ADVANCED REVIEW

WIREs COMPUTATIONAL STATISTICS   WILEY

Copulae: An overview and recent developments

Joshua Größer | **Ostap Okhrin** [ORCID]

Faculty of Transportation, Technische
Universitat Dresden, Dresden, Germany

Correspondence
Ostap Okhrin, Faculty of Transportation,
Technische Universitat Dresden, 01069
Dresden, Germany.
Email: ostap.okhrin@tu-dresden.de

Abstract

Over the decades that have passed since they were introduced, copulae still remain a very powerful tool for modeling and estimating multivariate distributions. This work gives an overview of copula theory and it also summarizes the latest results. This article recalls the basic definition, the most important cases of bivariate copulae, and it then proceeds to a sketch of how multivariate copulae are developed both from bivariate copulae and from scratch. Regarding higher dimensions, the focus is on hierarchical Archimedean, vine, and factor copulae, which are the most often used and most flexible ways to introduce copulae to multivariate distributions. We also provide an overview of how copulae can be used in various fields of data science, including recent results. These fields include but are not limited to time series and machine learning. Finally, we describe estimation and testing methods for copulae in general, their application to the presented copula structures, and we give some specific testing and estimation procedures for those specific copulae.

This article is categorized under:
    Statistical Models > Multivariate Models
    Statistical Models > Semiparametric Models
    Statistical and Graphical Methods of Data Analysis > Multivariate Analysis

KEYWORDS
copula, dependence modeling, multivariate distribution

1 | INTRODUCTION

In the modern era, vast amounts of data have become readily available. This offers the opportunity not only to investigate how single variables behave but also to examine the joint distribution of two or more variables, be it in economics (Kole et al., 2007; Oh & Patton, 2018; Salvatierra & Patton, 2015), medicine (Gomes et al., 2019; Kuss et al., 2014; Lapuyade-Lahorgue et al., 2017), climate research (Oppenheimer et al., 2016; Schölzel & Friederichs, 2008), hydrology and geophysics (Liu et al., 2018; Salvadori et al., 2007; Valle & Kaplan, 2019), engineering (Kilgore & Thompson, 2011), biology (Dokuzoğlu & Purutçuoğlu, 2017; Konigorski et al., 2014), transportation research (Bhat & Eluru, 2009; Huang et al., 2017; Ma et al., 2017), or any other field. However, describing and estimating the properties of these multivariate

In Memory of Abe Sklar.

distributions is complicated, even though the computation power has also increased massively in the last few decades. Copulae are special forms of bivariate and multivariate functions and they are an elegant way to express those multivariate distributions. Sklar (1959) showed that a copula exists for any multivariate distribution, such that the joint distribution equals the copula applied to the marginal distributions. This is very advantageous when estimating the multivariate distribution because the problem now reduces to estimating the univariate distributions, which is much easier, and their dependency structure. It also enables us to take a wider class of distribution functions into account, and model more complicated multivariate distributions more easily and efficiently. Apart from being interesting from a purely mathematical perspective, copulae have a wide range of applications. One of the first practical applications came from economics, where it was used to model economic risk; see Embrechts et al. (2002). The interest in copulae reached a temporary peak among practitioners during the financial crisis in 2008, where they were used to model complex financial derivatives, especially Collateralized Debt Obligations (see Li, 2000). Nowadays, their field of application ranges from financial risk modeling and asset prediction to weather forecasting (Schefzik, 2015) and detection of Martian sand dunes (Carrera et al., 2019). Essentially, copulae can be useful whenever we encounter a joint distribution. Over the last few decades, researchers from the field have proposed a list of text-books that are strongly recommended for everybody interested in copulae—Joe (1997), Cherubini et al. (2004), Nelsen (2006), Kurowicka and Cooke (2006), Mai and Scherer (2012), Joe (2014), Durante and Sempi (2015), Hofert et al. (2018), to name a few—and review articles, such as Patton (2012), Okhrin et al. (2017), Durante and Sempi (2010), and Genest and Nešlehová (2014). Some very useful overviews on the historical developments of copula theory are provided by Schweizer (1991) and Durante and Sempi (2010).

The rest of this article is structured as follows. First bivariate (Section 1) and then multivariate (Section 2) copulae are defined and the most prominent classes of copulae are introduced. Section 3 introduces copulae to other fields of statistics and Section 4 introduces copulae to machine learning. Section 5 takes a closer look at estimation, while Section 6 deals with dependency measures and testing copulae. The final section draws a conclusion.

2 | BIVARIATE COPULAE

In a simple and the most concise version, copulae can be defined as follows.

Definition 1 *A copula $C(u, v)$ is a restriction of a bivariate distribution to the unit square with uniform margins.*

This means in particular that for any rectangle $[a_1, a_2] \times [b_1, b_2]$, the corresponding volume $V([a_1, a_2] \times [b_1, b_2]) = C(a_2, b_2) - C(a_1, b_2) - C(a_2, b_1) + C(a_1, b_1)$ with respect to C is positive and that the respective boundary conditions $C(u, 1) = u$, $C(1, v) = v$, for all $u, v \in [0, 1]$ hold.

The first fundamental result regarding copula theory was stated by Sklar (1959). This lays the foundation for all research in this area as the statement allows the decomposition of any bivariate distribution into the univariate distributions together with the dependence structure.

Theorem 1 (Sklar's Theorem) *For any bivariate distribution function F with margins F_1 and F_2, there exists a copula C such that*

$$F(x, y) = C\{F_1(x), F_2(y)\}, \quad x, y \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}.$$

If F is continuous, then the copula C is unique. Otherwise, it is uniquely determined on $F_1(\overline{\mathbb{R}}) \times F_2(\overline{\mathbb{R}})$.

The converse is true as well: for any copula C and univariate distribution functions F_1 and F_2, the function $C\{F_1(x), F_2(y)\}$ is a bivariate distribution function with margins F_1 and F_2.

Interestingly, Sklar (1959) did not state a formal proof of the theorem now holding his name, although his work is a fundamental breakthrough in statistical thinking; see Durante and Sempi (2015). For many years, most results regarding copulae were contained in Schweizer and Sklar (1983). Since a copula is a distribution function, analogously to the density of any other differentiable distribution function, the density of a copula, if it exists, is given through:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}.$$

As shown in Nelsen (2006), copulae are invariant under strictly monotone transformations of the margins, which implies that they cannot model dependency structures that are not invariant under monotone transformation. To establish a better understanding of what copulae are and what they can look like, we will consider a few simple examples.

Probably the simplest and most straightforward copula is the independence copula $\Pi(u, v) = uv$. This is the copula of two stochastically independent random variables because for those it states that $F(x, y) = F_1(x)F_2(y)$. The other two very basic dependencies between two variables are perfect positive and perfect negative dependencies. These can also be modeled via copulae. For the positive case, this gives $M(u, v) = \min(u, v)$. Thus, for random variables X, Y with distribution functions $F_1(x)$, $F_2(y)$ and the joint distribution $M\{F_1(x), F_2(y)\}$, the random variable X is an increasing function of Y, and vice versa, because the copula is symmetric. The converse, with decreasing instead of increasing, holds true for the joint distribution function $W\{F_1(x), F_2(y)\}$ with $W(u, v) = \max(0, u+v-1)$, which is a copula again. Those two copulae establish boundaries as any other copula has to lie between M and W, which we call the Fréchet–Hoeffding upper and lower bounds, respectively. Because neither M nor W is differentiable in u or v, they do not possess a density.

The following subsection will introduce the most important and most frequently used ones.

2.1 | Extreme-value copulae

Assuming that X_1 and X_2 are two random variables with margins of F_1 and F_2, respectively, then the extreme-value (EV) copulae model the dependence between the maximum of n iid values generated from F_1 and the maximum of n iid values generated by F_2. A copula C is shown to be an EV copula if and only if it is max-stable, meaning

$$C(u,v) = C\left(u^{1/n}, v^{1/n}\right)^n,$$

for all u, v in [0, 1] and $n \in \mathbb{N}$. Any other copula C_F satisfying $C_F(u^{1/n}, v^{1/n}) \to C(u, v)$ as $n \to \infty$ is said to be in the domain of attraction of the EV copula C. This definition can easily be extended to arbitrary dimensions and these copulae are uniquely determined by a tail dependence function that is used when estimating them. EV copulae are often used in financial models and they are also used in hydrology to model flood (c.f. Sraj et al., 2014). More on detailed recent advances on EV copulae can be found in Gudendorf and Segers (2010), Kamnitui et al. (2019), Vettori et al. (2017), Andersen et al. (2018), and Kim et al. (2020).

2.2 | Elliptical family

The elliptical family arises from modeling, for example, the dependence structure underlying a bivariate Gaussian distribution using a copula. The bivariate Gaussian distribution has Gaussian margins. However, as the margins to which a copula is applied are not necessarily Gaussian, by applying the Gaussian copula to any marginal distributions one obtains not only a bivariate normal distribution but also any kind of two dimensional distribution with Gaussian dependency structure between the two variables. This is called *Meta-Gaussian* distribution. This can be directly achieved using Sklar's Theorem, as follows $C(u,v) = F\{F_1^{-1}(u), F_2^{-1}(v)\}$. The procedure works for any bivariate distribution $F(x, y)$ that has invertible marginal distributions F_1 and F_2, and can be extended to multivariate distributions. However, the crux of this approach is that those inverses frequently do not possess explicit representation. Thus, if one can explicitly give the inverse distribution, then this is a straightforward way to model the joint distribution via copulae. Let us have a closer look at the Gaussian copula with the correlation coefficient between the two variables being ρ. The copula is thus given using the inverse distribution of the univariate normal distribution and the distribution function of the bivariate normal distribution, while the density has a closed form representation.

$$C(u,v,\rho) = \Phi_\rho\{\Phi^{-1}(u), \Phi^{-1}(v)\},$$

$$c(u,v,\rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho)}(u^2+v^2-2\rho uv) + \frac{1}{2}(u^2+v^2)\right\},$$

where, Φ is the distribution function of a standard normal distributed random variable, Φ^{-1} its quantile function and Φ_ρ is the distribution function of the bivariate standard normal distribution with correlation coefficient ρ. Similarly, the t copula with ν degrees of freedom, and correlation coefficient ρ and its density are given as

$$C(u,v,\nu,\rho) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left\{1 + \frac{x^2-2\rho xy + y^2}{(1-\rho^2)\nu}\right\}^{\frac{k}{2}-1} dxdy,$$

$$c(u,v,\nu,\rho) = \frac{f_{\nu,\rho}\{t_\nu^{-1}(u), t_\nu^{-1}(v)\}}{f_\nu\{t_\nu^{-1}(u)\}f_\nu\{t_\nu^{-1}(v)\}},$$

where, $f_{\nu,\rho}$ and f_ν are the joint and marginal t-density functions, and t_ν^{-1} is the quantile function of the t-distribution. More details on elliptical and particularly t copulae can be found in Rachev et al. (2008), Demarta and McNeil (2005), and Luo and Shevchenko (2010).

