{X}_{a+1}) \\
 &= \mathbb{E}_{\mathbf{X}_{a+1}} \left\{ \prod_{j=1}^a \mathbb{P}(\mathbf{X}_j < \mathbf{X}_{a+1}) \right\} \\
 &= \mathbb{E}_{\mathbf{X}_{a+1}} \left\{ \prod_{j=1}^a F(\mathbf{X}_{a+1}) \right\} \\
 &= \mathbb{E}_{\mathbf{X}} \{ (F(\mathbf{X}))^a \}.
 \end{aligned}$$

Since by definition, $W = F(\mathbf{X})$, where $\mathbf{X} \sim F$, it follows readily that

$$\mathbb{E} \{ \mathcal{K}_a(\mathbf{X}_1, \dots, \mathbf{X}_{a+1}) \} = \mathbb{E}(W^a) = \mu_a.$$

Hence, the U-statistic with kernel \mathcal{K}_a is unbiased for μ_a . To conclude the proof, simply observe that because $W = F(\mathbf{X}) = C(\mathbf{U})$, where $\mathbf{U} = (F_1(X_1), \dots, F_d(X_d)) \sim C$, it is clear that $\mathbb{E} \{ \mathcal{K}_a(\mathbf{X}_1, \dots, \mathbf{X}_{a+1}) \} = \mathbb{E}(W^a) = \mathbb{E} \{ \mathcal{K}_a(\mathbf{U}_1, \dots, \mathbf{U}_{a+1}) \}$, which establishes that \mathcal{K}_a is marginal-free. □

A.6 Proof of Proposition 5

The result is a special case of Lemma 1 and mainly consists in deriving an expression for $\mathcal{K}_a^*(\mathbf{u}) = (a+1)[\mathbb{E} \{ \mathcal{K}_a(\mathbf{u}, \mathbf{U}_2, \dots, \mathbf{U}_{a+1}) \} - \mu_a]$. Upon recalling that

$$\mathcal{K}_a(\mathbf{x}_1, \dots, \mathbf{x}_{a+1}) = \frac{1}{a+1} \sum_{j=1}^{a+1} \mathbb{I} \left\{ \left(\max_{\ell \in \{1, \dots, a+1\}, \ell \neq j} \mathbf{x}_\ell \right) < \mathbf{x}_j \right\},$$

one computes

$$\begin{aligned}
 \mathcal{K}_a^*(\mathbf{u}) &= (a+1) \left[\mathbb{E} \{ \mathcal{K}_a(\mathbf{u}, \mathbf{U}_2, \dots, \mathbf{U}_{a+1}) \} - \mu_a \right] \\
 &= \mathbb{E} \left\{ \mathbb{I} \left(\mathbf{u} > \max_{2 \leq \ell \leq a+1} \mathbf{U}_\ell \right) \right\} \\
 &\quad + a \mathbb{E} \{ \mathbb{I} (\mathbf{U}_1 > \max(\mathbf{u}, \mathbf{U}_2, \dots, \mathbf{U}_a)) \} - (a+1) \mu_a \\
 &= \mathbb{P}(\mathbf{U}_2 < \mathbf{u}, \dots, \mathbf{U}_{a+1} < \mathbf{u}) \\
 &\quad + a \mathbb{P} (\mathbf{U}_1 > \mathbf{u}, \mathbf{U}_1 > \mathbf{U}_2, \dots, \mathbf{U}_1 > \mathbf{U}_a) - (a+1) \mu_a \\
 &= \prod_{\ell=2}^{a+1} \mathbb{P}(\mathbf{U}_\ell \leq \mathbf{u})
 \end{aligned}$$

$$\begin{aligned}
 & +a \mathbb{E}_{\mathbf{U}_1} \left\{ \mathbb{I}(\mathbf{U}_1 > \mathbf{u}) \prod_{\ell=2}^a \mathbb{P}(\mathbf{U}_\ell < \mathbf{U}_1) \right\} - (a + 1) \mu_a \\
 & = \{C(\mathbf{u})\}^a + a \mathbb{E} \left\{ \mathbb{I}(\mathbf{U} > \mathbf{u}) (C(\mathbf{U}))^{a-1} \right\} - (a + 1) \mu_a.
 \end{aligned}$$