2.3 | Archimedean family

The class of Archimedean copulae (AC) is a rich family of copulae that is not computed using Sklar's Theorem. They depend on the generator function $\phi : [0, \infty) \rightarrow [0, 1]$, with boundary conditions $\phi(0) = 1$, $\phi(\infty) = 0$ and complete monotonicity $(-1)^j\phi^{(j)} \geq 0$, for all $j \in \{1, 2, \dots\infty\}$. This generator can in general depend on several parameters, while most generators that are used to construct copulae depend on a single parameter ϑ. The definition of the Archimedean copula is as follows.

$$C(u,v) = \phi\{\phi^{-1}(u) + \phi^{-1}(v)\}.$$

This implies that AC is symmetric. If the inverse of the generator function does not exist, then the generalized inverse is used instead. An informative list of generators and their corresponding copulae can be found in Nelsen (2006). Here we recall a brief list of those that are most frequently used.

Gumbel (1960) copula

$$\phi(x,\vartheta) = \exp\left(-x^{\frac{1}{\vartheta}}\right), \phi^{-1}(x,\vartheta) = \{-\log(x)\}^\vartheta,$$

Clayton (1978) copula

$$\phi(x,\vartheta) = (\vartheta x + 1)^{-\frac{1}{\vartheta}}, \phi^{-1}(x,\vartheta) = \frac{x^{-\vartheta}-1}{\vartheta},$$

and Frank (1979) copula

$$\phi(x,\vartheta) = -\vartheta^{-1}\log\{1 - (1-e^{-\vartheta})e^{-x}\}, \phi^{-1}(x,\vartheta) = -\log\left\{\frac{\exp(-\vartheta x)-1}{1-\exp(-\vartheta)}\right\}.$$

While the Frank copula is the only elliptically contoured Archimedean copula, the Gumbel copula applied to extreme valued marginal distributions is the only bivariate extreme value distribution that belongs to the Archimedean family, as shown by Genest and Rivest (1989). In addition, any other Archimedean copula is in the domain of attraction of the

Gumbel copula. To get a better intuition of what copulae look like, some contour plots of copulae applied to standard normal or t distributed (with $\nu = 2$) margins are supplied; see Figure 1. A powerful property of AC is that they can be extended to higher dimensions, as we will see in the next section.
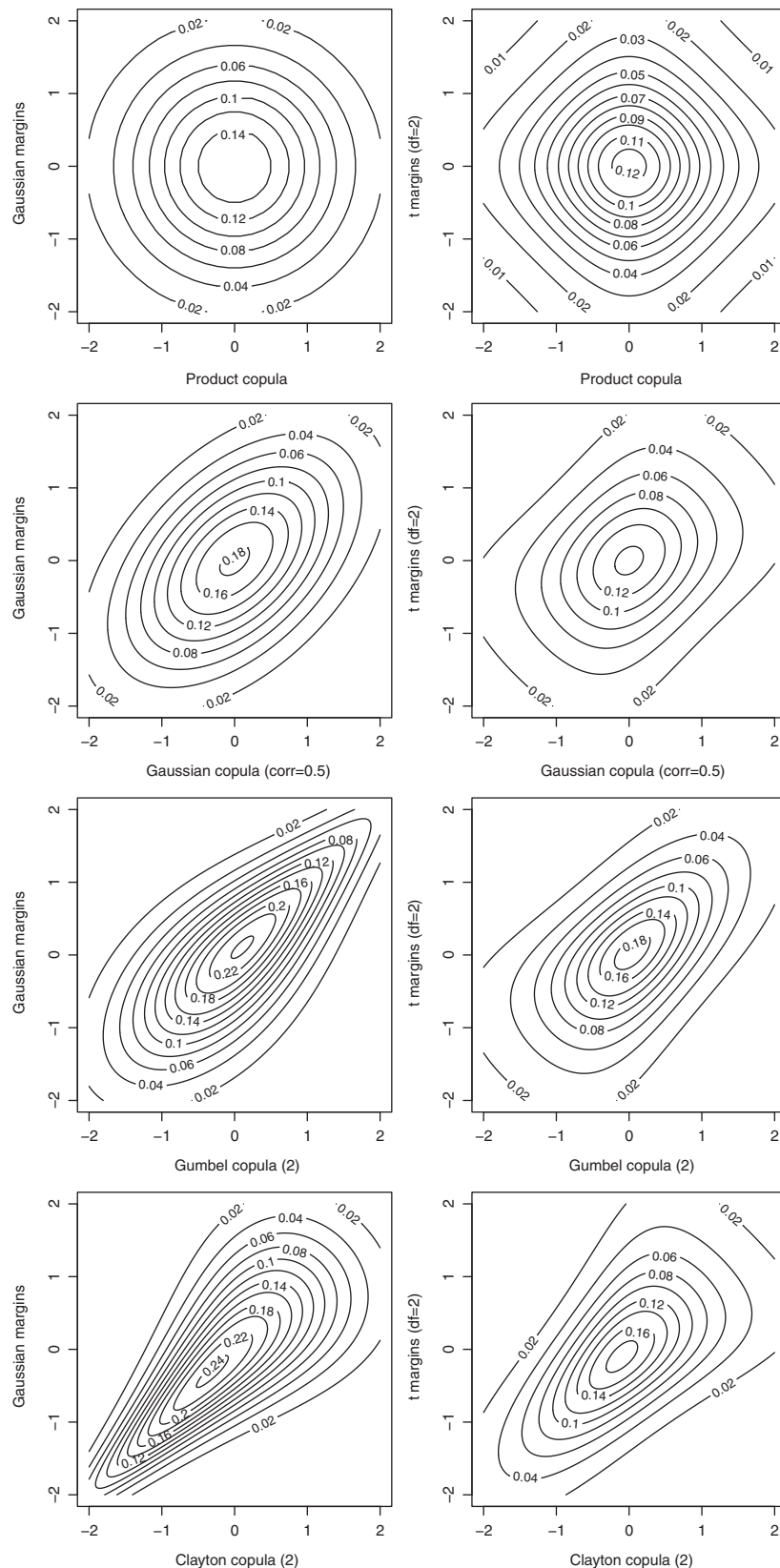

FIGURE 1 Contour plots of different copulae with Gaussian and t-distributed margins

2.4 | Archimax copula

The combination of AC and EV copulae provides Archimax copulae, which is a class of copulae that allows us to model any asymptotic dependence between extremes; which was first introduced in Capéràa et al. (2000). It depends on the Archimedean generator ϕ, which blurs the extreme-value dependence structure, and a convex function A, which is the so-called Pickands dependence function, satisfying $A(t) \geq \max\{t, 1 - t\}$ for $t \in [0, 1]$.

$$C_{\phi,A}(u,v) = \phi\left[\{\phi^{-1}(u) + \phi^{-1}(v)\}A\left\{\frac{\phi^{-1}(u)}{\phi^{-1}(u) + \phi^{-1}(v)}\right\}\right]$$

If $A(t) = 1$, then the copula is just a regular Archimedean copula. Meanwhile, if $\phi(t) = \exp(-t)$, then the copula reduces to an EV copula, and thus the name Archimax. More information on Archimax copulae and their generalizations to higher dimensions can be found in Bacigal et al. (2011), Wysocki (2013), Mesiar and Jágr (2013), and Charpentier et al. (2014). Genest and Jaworski (2020) observe how a prespecification of the degree of association affects the size of the class of Archimax copulae. This is done in general as well as for many of the dependence measures mentioned in Section 6.

2.5 | Generalized Pareto copula

Pareto copula can further help to find and model multivariate extremes. They are defined in the following way: if there is a norm $\|\cdot\|$ such that $C(u, v) = 1 - \|(u, v) - (1, 1)\|$, then C is a Pareto copula. Pareto copulae can be easily extended to higher dimensions by using norms on $\mathbb{R}^d$ instead of $\mathbb{R}^2$. It is worth mentioning here that not every norm induces a copula. It can be shown (see Aulbach et al., 2012) that a copula satisfies the extreme-value condition if and only if it is a generalized Pareto copula (GPC), implying that there is a connection between extreme-value copulae and GPC. The so-called δ-neighborhood of a GPC contains, roughly speaking, the copulae whose maxima have a polynomial rate of convergence of maxima in the sense that

$$\sup_{x \in (-\infty, 0]^d} \left| C^n\left\{1 + \frac{1}{n}x\right\} - \exp(-\|x\|) \right| = O(n^{-\delta}),$$

where, $\|\cdot\|$ is the norm corresponding to the Pareto copula; for further details see Falk et al. (2011). Thus, it is of interest whether a copula is in the neighborhood of a GPC. A test based on χ^2 to verify if this is the case can be found in Aulbach et al. (2019). GPC are of use in extreme value analysis to estimate exceedance probability (see Falk et al., 2019).

2.6 | Marshall–Olkin copulae

When describing the failure time of one specific part of a system, one often assumes it to be exponentially distributed. When a system consists of several components, the failure times are generally assumed to be independent of each other. Then the whole system's failure time, assuming one component's failure makes the system fail, can be computed straightforwardly. However, in reality, this independence assumption is rarely fulfilled. For example, common factors such as maintenance or common shocks can affect the lifetime of all of the components of a system. The Marshall–Olkin distribution takes this into account. This distribution was introduced in Marshall and Olkin (1967) and it relies on independent shocks affecting the lifetime of one or more components. One can then derive copulae describing this kind of dependency. An overview of how to define those copulae in general with their properties, for example, tail dependence coefficients, can be found in Lin and Li (2014) and Botev et al. (2016). When the univariate distributions are all continuous, and there is only one shock that affects all the variables at once, then the construction of the Marshall–Olkin copula is quite simple using a function depending on the shock distribution, as shown in Durante et al. (2016). Durante et al. (2015) provide an idea of how to combine different Marshall–Olkin copulae. AC can also be applied to the study of Marshall–Olkin distributions because the resulting dependence structure can be modeled using the generator of a trivariate Archimedean copula. This limits this approach's applicability because it implies exchangeability of the different survival times, which could instead be tackled using hierarchical AC (see later sections). However, it is still a widely usable approach that is easy to understand and to

implement. Mulinacci (2018) investigated under which conditions the memory-less property of the Marshall–Olkin distribution also holds for the Archimedean based Marshall–Olkin distribution.

3 | MULTIVARIATE COPULAE

In practice, multivariate dependencies are in many cases more relevant than bivariate ones. Consequently, the results obtained for bivariate distributions should be extended to multivariate ones. To that extent, Sklar's Theorem can be extended to its multivariate version.

Theorem 2 *For any multivariate distribution function F with margins $F_1,..., F_d$ there exists a copula C such that*

$$F(x_1,...,x_d) = C\{F_1(x_1),...,F_d(x_d)\}, x_1,...,x_d \in \bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}.$$

If F is continuous, then the copula C is unique. Otherwise, it is only uniquely determined on $F_1(\bar{\mathbb{R}}) \times ... \times F_d(\bar{\mathbb{R}})$. The converse is also true: for any copula C and univariate distribution functions $F_1, ..., F_d$, the function $C\{F_1(x_1), ..., F_d(x_d)\}$ is a multivariate distribution function with margins $F_1, ..., F_d$.

Some multivariate copulae arise practically from the bivariate copulae that were previously established. The independence copula and the upper Fréchet–Hoeffding bound can be extended directly. In contrast, the lower Fréchet–Hoeffding bound is not a copula in dimensions higher than two because the perfect negative dependence in higher dimensions cannot be defined. The family of elliptical copulae can also be extended to the multivariate case. This is done in the same way as constructing bivariate elliptical copulae, but this time one can use the multivariate distribution and the inverses of the marginals of that joint distribution. For AC, the extension to higher dimensions becomes a little trickier.

3.1 | Hierarchical Archimedean copula

With very little effort, one gets a multivariate Archimedean copula by $\phi\{\phi^{-1}(u_1) + ... + \phi^{-1}(u_d)\}$. But this limits their applicability because the distribution depends only on the parameter of the only one Archimedean generator that is used. The variables are also exchangeable because the sum commutes. This is where hierarchical Archimedean copulae (HAC) are useful: to construct a three-dimensional distribution, one replaces one of the margins in the bivariate AC with another bivariate AC. So, with ϕ_2 and ϕ_1 being generators of bivariate AC, $C(u,v,w) = \phi_2\left(\phi_2^{-1}(u) + \phi_2^{-1}\left[\phi_1\{\phi_1^{-1}(v) + \phi_1^{-1}(w)\}\right]\right)$, satisfying $\phi_{d-i}^{-1}\left(\phi_{d-j}\right) : [0, \infty) \rightarrow [0, \infty)$, $\phi_{d-i}^{-1}\left\{\phi_{d-j}(0)\right\} = 0, \phi_{d-i}^{-1}\left\{\phi_{d-j}(\infty)\right\} = \infty$, $(-1)^{j-1}\phi_{d-i}^{-1}\left(\phi_{d-j}\right)^{(j)} \geq 0$ where, d is the dimension of the distribution, $j = 1, ..., d$ and $i < j$. This procedure can be further extended up to any arbitrary dimension by replacing one of the nested margins with another AC. In addition to this fully nested approach it is also possible to construct partially nested copulae. In this case, when extending the above mentioned to higher dimensions, then one of the not yet nested marginals gets replaced with an AC.

The dependency structure partly determines the HAC: ϕ_1 models the dependency between two of the marginals, ϕ_2 models the dependency between the resulting bivariate random variable and another marginal, and so on. This results in a tree-like structure on the marginals, see Figure 2.

Further informations on HAC, their structure, properties, corresponding simulation procedures and some extensions can be found in Hofert and Scherer (2011), Hofert (2011), Górecki et al. (2016), Segers and Uyttendaele (2014), Okhrin et al. (2013a), and Cossette et al. (2019) while a useful R package HAC was thoroughly discussed in Okhrin and Ristig (2014). Recent results by Górecki et al. (2021) show that HAC can be vastly improved by outer power transformation.

3.2 | Multivariate Archimax copulae

As mentioned earlier, Archimax copulae can also be generalized to dimensions higher than two. Therefore, as introduced in Charpentier et al. (2014), one uses the characterization of the EV copula via tail dependence function l, which
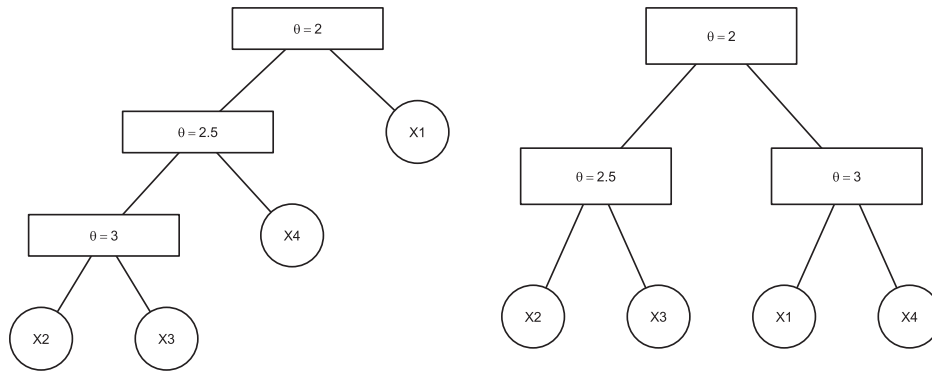
FIGURE 2 The structure of a fully nested hierarchical Gumbel copula with parameters 2, 2.5, and 3 on the left side and the partially nested hierarchical Gumbel copula with parameters 2, 2.5, and 3 on the right side

satisfies that there exists an EV copula D such that $l(x_1,...,x_d) = -\log\{D(e^{-x_1},...,e^{-x_d})\}$ for all x_1, ..., $x_d \in [0, \infty)$. Then, one defines the d-dimensional Archimax copula as

$$C_{\phi,l}(u_1,...,u_d) = \phi\left[l\{\phi^{-1}(u_1),...,\phi^{-1}(u_d)\}\right],$$

where, ϕ is the generator of a multivariate Archimedean copula. Again, if ϕ is the generator of the independence copula, this is an EV copula; while for $l(x_1,...,x_d) = -(e^{-x_1}\cdots e^{-x_d})$, this is a multidimensional Archimedean copula. Therefore, the name Archimax copula still makes sense. Furthermore, for $d = 2$, this returns the definition of the bivariate Archimax copula. Further details on multivariate Archimax copulae can be found, for example, in Chatelain et al. (2020) and Mesiar and Jágr (2013).

3.3 | Factor copulae

Factor copulae are another family of copulae that is more than just an extension from the bivariate case. The idea of factor analysis is to explain observed realizations from the random variables via latent variables, see Härdle and Simar (2015). Factor copulae are based on this approach and arise from the representation of X_i via a functional applied to latent variables and an error term. For example, using the linear additive model, $X_i = \sum_{j=1}^{m} \alpha_{ij} W_j + \varepsilon_i$, where W_j are the latent variables and ε_i are the mutually independent error terms. Replacing linearity by a non-linear function and Gaussian factors by non-Gaussian ones, we can even, among many others, represent classical AC (see Oh & Patton, 2015).

Another more general construction of factor copulae arises from the idea that the dependency between U_1, ..., U_d comes solely from other random variables V_1, ..., V_m, meaning that conditioned on V_j, the U_i are independent. Assuming that all random variables are uniform on [0, 1], this gives us the following copula:

$$C(u_1,...,u_d) = \int_{[0,1]^m} \prod_{i=1}^{d} C_{U_i|V_1,...,V_m}(u_i|v_1,...,v_m)dv_1...dv_m.$$

The product in this formula comes from the fact that the U_i are independent when conditioned on the V_j and the joint distribution function of independent random variables equals the product of those distribution functions. Assuming that $m = 1$, C_{U_i,V_1}, and c_{U_i,V_1} are the joint cdf and density of a pair U_i, V_1, then one can write the density of $C(u_1,...,u_d)$ as

$$c(u_1,...,u_d) = \frac{\partial^d C(u_1,...,u_d)}{\partial u_1,...,\partial u_d} = \int_{[0,1]} \prod_{i=1}^{d} \frac{\partial^2 C_{U_i,V_1}(u_i,v)}{\partial u_i \partial v}\Big|_{v=v_1} dv_1. \tag{1}$$

An overview over factor copula theory can be found in Joe (2014), and for further details we refer the reader to Krupskii and Joe (2013), Nikoloulopoulos and Joe (2015), and Krupskii and Joe (2015).

3.4 | Vine copulae

Vine copulae arise by dividing the dependency structure of d variables into $\frac{1}{2}d(d-1)$ bivariate copulae, not only between two variables but also between two variables, conditioned on one or more others. One typically starts with a graphical representation of the dependence structure. Aiming to decompose the dependency structure into bivariate copulae, one starts with a tree, where the nodes represent the single random variables and the edges represent the copulae between those random variables. The second step is a tree with the former edges as new nodes. The edges represent copulae conditioned on one random variable. This procedure continues until the tree containing only two nodes is left. The idea of vine copulae was introduced in Joe (1996), and they were developed enormously in the 2000s, in particular by Bedford and Cooke (2002), Bedford and Cooke (2001), Kurowicka and Cooke (2006), and Aas et al. (2009).

There are three different main dependency structures: the canonical vine (C-vine), the drawable vine (D-vine) and the regular vine (R-vine). The canonical vine assumes that in each level, one node is connected with every other node in that tree. For example, having random variables $X_1, ..., X_d$, in the first level, there is one copula for the dependence between X_1 and X_2, another for the dependence between X_1 and X_3 and so on, up to the one for the dependence between X_1 and X_d. The copula density corresponding to the graphical representation of a four dimensional C-vine is given through

$$
\begin{aligned}
&c\{F_1(x_1),...,F_4(x_4)\} \\
&= c_{12}\{F_1(c_1),F_2(x_2)\}\cdot c_{13}\{F_1(c_1),F_3(x_3)\}\cdot c_{14}\{F_1(c_1),F_4(x_4)\} \\
&\quad \cdot c_{23;1}\{F(x_2|x_1),F(x_3|x_1)\}\cdot c_{24;1}\{F(x_2|x_1),F(x_4|x_1)\} \\
&\quad \cdot c_{3,4;12}\{F(x_3|x_{12}),F(x_4|x_{12})\}.
\end{aligned}
$$

For the D-vine, at each level, for any node there are at most two edges connecting this node to other nodes, resulting in a line of copulae in the graphical representation. The copula density is thus

$$
\begin{aligned}
&c\{F_1(x_1),...,F_4(x_4)\} \\
&= c_{12}\{F_1(c_1),F_2(x_2)\}\cdot c_{23}\{F_2(c_2),F_3(x_3)\}\cdot c_{34}\{F_3(c_3),F_4(x_4)\} \\
&\quad \cdot c_{13;2}\{F(x_1|x_2),F(x_3|x_2)\}\cdot c_{24;3}\{F(x_2|x_3),F(x_4|x_3)\} \\
&\quad \cdot c_{1,4;23}\{F(x_1|x_{23}),F(x_4|x_{23})\}.
\end{aligned}
$$

This class is useful in modeling the quantile regression, see Kraus and Czado (2017). For an R-vine, it is necessary that in each tree there is an edge between two nodes if and only if those nodes represent edges from the former tree which have a node in common. For example, if there is a copula modeling the dependence between X_1 and X_2, and another one modeling the dependence between X_2 and X_3, then on the next level there is a copula modeling the dependence between X_1 and X_3 conditioned on X_2. C-vines and D-vines are special cases of R-vines. We refer readers who are interested in R-vines and recent advances to Joe (2014), Killiches et al. (2018), and Torre et al. (2019). A good overview of vine copulae can also be found in https://www.groups.ma.tum.de/statistics/forschung/vine-copula-models/.