From an application of Theorem 2, p. 76 of Lee (1990), $\sqrt{T}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})$ is asymptotically L -variate Normal with vector of means $(\mathbb{E}\{\mathcal{K}_1(\mathbf{U})\}, \dots, \mathbb{E}\{\mathcal{K}_L(\mathbf{U})\}) = (0, \dots, 0)$ and variance-covariance matrix $\Sigma \in \mathbb{R}^{L \times L}$ such that for any $a, a' \in \{1, \dots, L\}$,

$$\Sigma_{aa'} = \text{Cov} \left\{ \mathcal{K}_a^*(\mathbf{U}), \mathcal{K}_{a'}^*(\mathbf{U}) \right\} = \mathbb{E} \left\{ \mathcal{K}_a^*(\mathbf{U}) \mathcal{K}_{a'}^*(\mathbf{U}) \right\},$$

which completes the proof. □

A.7 Proof of Proposition 6

First note that $\mathbb{E}(W^a) = \mathbb{P}(\mathbf{X}^* - \mathbf{X}_1 > 0, \dots, \mathbf{X}^* - \mathbf{X}_a > 0)$, where $\mathbf{X}^*, \mathbf{X}_1, \dots, \mathbf{X}_a$ are i.i.d. from the elliptical distribution characterized by the radial variable R . For each $j \in \{1, \dots, a\}$, the distribution of $\mathbf{Y}_j = \mathbf{X}^* - \mathbf{X}_j$ is elliptically contoured, so that it admits the stochastic representation $\mathbf{Y}_j = \mathcal{G}_j \mathbf{Z}_j$ for some positive random variable \mathcal{G}_j and where \mathbf{Z}_j is d -variate normal with some covariance matrix Σ . As a consequence, one has the stochastic representation $(\mathbf{X}^* - \mathbf{X}_1, \dots, \mathbf{X}^* - \mathbf{X}_a) = (\mathcal{G}_1 \mathbf{Z}_1, \dots, \mathcal{G}_a \mathbf{Z}_a)$, where $\mathbf{Z}_1, \dots, \mathbf{Z}_d$ are dependent normal vectors. One can then write $\mathbb{E}(W^a) = \mathbb{P}(\mathcal{G}_1 \mathbf{Z}_1 > 0, \dots, \mathcal{G}_a \mathbf{Z}_a > 0) = \mathbb{P}(Z_1 > 0, \dots, Z_a > 0)$, which ends the proof. □

Acknowledgements The authors acknowledge financial support by individual grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Declarations

Conflict of interest We state that there is no Conflict of interest.

References

Andrews DWK (1994) Asymptotics for semiparametric econometric models via stochastic equicontinuity. *Econometrica* 62:43–72

Bahraoui T, Quessy J-F (2017) Tests of radial symmetry for multivariate copulas based on the copula characteristic function. *Electron J Stat* 11:2066–2096

Bárdossy A (2006) Copula-based geostatistical models for groundwater quality parameters. *Water Resour Res* 42:1–12

Biau G, Wegkamp M (2005) A note on minimum distance estimation of copula densities. *Stat Probab Lett* 73:105–114

Brahimi B, Necir A (2012) A semiparametric estimation of copula models based on the method of moments. *Stat Methodol* 9:467–477

Cambanis S, Huang S, Simons G (1981) On the theory of elliptically contoured distributions. *J Multivar Anal* 11:368–385

Cherubini U, Luciano E, Vecchiato W (2004) *Copula methods in finance*. Wiley, Chichester