4 | FURTHER APPLICATIONS OF COPULAE

4.1 | Copulae in time series

Time series analysis is an important field of statistics, with applications from economics through biometrics to meteorology. There are important dependencies to be studied because there are either joint distributions of one random variable at different time points, such as the exchange rate EUR-USD at different days, see Patton (2006), or joint distributions of several random variables at different time points, such as the temperature in different capitals on

different days. Note that all of the statements in this section are made under the assumption that the random variables are continuous.

Sklar's theorem also holds in a conditional version (see Patton, 2004). This means that, given a random vector Y_t and an information set $\mathscr{F}_{t-1}$ being the canonical σ-algebra generated by $Y_{t-1}, Y_{t-2}, \ldots$, the multivariate conditional distribution of Y_t given $\mathscr{F}_{t-1}$ can be expressed via a copula applied to the conditional marginal distributions:

$$F(y_1, \ldots, y_d | \mathscr{F}_{t-1}) = C\{F_1(y_1 | \mathscr{F}_{t-1}), \ldots, F_d(y_d | \mathscr{F}_{t-1})\},$$

where, $Y_{it} | \mathscr{F}_{t-1} \sim F_i(\cdot | \mathscr{F}_{t-1})$. An introduction of copulae to ARCH, GARCH and t-GARCH models can be found in Patton (2004) and (2006). A more general approach is considered in Patton (2012), where the marginals of Y_t are distributed as

$$Y_{it} = \mu_i(Z_{t-1}, b_i) + \sigma_i(Z_{t-1}, b_i)\epsilon_{it},$$

with $Z_{t-1} \in \mathscr{F}_{t-1}$ being the information obtained before time t, b is the vector of parameters, $\epsilon_{it} | \mathscr{F}_{t-1} \sim F_{it}$, so ϵ_{it} is conditionally distributed via the marginal distribution, and $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt}) | \mathscr{F}_{t-1} \sim F_{\epsilon t} = C_t(F_{1t}, \ldots, F_{dt})$, which means that the joint distribution of the ϵ_{it} is captured by a copula applied to the marginals. Both parametric and nonparametric approaches are possible for the distribution F_{it}.

The use of factor copulae in time series applications is discussed in depth in Oh and Patton (2018), while Stöber and Czado (2012) observe how vine copulae can be used to model dependencies over time.

Markov chains can be understood as time series, and thus the corresponding finite-dimensional distributions can be uniquely specified by the distribution function F and the copula C modeling the dependence between two *consecutive* time points. In that respect, the advances of copulae are, similar to the general use of copulae, that heavy-tailed dependency structures can be expressed this way. The conditions under which the Markov chain defined by a copula is stationary with geometric mixing rates can be found in Beare (2010). It was then shown by Beare (2012) that Markov chains generated using AC are ergodic if the generator of the copula is regularly varying at 0 and 1. Furthermore, Beare and Seo (2015) define the M-vine, which is unique on each set of nodes. The Markov property for this M-vine is achieved by requiring certain copulae in the vine copula construction to be the independence copula, while stationarity can be achieved by requiring some copulae to be equal.

4.2 | Clustering methods

No later than 2009, due to the financial crisis, clustering of time series data not only became important for business analysts to optimize portfolios but also it became useful in climate research. Copulae are very useful in clustering, as shown by Luca and Zuccolotto (2017) and Disegna et al. (2017). The clustering methods presented in these articles rely on different procedures. The first one relies on time-varying copula-based estimators via minimization of the value-at-risk instead of other dependency measures. The other clustering algorithm COFUST estimates the copula, it then measures the distance between the copula and the Fréchet upper-bound copula M, which can be further updated using spatial information. Finally, a fuzzy clustering algorithm is applied.

The question of how to cluster data not only refers to time series models. Copulae can also be used to improve existing clustering methods to enable adjusting to more flexible dependencies: many existing finite-mixture approaches rely on the marginals all belonging to the same distribution. This can be improved with Gaussian mixtures, while not increasing the computational effort endlessly for high dimensional data as in Zhang and Baek (2019). Another approach to clustering was taken in Joe and Sang (2016), where the clustering algorithm relies on aggregation variables. When clustering hierarchically, an improvement to clustering methods from Brechmann (2014) can be found in Su et al. (2019). Note that these clustering methods rely on hierarchical Kendall copula with Archimedean clusters.

4.3 | Blind source separation

Statistical methods can also be applied to separate a set of source signals from a set of mixed signals with hardly any knowledge about the mixing process. This is useful when dealing with different signals; for example, when trying to listen to a specific conversation in a crowded room, in image processing, or in medical applications such as EEG or

medical imaging. A special case of blind source separation is independent component analysis (ICA), where one assumes the source signals to be non-Gaussian independent variables that are mixed according to an invertible matrix. Consequently, the remaining task to recover these source signals is to estimate the matrix. Several approaches to do so were recently developed; an overview can be found in Hyvärinen (2012), while an approach using non-local learning methods was provided by Isomura and Toyoizumi (2016). However, copulae can also be used to perform ICA, leading to COPICA, an approach introduced by Chen et al. (2015). Furthermore, copulae can also be used to generalize ICA by replacing the independence assumption with a dependence structure represented by a copula (see Ma & Sun, 2007).

5 | USING COPULAE IN MACHINE LEARNING

Artificial intelligence and machine learning in general are some of the most popular trending topics in both scientific research and popular science. As there is a broad intersection between statistical methods and machine learning, there is an active area of the use of copulae in the improvement of existing machine learning methods or the development of new ones.

Bayesian networks are a kind of expert system and are useful thanks to their ability to enable deductive as well as abductive reasoning. However, when using them in a fully automated way, both the structure and the parameters need to be estimated or learned, as it is called in machine learning literature. One of the first overviews of how copulae can be used to describe properties of Bayesian networks can be found in Elidan (2013). Establishing a Bayesian network neither on expert knowledge nor multivariate normal distributions but vine copulae is done in Pircalabelu et al. (2017). When the observed random variables are not only continuous but also of discrete nature, then the presented models need to be adapted because copulae are only unique if the marginals are continuous. In the case of Bayesian networks, this is solved by Karra and Mili (2016).

Copulae are of further use in machine learning. The so-called Information Bottleneck (IB) is a technique from information theory that is used in machine learning, which was initially used to find a tradeoff between compressing data and obtaining all information. However, in machine learning it is used to find a trade-off between the information that any hidden layer of a deep neuronal network contains about the input and the output. Copula can be used to describe the IB equation, which reformulates the minimization of the IB to find the best fitting copula density for the dependence between input and hidden layer, as shown in Rey (2015). The classical IB methods face some difficulties (see Wieczorek et al., 2018), such as problems with monotone transformations and assumptions on the marginals that might not be true in practice. These can be overcome by transforming the observed random variables to Gaussian ones, which are called *normal scores*.

Fusing data from different sources is a field that also benefits from copulae. When combining the results from different classifiers, copulae are used in combining the respective probability scores (see Ozdemir et al., 2018). Different levels of deep neural networks and their corresponding classifiers are a special case of a general classifier. In Zhang et al. (2019), the authors fused different classifiers when detecting human activity recognition using vine copulae.

6 | ESTIMATION METHODS FOR COPULAE

For notation purposes, note that, while X is usually a random variable, x a realization of this random variable X. Under some regularity conditions, the full maximum likelihood (ML) estimator is efficient and asymptotically normal but it takes a lot of time and computing power, especially for higher dimensions. A computationally easier but not efficient estimator is given by the following procedure called inference for margins. First, the parameters of the margins are estimated using ML under the assumption that the random variables are independent and thus that the copula is the product copula. Then, the parameter of the copula is estimated via ML, using the estimated parameters of the marginals distributions.

To estimate with non-parametric marginals, one can use the rescaled empirical distribution function $\hat{F}_{nj}(x) = \frac{1}{n+1} \sum_{i=1}^{n} I(x_{ij} \leq x)$ for $j = 1, ..., d$ or kernel-based estimators. Then, the estimator of the copula parameter is obtained by maximizing the likelihood of the copula applied to the estimated marginals.

For a complete nonparametric approach to estimating the copula, the rank-based empirical copula

$$C_n(u_1, ..., u_d) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} I \left\{ \frac{1}{n} \sum_{t=1}^{n} I(x_{tj} \leq x_{ij}) \leq u_j \right\} \qquad (2)$$

is very often used, particularly for the testing procedures. However, this is piecewise constant and not continuous. To smooth this out, one can use, for example, the empirical beta copula, which replaces the indicator function with the distribution function of the beta distribution. Further details can be found in Segers et al. (2017), which also provides asymptotics of the empirical beta copula and the empirical Bernstein copula. In contrast, the empirical beta copula is a special case of the empirical Bernstein copula, which relies on Bernstein polynomials; see Sancetta and Satchell (2004). When nonparametrically estimating a bivariate AC using Bernstein polynomials, the Kendall distribution function can be estimated using B-splines or Bezier curve; as in Susam and Ucer (2020). Lin et al. (2017) proposed a procedure for nonparametric estimation of multivariate conditional copulae using local linear smoothing in the kernel estimation and then using the Newton–Raphson method to minimize the estimated likelihood. Another approach to estimating the density of copulae, in both the parametric and nonparametric cases, is based on Legendre multiwavelets and was introduced by Chatrabgoun et al. (2017).

One of the most practical procedures to estimate a copula is via the method of moments. In the bivariate case, estimating the (usually univariate) parameter is done by matching Kendall's τ; as proposed in Genest and Rivest (1993).

6.1 | Estimating hierarchical Archimedean copulae

The problem in estimating HAC comes from two perspectives: first, the bivariate parameters of all of the AC, which can be of the same or different families (e.g., Frank and Clayton), have to be estimated; and second, the dependency structure is not clear. The big focus on research regarding HAC is on structures, where all copulae are of the same family but have different parameters. An estimator for this case was proposed by Okhrin et al. (2013a), using the fact that the dependencies between variables should decrease from the lowest to the highest hierarchical level of the copula structure, meaning that not all different possible structures should be evaluated. This procedure was further generalized by Górecki et al. (2017a) using Kendall's τ for estimating both the structure and the parameters of the copulae simultaneously. Their work includes a proof that the structure of a HAC can be recovered from agglomerative clustering using Kendall's τ. Furthermore, Górecki et al. (2017b) tackle the problem of estimating HAC, which include copulae from different Archimedean families, by establishing an estimator that uses goodness-of-fit test directly in the process of estimation. Under certain regularity conditions, the presented method guarantees a copula. This approach by Górecki et al. (2017b) is based on the estimation of a HAC using only one Archimedean family. The difference is that the bivariate copula is estimated for all admissible families in every step. One then chooses the family which suits best according to a goodness-of-fit test. While presented using the estimation method from Górecki et al. (2017b) for the bivariate copulae, this procedure also works when applied with other estimation methods for the bivariate copulae and the dependency structure. Gunawan et al. (2019) introduce a method to estimate high dimensional AC (not HAC) with discrete margins, which is less computationally complex. This is achieved by using unbiased Bayesian estimators of the likelihood rather than the likelihood itself. A comparison between empirical Bayesian approaches and fully Bayesian approaches, which are computationally more efficient for experiments with mixed outcomes, can be found in Schifano et al. (2020).

6.2 | Estimating Archimax copulae

The problem with Archimax copulae in practice is that estimating them is quite difficult. The first results not limited to the two-dimensional or three-dimensional case, when both the Archimedean generator and the tail dependence function belong to parametric Archimedean families, were obtained by Chatelain et al. (2020). The presented method of moments estimation assumes that estimating the Archimedean generator is a parametric problem. This means that only finitely many Archimedean families are considered, while no further information on the tail dependence function is given. Therefore, the parameter of the Archimedean generator first needs to be estimated. Having estimated the parameter of the Archimedean generator, this can be used to estimate the tail dependence function. Actually, the Pickands dependence function determining the tail dependence function is estimated instead of the tail dependence function

itself. Therefore, with a random sample from a d-dimensional distribution generated from the Archimax copula applied to the margins and $\hat{F}_{nj}$, the empirical distribution function of $x_{1j}, ..., x_{nj}$ and Z is a random variable with survival function ϕ. One then defines $\hat{u}_{ij} = \hat{F}_{nj}(x_{ij})$ and for every ω in the d-dimensional simplex $\hat{\xi}_i(\omega) = \min\{\phi(\hat{u}_{i1})/\omega_1, ..., \phi(\hat{u}_{id})/\omega_d\}$. Based on these, the Pickands (1981) estimator and the Capéràa et al. (1997) type estimator are calculated. Hence these estimators do not necessarily provide a Pickands function, they have to be further constrained; as pointed out in Rojo et al. (2001), who also provided some estimators satisfying the convexity condition and demonstrated the strong uniform convergence. Ahmadabadi and Ucer (2017) introduced a new estimator for the Pickands function based on Bernstein copula and kernel regression. However, its inclusion in the Archimax copula estimation has not yet been studied, while in Kiriliouk et al. (2018) the empirical beta copula was used.

6.3 | Estimating factor copulae

The demanding part of the estimation of factor copulae is that, in the case of non-Gaussian margins and non-Gaussian factors, an analytical form of the density function is in general not available; and therefore ML-estimation is not possible. To cope with this problem, Oh and Patton (2013) introduced a moment-based estimator, which is under weak regularity conditions consistent and asymptotically normal. However, as mentioned earlier, moment estimators for copulae tend to be highly inefficient. Thus, another estimation method, which is based on the copula representation in Equation (1) and the assumption of a parametric copula density, was proposed by Krupskii and Joe (2013). This estimation works similarly to the general approach on copula estimation by first estimating the margins and then using maximum likelihood to estimate the copula parameters.

6.4 | Estimating vine copulae

When estimating vine copulae, similar to HAC, both the copula structure and the resulting density function need to be estimated. The copula structure itself may be estimated using some greedy algorithm, capturing the largest correlations in lower trees, in the spirit of Okhrin et al. (2013b), which suffers from the usual problem that greedy algorithms have. More sophisticated results come from an algorithm proposed in Zhu et al. (2020), which starts with a structure that might be obtained from a greedy or any other algorithm. The algorithm then searches for a regular vine structure that better fits while having at least two common sampling orders with the initial structure. Better results than with the greedy algorithm may be obtained using the procedure proposed in Müller and Czado (2019), which involves structural equation models and the LASSO regularization. In Chang et al. (2019), the tree-like structure of vine copulae is exploited using Monte Carlo tree search to determine the dependency structure. There are several approaches to estimate the density function. The first and most straightforward is the ML estimation, either in one step or two; as discussed in Section 5. However, a semi-parametric approach as introduced in Haff (2013) is more useful. Because the density of a vine copula can be decomposed into a product of conditional bivariate copula densities, the log-likelihood can be decomposed into the sum of logarithms of conditional bivariate densities. By replacing the conditional densities with unconditional ones, one gets a multistep ML procedure by first estimating the parameters in the first level and then using the estimated parameters to estimate the parameters of the copulae in the second level, and so on. This procedure can be further improved by replacing the unconditional densities with conditional densities that only depend on the principal factor and not on all factors, which can be quite a lot in higher dimensions. Consequently, this is a tradeoff between decreased computation time and precision. Details on this procedure and determining the principal factor can be found in Schellhase and Spanhel (2018). Another approach of estimating the conditional densities was developed by Zhang and Beford (2018), where the authors use a mixture of basis functions.

6.5 | Estimating and testing time series copulae

When estimating the multivariate distribution in a copula-based time series, one first estimates all the parameters of the univariate series; such as means, variances, and residuals ϵ_{it}. The latter can be estimated using either parametric or nonparametric approaches. In the case of a nonparametric approach, the distribution is mostly assumed to be constant in time and is then estimated using the empirical distribution or kernel; see Fermanian and Scaillet (2003). When

performing a fully parametric approach, also assuming the copula to be parametric, ML is the most efficient estimator. However, multistage ML is often not less efficient but is much easier to compute; see Patton (2012).

One of the main advantages when using copulae in time series is that one can choose a semiparametric approach: while estimating the marginals using nonparametric methods, the copula itself can then be estimated parametrically using ML. While standard ML procedure cannot be applied as the likelihood depends on infinite-dimensional parameters, Chen and Fan (2006) showed that, under certain conditions, the asymptotic normal distribution of the estimator is achieved.

When testing whether the stochastic behavior of a set of random variables changes over time, most studies deal with the case where the change in the stochastic behavior happens notably at one specific time point rather than a step by step change. Quessy (2019) introduces a test procedure that allows us to test for rather gradual changes, which is what often happens in reality, in the copula as well as in the marginal distributions. Further information on estimating time series models can be found in Härdle et al. (2013) and (2015).

A completely model-free approach using high-frequency data was proposed by Fengler and Okhrin (2016), where the copula is estimated using the realized covariance matrix computed using high-frequency data and later applied to an autoregressive time series model. The obtained estimator is compared with other more classical estimation methods, such as the ML method, with respect to value at risk.

Vine copulae can also be applied to time series problems. This can be done by choosing time-dependent parameters. Using this to model dependencies leads to dynamic vine copula models, where the bivariate copulae, and also the dependence structure, can change over time; as discussed in Kreuzer and Czado (2019), which also includes a Bayesian approach to estimate parameters. The presented class of vine copulae is a generalization of C-vines and D-vines. A nonparametric approach to estimating dynamical vines is presented in Acar et al. (2019).

7 | DEPENDENCE MEASURES AND GOODNESS-OF-FIT TESTS FOR COPULAE

7.1 | Bivariate copulae and dependence measures

Because copulae model the dependence between random variables, there exists a connection between copulae and dependence measures; such as Kendall's τ, Spearman's ρ, and Blomqvist's β. In contrast to Pearson's correlation coefficient, they do not measure the linear dependence but instead measure monotone dependency, and therefore are perfectly suitable for applications on copulae. All τ, ρ, and β are equal to one or minus one in the case of perfect positive or negative dependence, respectively, and are invariant under strictly increasing transformations.

Let (X_1, X_2) and (X_1', X_2') be two independent pairs of random variables, which are both distributed via a bivariate distribution function F, and let C be the corresponding copula. Then, Kendall's τ is defined as follows

$$
\begin{aligned}
\tau &= P\left\{ \left(X_1 - X_1'\right)\left(X_2 - X_2'\right) > 0 \right\} - P\left\{ \left(X_1 - X_1'\right)\left(X_2 - X_2'\right) < 0 \right\} \\
&= 2P\left\{ \left(X_1 - X_1'\right)\left(X_2 - X_2'\right) > 0 \right\} - 1 = 4 \int_{[0,1]^2} C(u,v) dC(u,v) - 1.
\end{aligned}
$$

Using this, for many copulae there is a one-to-one relationship between the parameter and Kendall's τ. For example, Kendall's τ of a Gaussian copula equals $\frac{2}{\pi}\arcsin(\rho)$. As mentioned in Section 5, this can be used to estimate this parameter by estimating Kendall's τ; for example, via $\hat{\tau} = \frac{4}{n(n-1)}P_n - 1$ with P_n being the number of concordant pairs, see Genest et al. (1995).

Similar results can be obtained for Spearman's ρ, assuming that X_1 is distributed via F_1 and X_2 is distributed via F_2 in addition to the assumptions made for Kendall's τ:

$$
\rho = 12 \int_{\mathbb{R}^2} F_1(x_1) F_2(x_2) dF(x_1, x_2) - 3 = 12 \int_{[0,1]^2} uv dC(u,v) - 3.
$$

Schmid and Schmidt (2007) show that ρ can be expressed via

$$\frac{\int_{[0,1]^2} uv dC(u,v) - \int_{[0,1]^2} uv d\Pi(u,v)}{\int_{[0,1]^2} uv dM(u,v) - \int_{[0,1]^2} uv d\Pi(u,v)},$$

thus Spearman's ρ can be interpreted as the normalized average distance between C and the independence copula Π. However, in contrast to Kendall's τ, in case of an Archimedean copula there is rarely an explicit representation with respect to the generator. Both Kendall's τ and Spearman's ρ can be used for a nonparametric estimation of the copula in case of discrete margins, see Zhang et al. (2020) and Blumentritt and Schmid (2014). The exact region that is determined by Spearman's ρ and Kendall's τ is specified in Schreyer et al. (2017).