- Czado C (2019) Analyzing dependent data with vine copulas, vol. 222 of Lecture Notes in Statistics. Springer, Cham. A practical guide with R
- Demarta S, McNeil AJ (2005) The t copula and related copulas. *Int Stat Rev* 73:111–129
- Fang H-B, Fang K-T, Kotz S (2002) The meta-elliptical distributions with given marginals. *J Multivar Anal* 82:1–16
- Favre A-C, Quessy J-F, Toupin M-H (2018) The new family of Fischer copulas to model upper tail dependence and radial asymmetry: properties and application to high-dimensional rainfall data. *Environmetrics* 29(e2494):17
- Genest C, Ghoudi K, Rivest L-P (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82:543–552
- Genest C, Nešlehová J, Ben Ghorbal N (2011) Estimators based on Kendall's tau in multivariate copula models. *Aust N Z J Stat* 53:157–177
- Genest C, Quessy J-F, Rémillard B (2006) Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand J Stat* 33:337–366
- Genest C, Rivest L-P (1993) Statistical inference procedures for bivariate Archimedean copulas. *J Am Stat Assoc* 88:1034–1043
- Gouriéroux C, Monfort A, Renault E (1996) Two-stage generalized moment method with applications to regressions with heteroscedasticity of unknown form. *J Stat Plann Inference* 50:37–63
- Higham NJ (2002) Computing the nearest correlation matrix—a problem from finance. *IMA J Numer Anal* 22:329–343
- Joe H (1990) Multivariate concordance. *J Multivar Anal* 35:12–30
- Joe H (2015) Dependence modeling with copulas. Monographs on statistics and applied probability. CRC Press, Boca Raton
- Kendall MG, Smith BB (1940) On the method of paired comparisons. *Biometrika* 31:324–345
- Kozubowski TJ, Podgórski K, Rychlik I (2013) Multivariate generalized Laplace distribution and related random fields. *J Multivar Anal* 113:59–72
- Krupskii P, Joe H (2013) Factor copula models for multivariate data. *J Multivar Anal* 120:85–101
- Krupskii P, Joe H (2015) Structured factor copula models: theory, inference and computation. *J Multivar Anal* 138:53–73
- Lee AJ (1990) U -statistics statistics: textbooks and monographs. Marcel Dekker Inc, New York
- Mazo G, Girard S, Forbes F (2016) A flexible and tractable class of one-factor copulas. *Stat Comput* 26:965–979
- McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57:995–1026
- McNeil AJ, Nešlehová J (2009) Multivariate Archimedean copulas, d -monotone functions and l_1 -norm symmetric distributions. *Ann Stat* 37:3059–3097
- Nelsen RB (2007) Extremes of nonexchangeability. *Stat Pap* 48:329–336
- Newey WK, McFadden D (1994) Large sample estimation and hypothesis testing. In *Handbook of econometrics*, Vol. IV, vol. 2 of *Handbooks in Econometrics*. North-Holland, Amsterdam, pp 2111–2245
- Oakes D (1994) Multivariate survival distributions. *J Nonparametr Stat* 3:343–354
- Oh DH, Patton AJ (2013) Simulated method of moments estimation for copula-based multivariate models. *J Am Stat Assoc* 108:689–700
- Pakes A, Pollard D (1989) Simulation and the asymptotics of optimization estimators. *Econometrica* 57:1027–1057
- Quessy J-F (2009) Theoretical efficiency comparisons of independence tests based on multivariate versions of Spearman's rho. *Metrika* 70:315–338
- Quessy J-F, Durocher M (2019) The class of copulas arising from squared distributions: properties and inference. *Econom Stat* 12:148–166
- Quessy J-F, Rivest L-P, Toupin M-H (2016) On the family of multivariate chi-square copulas. *J Multivar Anal* 152:40–60
- Shih JH, Louis TA (1995) Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* 51:1384–1399
- Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Stat Univ Paris* 8:229–231
- Tsukahara H (2005) Semiparametric estimation in copula models. *Can J Stat* 33:357–375
- van der Vaart AW (1998) Asymptotic statistics, Cambridge series in statistical and probabilistic mathematics, vol 34. Cambridge University Press, Cambridge

Weiß G (2011) Copula parameter estimation by maximum-likelihood and minimum-distance estimators: a simulation study. *Comput Stat* 26:31–54

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.