Blomqvist's β is defined using the median $\tilde{X}_1$ of X_1 and the median $\tilde{X}_2$ of X_2

$$\beta = P\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\} - P\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\}.$$

Compared with Kendall's τ and Spearman's ρ, its estimation needs less computation time as only medians have to be estimated. Using the facts that $P\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\} = 2P\{(X_1 - \tilde{X}_1) < 0, (X_2 - \tilde{X}_2) < 0\} = 2C(\frac{1}{2}, \frac{1}{2})$, $\Pi(\frac{1}{2}, \frac{1}{2}) = \bar{\Pi}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$ and $M(\frac{1}{2}, \frac{1}{2}) = \bar{M}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$, where for any copula C, $\bar{C}(u,v) = P(X_1 > u, X_2 > v)$ is the corresponding survival copula, one can express Blomqvist's β using copula notations in the following way

$$\beta = \frac{C(\frac{1}{2}, \frac{1}{2}) + \bar{C}(\frac{1}{2}, \frac{1}{2}) - \Pi(\frac{1}{2}, \frac{1}{2}) - \bar{\Pi}(\frac{1}{2}, \frac{1}{2})}{M(\frac{1}{2}, \frac{1}{2}) + \bar{M}(\frac{1}{2}, \frac{1}{2}) - \Pi(\frac{1}{2}, \frac{1}{2}) - \bar{\Pi}(\frac{1}{2}, \frac{1}{2})},$$

so Blomqvist's β can also be understood as some normalized distance between the copula C and the independence copula Π similar to Spearman's ρ. This definition can be used to extend β to higher dimensions and to other measures of tail dependence by replacing $\frac{1}{2}$ with other values between 0 and 1. Further details can be found in Schmid and Schmidt (2006), Genest et al. (2013), and Bukovsek et al. (2019).

The Kullback–Leibler divergence is a measure of how much one distribution differs from another. This generally depends on the marginal distributions and the dependency structure. When expressing the KL divergence for a random vector X, whose distribution can be expressed using the C with density c on the marginals of X, one gets that

$$KL(X) = \int_{[0,1]^d} c(u) \log\{c(u)\} du = \mathbb{E}_C[\log\{c(U)\}]$$

where, U is a d-dimensional uniform random variable with independent margins. The KL divergence no longer depends on the marginals. Then, as shown in Blumentritt and Schmid (2012), nonparametric estimators for a copula can be based on the KL divergence.

7.2 | Goodness-of-fit tests

Several goodness-of-fit tests have been developed to test the quality of the estimated copula's fit to the data. Under the H_0 hypothesis, one assumes that the true copula C belongs to some parametric family. A straightforward way is to measure the difference of the estimated copula from the empirical copula Equation (2), which is a consistent estimator of the true copula (see Fermanian et al., 2004; Gaensler & Stute, 1987). Then, tests can be defined using the empirical copula process

$$\mathbb{C}_n(u_1, ..., u_d) = \sqrt{n}\{C_n(u_1, ..., u_d) - C_{\hat{\theta}}(u_1, ..., u_d)\}.$$

The empirical copula process is shown to have smaller variance than the standard empirical process under certain regularity conditions in Genest, Mesfioui, and Nešlehová (2019a). If the marginal distributions are continuous, then the empirical copula process converges weakly to a centered Gaussian process. To extend this to arbitrary margins, a

multilinear extension of the empirical copula—the so-called empirical checkerboard copula, see Deheuvels (1979)—was introduced. As shown by Genest et al. (2017), for arbitrary marginal distributions, the corresponding checkerboard copula process is asymptotically a centered Gaussian process as long as there exists an open subset of $[0, 1]^d$ on which the partial derivatives of the checkerboard copula exist and are continuous.

Some of the tests proposed by Fermanian (2005) and Genest and Rémillard (2008) are based on the Cramer–von-Mises or Kolmogorov distances between the estimated parametric and the empirical copula:

$$S_n^E = \int_{[0,1]^d} \mathbb{C}_n(u_1,...,u_d)^2 dC_n(u_1,...,u_d),$$

$$T_n^E = \sup_{[u_1,...,u_d]\in[0,1]^d} |\mathbb{C}_n(u_1,...,u_d)|.$$

Since these tests all rely on the empirical copula, one can also construct tests using Kendall's distribution instead (see Genest & Rivest, 1993; Genest et al., 2006; Wang & Wells, 2000). Analogously to the empirical copula process, one defines the empirical Kendall's distribution K_n as

$$K_n(v) = \frac{1}{n}\sum_{i=1}^{n} I\left[C_n\{\hat{F}_{n1}(x_{i1}),...,\hat{F}_{nd}(x_{id})\} \leq v\right]$$

while Kendall's distribution is the distribution of the copula value $C(U_1,...,U_d) \sim K$. Thus, Kendall's process is

$$\mathbb{K}_n(v) = \sqrt{n}\{K_n(v) - K_{\hat{\theta}}(v)\}.$$

Subsequently, one can measure the deviation of the estimated Kendall's distribution from the empirical one using Cramer–von Mises or Kolmogorov–Smirnov distances. However, the null hypothesis when testing using Kendall's distribution is only equivalent to the initial H_0 when testing a bivariate Archimedean copula. Moreover, the approach using Kendall's distribution can also be used together with Bernstein polynomials and Bezier curves. Here we have the advantage that the power of the test can be adapted by choosing the degree of the Bernstein polynomial or Bezier curves, as shown in Susam and Ucer (2018) and (2020).

Another integral transformation based approach relies on the Rosenblatt (1952) transformation of a copula C, which is obtained via the mapping $\mathfrak{R}: (0, 1)^d \rightarrow (0, 1)^d$, $(u_1,...,u_d)\mapsto(e_1,...,e_d)$ with $u_1 = e_1$ and $e_i = C_d(u_i| u_1,...,u_{i-1})$, where $C_d(u_i| u_1,...,u_{i-1}) = P(U_i \leq u_i| U_1 = u_1,...,U_{i-1} = u_{i-1})$ with U_j being iid uniform $[0, 1]$ random variables. Then, the initial H_0 is equivalent to e_i being independent of each other. Tests using this transformation rely on Anderson–Darling test statistics (Breymann et al. (2003)) or deviations between estimated density functions, Chen et al. (2004). A substantial improvement regarding the power of those tests can be achieved by using the copula of the transformed data, see Genest et al. (2009). One then uses the empirical distribution function $D_n(u)$ of the e_i and afterward computes the Cramer–von Mises distance

$$S_n = n\int_{[0,1]^d} \{D_n(u_1,...,u_d) - \Pi(u_1,...,u_d)\}du_1...du_d.$$

Further goodness-of-fit tests can be designed using kernel density estimators (see Scaillet, 2007) or the comparison between two-step pseudo maximum likelihood with the delete-one block pseudo maximum likelihood (see Zhang et al., 2016). The asymptotic distributions of most of those tests are not given in explicit form because p-values can be obtained using parametric bootstrap methods (see Okhrin et al., 2017).

Some further interesting testing procedures can be sketched as follows: (a) When looking at the experiment with two binary outcomes, a procedure to maximize the weighted probabilities of the different results (e.g., being healed and not poisoned by a drug) under different copulae models is introduced in Deldossi et al. (2019). Because this maximization may depend on the copula model, it is also important to consider which copula fits the data better and is therefore more likely to be the real copula (and how much more likely); (b) To test whether a copula is symmetric, either in the sense of exchangeability ($C(U, V) = C(V, U)$ in distribution) or radial symmetry ($C(u,v) = \bar{C}(u,v)$), Beare and Seo (2020)

introduced test statistics to check these hypotheses; (c) To test for independence, rank-based test statistics can be used in any dimension and with any marginal distributions; see Genest, Nešlehová, et al. (2019b); (d) As shown by Bianchi et al. (2020), copulae can be used to test for conditional independence, and the presented methods have power comparable to current state-of-the-art methods; (e) Copulae may also be useful when testing several hypothesizes simultaneously. However, the question of whether or not copulae-based tests are in general superior to classical multiple tests such as Bonferroni has yet to be examined. An introduction and a comparison for some Archimedean cases can be found in Neumann and Dickhaus (2020).

8 | CONCLUSION

Copula theory and its applications is an important topic and it has many advantages for practitioners in almost every discipline. Even though many results already exist, there is still a lot of work to be done, especially when estimating higher dimensional dependencies and establishing new copulae.

CONFLICT OF INTEREST
The authors have declared no conflicts of interest for this article.

AUTHOR CONTRIBUTIONS
Joshua Größer: Conceptualization; formal analysis; project administration; supervision; writing-original draft; writing-review & editing. **Ostap Okhrin:** Conceptualization; formal analysis; project administration; supervision; writing-original draft; writing-review & editing.

DATA AVAILABILITY STATEMENT
Data sharing not applicable – no new data generated, or the article describes entirely theoretical research.

ORCID
Ostap Okhrin https://orcid.org/0000-0002-8909-4861

RELATED WIRES ARTICLES
A review of multivariate distributions for count data derived from the Poisson distribution

REFERENCES
Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, *44*(2), 182–198.

Acar, E. F., Czado, C., & Lysy, M. (2019). Flexible dynamic vine copula models for multivariate time series data. *Econometrics and Statistics*, *12*, 181–197.

Ahmadabadi, A., & Ucer, B. H. (2017). Bivariate nonparametric estimation of the Pickands dependence function using Bernstein copula with kernel regression approach. *Computational Statistics*, *32*, 1515–1532.

Andersen, L. N., Laub, P. J., & Rojas-Nandayapa, L. (2018). Efficient simulation for dependent rare events with applications to extremes. *Methodology and Computing in Applied Probability*, *20*, 385–409.

Aulbach, S., Bayer, V., & Falk, M. (2012). Multivariate piecing-together approach with an application to operational loss data. *Bernoulli*, *18*(2), 455–475.

Aulbach, S., Falk, M., & Fuller, T. (2019). Testing for a $_\delta$-neighbourhood of a generalized Pareto copula. *Annals of the Institute of Statistical Mathematics*, *71*, 599–626.

Bacigal, T., Jagr, V., & Mesiar, R. (2011). Non-exchangeable random variables, Archimax copulas and their fitting to real data. *Kybernetika*, *47*(4), 519–531.

Beare, B. K. (2010). Copulas and temporal dependence. *Econometrica*, *78*(1), 395–410.

Beare, B. K. (2012). Archimedean copulas and temporal dependence. *Econometric Theory*, *28*(6), 1165–1185.

Beare, B. K., & Seo, J. (2015). Vine copula specifications for stationary multivariate Markov chains. *Journal of Time Series Analysis*, *36*(2), 228–246.

Beare, B. K., & Seo, J. (2020). Randomization tests of copula symmetry. *Econometric Theory*, *36*(6), 1025–1063.

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, *32*, 245–268.

Bedford, T., & Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, *30*(4), 1031–1068.

Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B: Methodological*, *43*(7), 749–765.

Bianchi, P., Elgui, K., and Portier, F. (2020). Conditional independence testing via weighted partial copulas. Working Paper. https://arxiv.org/pdf/2006.12839.pdf

Blumentritt, T., & Schmid, F. (2012). Mutual information as a measure of multivariate association: Analytical properties and statistical estimation. *Journal of Statistical Computation and Simulation*, *82*(9), 1257–1274.

Blumentritt, T., & Schmid, F. (2014). Nonparametric estimation of copula-based measures of multivariate association from contingency tables. *Journal of Statistical Computation and Simulation*, *84*(4), 781–797.

Botev, Z. I., L'Ecuyer, P., Simard, R., & Tuffin, B. (2016). Static network reliability estimation under the Marshall-Olkin copula. *ACM Transactions on Modeling and Computer Simulation*, *26*(2), 1–28.

Brechmann, E. C. (2014). Hierarchical Kendall copulas: Properties and inference. *The Canadian Journal of Statistics*, *42*(1), 78–108.

Breymann, W., Dias, A., & Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, *3*(1), 1–14.

Bukovsek, D. K., Kosir, T., Mojskerc, B., and Omladic, M. (2019). Relation between Blomqvist's beta and other measures of concordance of copulas. Working Paper. https://arxiv.org/pdf/1911.03467.pdf

Capéràa, P., Fougères, A.-L., & Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, *84*(3), 567–577.

Capéràa, P., Fougères, A.-L., & Genest, C. (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, *72*(1), 30–49.

Carrera, D., Bandeira, L., Santana, R., & Lozano, J. A. (2019). Detection of sand dunes on mars using a regular vine-based classification approach. *Knowledge-Based Systems*, *163*, 858–874.

Chang, B., Pan, S., & Joe, H. (2019). Vine copula structure learning via Monte Carlo tree search. *Proceedings of Machine Learning Research*, *89*, 353–361.

Charpentier, A., Fougères, A.-L., Genest, C., & Nešlehová, J. G. (2014). Multivariate Archimax copulas. *Journal of Multivariate Analysis*, *126*, 118–136.

Chatelain, S., Fougères, A.-L., & Nešlehová, J. G. (2020). Inference for Archimax copulas. *The 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*, 673–690.

Chen, R.-B., Guo, M., Härdle, W., & Huang, S.-F. (2015). COPICA: Independent component analysis via copula techniques. *Statistics and Computing*, *25*, 273–288.

Chen, X., & Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, *135*(1–2), 125–154.

Chen, X., Fan, Y., and Patton, A. (2004). Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. Discussion paper 483, Financial Markets Group, London School of Economics.

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

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*(1), 141–151.

Cossette, H., Marceau, E., Nguyen, Q. H., & Robert, C. Y. (2019). Tail approximation for sums of dependent regularly varying random variables under Archimedean copula models. *Methodology and Computing in Applied Probability*, *21*, 461–490.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. *Bulletin de la Classe des Sciences*, *65*, 274–292.

Deldossi, L., Osmetti, S. A., & Tomassi, C. (2019). Optimal design to discriminate between rival copula models for a bivariate binary response. *TEST*, *28*, 147–165.

Demarta, S., & McNeil, A. J. (2005). The *t* copula and related copulas. *International Statistical Review*, *73*(1), 111–129.

Disegna, M., D'Urso, P., & Durante, F. (2017). Copula-based fuzzy clustering of spatial time series. *Spatial Statistics*, *21*, 209–225.

Dokuzoğlu, D., & Purutçuoğlu, V. (2017). Comprehensive analyses of Gaussian graphical model under different biological networks. *Acta Physica Polonica, A*, *132*(3), 1106–1111.

Durante, F., Girard, S., & Mazo, G. (2015). Copulas based on Marshall-Olkin machinery. In U. Cherubini, F. Durante, & S. Mulinacci (Eds.), *Marshall Olkin distributions: Advances in theory and applications* (pp. 15–31). Springer.

Durante, F., Girard, S., & Mazo, G. (2016). Marshall-Olkin type copulas generated by a global shock. *Journal of Computational and Applied Mathematics*, *296*, 638–648.

Durante, F., & Sempi, C. (2010). Copula theory: An introduction. In P. Jaworski, F. Durante, W. K. Härdle, & T. Rychlik (Eds.), *Copula theory and its applications* (pp. 3–31). Springer.

Durante, F., & Sempi, C. (2015). *Principles of copula theory*. Chapman and Hall/CRC.

Elidan, G. (2013). Copulas in machine learning. In Proceedings of the Workshop on Copulae in Mathematical and Quantitative Finance.

Embrechts, P., McNeil, A. J., & Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster (Ed.), *Risk management* (pp. 176–223). Cambridge University Press.

Falk, M., Hüsler, J., & Reiss, R.-D. (2011). *Laws of small numbers: Extremes and rare events*. Springer.

Falk, M., Padoan, S. A., & Wisheckel, F. (2019). Generalized pareto copulas: A key to multivariate extremes. *Journal of Multivariate Analysis*, *174*, 174–212.

Fengler, M. R., & Okhrin, O. (2016). Managing risk with a realized copula parameter. *Computational Statistics and Data Analysis*, *100*, 131–152.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *Journal of Multivariate Analysis*, *95*(1), 119–152.

Fermanian, J.-D., Radulovic, D., & Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, *10*(5), 847–860.

Fermanian, J.-D., & Scaillet, O. (2003). Nonparametric estimation of copulas for time series. *Risk*, *5*(4), 25–54.

Frank, M. J. (1979). On the simultaneous associativity of f(x,y) and x+y-f(x,y). *Aequationes mathematicae*, *19*, 194–226.

Gaensler, P., & Stute, W. (1987). *Seminar on empirical processes*. Springer Basel AG.

Genest, C., Carabarín-Aguirre, A., & Harvey, F. (2013). Copula parameter estimation using Blomqvist's beta. *Journal de la Société Française de Statistique*, *154*(1), 5–24.

Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, *82*(3), 543–552.

Genest, C., & Jaworski, P. (2020). On the class of bivariate archimax copulas under constraints. *Fuzzy Sets and Systems*.

Genest, C., Mesfioui, M., & Nešlehová, J. G. (2019a). On the asymptotic covariance of the multivariate empirical copula process. *Dependence Modeling*, *7*, 279–291.

Genest, C., & Nešlehová, J. (2014). *Copulas and copula models (statistics reference online)*. Wiley StatsRef.

Genest, C., Nešlehová, J. G., & Rémillard, B. (2017). Asymptotic behavior of the empirical multiliniar copula process under broad conditions. *Journal of Multivariate Analysis*, *159*, 82–110.

Genest, C., Nešlehová, J. G., Rémillard, B., & Murphy, O. A. (2019b). Testing for independence in arbitrary distributions. *Biometrika*, *106*(1), 47–68.

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.

Genest, C., & Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, *44*(6), 1096–1127.

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, *44*(2), 199–213.

Genest, C., & Rivest, L.-P. (1989). A characterization of Gumbel family of extreme value distributions. *Statistics & Probability Letters*, *8*(3), 207–211.

Genest, C., & Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, *88*(423), 1034–1043.

Gomes, M., Radice, R., Camarena Brenes, J., & Marra, G. (2019). Copula selection models for non-Gaussian outcomes that are missing not at random. *Statistics in Medicine*, *38*(3), 480–496.

Górecki, J., Hofert, M., & Holeňa, M. (2016). An approach to structure determination and estimation of hierarchical Archimedean copulas and its application to Bayesian classification. *Journal of Intelligent Information Systems*, *46*, 21–59.

Górecki, J., Hofert, M., & Holeňa, M. (2017a). Kendall's tau and agglomerative clustering. *Dependence Modeling*, *5*, 75–87.

Górecki, J., Hofert, M., & Holeňa, M. (2017b). On structure, family and parameter estimation of hierarchical Archimedean copulas. *Journal of Statistical Computation and Simulation*, *87*(17), 3261–3324.

Górecki, J., Hofert, M., & Okhrin, O. (2021). Outer power transformations of hierarchical Archimedean copulas: Construction, sampling and estimation. *Computational Statistics and Data Analysis*, *155*.

Gudendorf, G., & Segers, J. (2010). Extreme-value copulas. In P. Jaworski, F. Durante, W. K. Härdle, & T. Rychlik (Eds.), *Copula theory and its applications* (pp. 127–145). Springer.

Gumbel, E. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, *55*(292), 698–707.

Gunawan, D., Tran, M.-N., Suzuki, K., Dick, J., & Kohn, R. (2019). Computationally efficient Bayesian estimation of high-dimensional Archimedean copulas. *Statistics and Computing*, *29*, 933–946.

Haff, I. H. (2013). Parameter estimation for pair-copula constructions. *Bernoulli*, *19*(2), 462–491.

Härdle, W. K., Okhrin, O., & Okhrin, Y. (2013). Dynamic structured copula models. *Statistics and Risk Modeling*, *30*, 361–388.

Härdle, W. K., Okhrin, O., & Wang, W. (2015). Hidden Markov structures for dynamic copulae. *Econometric Theory*, *31*(5), 981–1015.

Härdle, W. K., & Simar, L. (2015). *Applied multivariate statistical analysis*. Springer.

Hofert, M. (2011). Efficiently sampling nested Archimedean copulas. *Computational Statistics & Data Analysis*, *55*(1), 57–70.

Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2018). *Elements of copula modeling with R*. Springer Use R! Series.

Hofert, M., & Scherer, M. (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance*, *11*(5), 775–787.

Huang, K., Dai, L., Yao, M., Fan, Y., & Kong, X. (2017). Modelling dependence between traffic noise and traffic flow through an entropy-copula method. *Journal of Environmental Informatics*, *29*(2), 134–151.

Hyvärinen, A. (2012). Independent component analysis: Recent advances. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*, 1–19.

Isomura, T., & Toyoizumi, T. (2016). A local learning rule for independent component analysis. *Scientific Reports*, *6*, 1–17.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf, B. Schweizer, & M. Taylor (Eds.), *Distribution with fixed marginals and related topics, IMS lecture notes – Monograph series*. Institute of Mathematical Statistics.

Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman & Hall.

Joe, H. (2014). *Dependence modeling with copulas*. Chapman and Hall/CRC.

Joe, H., & Sang, P. (2016). Multivariate models for dependent clusters of variables with conditional independence given aggregation variables. *Computational Statistics and Data Analysis*, *97*, 114–132.

Kamnitui, N., Genest, C., Jaworski, P., & Trutschnig, W. (2019). On the size of the class of bivariate extreme-value copulas with a fixed value of Spearman's ρ or Kendall's τ. *Journal of Mathematical Analysis and Applications*, *472*(1), 920–936.

Karra, K. and Mili, L. (2016). Hybrid copula Bayesian networks. In Proceedings of the Eighth International Conference on Probabilistic Graphical Models, PMLR (pp. 240–251).

Kilgore, R. T., & Thompson, D. B. (2011). Estimating joint flow probabilities at stream confluences by using copulas. *Transportation Research Record*, *2262*, 200–206.

Killiches, M., Kraus, D., & Czado, C. (2018). Model distances for vine copulas in high dimensions. *Statistics and Computing*, *28*, 323–341.

Kim, S., Kim, K.-K., & Ryu, H. (2020). Robust quantile estimation under bivariate extreme value models. *Extremes*, *23*, 55–83.

Kiriliouk, A., Segers, J., & Tafakori, L. (2018). An estimator of the stable tail dependence function based on the empirical beta copula. *Extremes*, *21*, 581–600.

Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Banking and Finance*, *31*(8), 2405–2423.

Konigorski, S., Yilmaz, Y. E., & Bull, S. B. (2014). Bivariate genetic association analysis of systolic and diastolic blood pressure by copula models. *BMC Proceedings*, *8*(1), S72.

Kraus, D., & Czado, C. (2017). D-vine copula based quantile regression. *Computational Statistics and Data Analysis*, *110*, 1–18.

Kreuzer, A. and Czado, C. (2019). Bayesian inference for dynamic vine copulas in higher dimensions. https://arxiv.org/abs/1911.00702

Krupskii, P., & Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis*, *120*, 85–101.

Krupskii, P., & Joe, H. (2015). Structured factor copula models: Theory, inference and computation. *Journal of Multivariate Analysis*, *138*, 53–73.

Kurowicka, M., & Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence Modelling*. John Wiley & Sons.

Kuss, O., Hoyer, A., & Solms, A. (2014). Meta-analysis for diagnostic accuracy studies: A new statistical model using beta-binomial distributions and bivariate copulas. *Statistics in Medicine*, *33*(1), 17–30.

Lapuyade-Lahorgue, J., Xue, J.-H., & Ruan, S. (2017). Segmenting multi-source images using hidden Markov fields with copula-based multivariate statistical distributions. *IEEE Transactions on Image Processing*, *26*(7), 3187–3195.

Li, D. X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income*, *9*(4), 43–54.

Lin, J., & Li, X. (2014). Multivariate generalized Marshall-Olkin distributions and copulas. *Methodology and Computing in Applied Probability*, *16*, 53–78.

Lin, J.-G., Zhang, K.-S., & Zhao, Y.-Y. (2017). Nonparametric estimation of multivariate multiparameter conditional copulas. *Journal of the Korean Statistical Society*, *46*(1), 126–136.

Liu, Z., Guo, S., Xiong, L., & Xu, C.-Y. (2018). Hydrological uncertainty processor based on a copula function. *Hydrological Sciences Journal*, *63*(1), 74–86.

Luca, G. D., & Zuccolotto, P. (2017). Dynamic tail dependence clustering of financial time series. *Statistical Papers*, *58*, 641–657.

Luo, X., & Shevchenko, P. V. (2010). The t-copula with multiple parameters of degrees of freedom: Bivariate characteristics and application to risk management. *Quantitative Finance*, *10*(9), 1039–1054.

Ma, J., & Sun, Z. (2007). Copula component analysis. In M. E. Davies, C. J. James, S. A. Abdallah, & M. D. Plumbley (Eds.), *Independent component analysis and signal separation* (pp. 73–80). Springer.

Ma, X., Luan, S., Du, B., & Yu, B. (2017). Spatial copula model for imputing traffic flow data from remote microwave sensors. *Sensors*, *17*(10), 2160.

Mai, J.-F., & Scherer, M. (2012). *Simulating copulas: Stochastic models, sampling algorithms, and applications*. Imperial College Press.

Marshall, A. W., & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, *62*(317), 30–44.

Mesiar, R., & Jágr, V. (2013). *d*-dimensional dependence functions and Archimax copulas. *Fuzzy Sets and Systems*, *228*, 78–87.

Müller, D., & Czado, C. (2019). Selection of sparse vine copulas in high dimensions with the LASSO. *Statistics and Computing*, *29*, 269–287.

Mulinacci, S. (2018). Archimedean-based Marshall-Olkin distributions and related dependence structures. *Methodology and Computing in Applied Probability*, *20*, 205–236.

Nelsen, R. (2006). *An introduction to copulas*. Springer.

Neumann, A., & Dickhaus, T. (2020). Nonparametric Archimedean generator estimation with implications for multiple testing. *AStA Advances in Statistical Analysis*, *104*, 309–323.

Nikoloulopoulos, A., & Joe, H. (2015). Factor copula models for item response data. *Psychometrika*, *80*, 126–150.

Oh, D. H., & Patton, A. (2013). Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association*, *108*(502), 689–700.

Oh, D. H., & Patton, A. (2015). Modelling dependence in high dimensions with factor copulas. In *Finance and economics discussion series 2015–051*. Board of Governors of the Federal Reserve System.

Oh, D. H., & Patton, A. J. (2018). Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads. *Journal of Business and Economic Statistics*, *36*(2), 181–195.

Okhrin, O., Okhrin, Y., & Schmid, W. (2013a). On the structure and estimation of hierarchical Archimedean copulas. *Journal of Econometrics*, *173*(2), 189–204.

Okhrin, O., Okhrin, Y., & Schmid, W. (2013b). Properties of hierarchical Archimedean copulas. *Statistics & Risk Modeling*, *30*, 21–53.

Okhrin, O., & Ristig, A. (2014). Hierarchical Archimedean copulae: The HAC package. *Journal of Statistical Software*, *58*(4), 1–20.

Okhrin, O., Ristig, A., & Xu, Y.-F. (2017). Copulae in high dimensions: An introduction. In W. K. Härdle, C. Y.-H. Chen, & L. Overbeck (Eds.), *Applied quantitative finance*. Springer.

Oppenheimer, M., Little, C. M., & Cooke, R. M. (2016). Expert judgement and uncertainty quantification for climate change. *Nature Climate Change*, *6*, 445–451.

Ozdemir, O., Allen, T. G., Choi, S., Wimalajeewa, T., & Varshney, P. K. (2018). Copula based classifier fusion under statistical dependence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(11), 2740–2748.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, *2*(1), 130–168.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*(2), 527–556.

Patton, A. J. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis*, *110*, 4–18.

Pickands, J. (1981). Multivariate extreme value distributions. In Proceedings of the 43rd session of the International Statistical Institute (pp. 245–253).

Pircalabelu, E., Claeskens, G., & Gijbels, I. (2017). Copula directed acyclic graphs. *Statistics and Computing*, *27*, 55–78.

Quessy, J.-F. (2019). Consistent nonparametric tests for detecting gradual changes in the marginals and the copula of multivariate time series. *Statistical Papers*, *60*, 717–746.

Rachev, S., Stoyanov, S., & Fabozzi, F. (2008). *Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures*. John Wiley & Sons.

Rey, M. (2015). Copula models in machine learning (PhD Thesis). Universität Basel.

Rojo, J., Villa-Diharce, E., & Flores, M. (2001). Nonparametric estimation of the dependence function in bivariate extreme value distributions. *Journal of Multivariate Analysis*, *76*(2), 159–191.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, *23*(3), 470–472.

Salvadori, G., Michele, C. D., Kottegoda, N. T., and Rosso, R. (2007). *Extremes in nature*. Berlin/Heidelberg: Springer.

Salvatierra, I. D. L., & Patton, A. J. (2015). Dynamic copula models and high frequency data. *Journal of Empirical Finance*, *30*, 120–135.

Sancetta, A., & Satchell, S. (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, *20*(3), 535–562.

Scaillet, O. (2007). Kernel-based goodness-of-fit tests for copulas with fixed smoothing parameters. *Journal of Multivariate Analysis*, *98*(3), 533–543.

Schefzik, R. (2015). *Multivariate discrete copulas, with applications in probabilistic weather forecasting*. Heidelberger Institut für Theoretische Studien.

Schellhase, C., & Spanhel, F. (2018). Estimating non-simplified vine copulas using penalized splines. *Statistics and Computing*, *28*, 387–409.

Schifano, E. D., Jeong, H., Deshpande, V., & Dey, E. K. (2020). Fully and empirical Bayes approaches to estimating copula-based models for bivariate mixed outcomes using Hamiltonian Monte Carlo. *TEST*.

Schölzel, C., & Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research: Introduction to the copula approach. *Nonlinear Processes in Geophysics*, *15*(5), 761–772.

Schmid, F., & Schmidt, R. (2006). Nonparametric inference on multivariate versions of Blomqvist's beta and related measures of tail dependence. *Metrika*, *66*, 323–354.

Schmid, F., & Schmidt, R. (2007). Multivariate extensions of Spearman's rho and related statistics. *Statistics and Probability Letters*, *77*(4), 407–416.

Schreyer, M., Paulin, R., & Trutschnig, W. (2017). On the exact region determined by Kendall's τ and Spearman's ρ. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *79*(2), 613–633.

Schweizer, B. (1991). *Thirty years of copulas*. Springer.

Schweizer, B., & Sklar, A. (1983). *Probabilistic metric spaces*. North-Holland Publishing Co.

Segers, J., Sibuya, M., & Tsukahara, H. (2017). The empirical beta copula. *Journal of Multivariate Analysis*, *155*, 35–51.

Segers, J., & Uyttendaele, N. (2014). Nonparametric estimation of the tree structure of a nested Archimedean copula. *Computational Statistics & Data Analysis*, *72*, 190–204.

Sklar, A. (1959). Fonctions de répartition à n dimension et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, *8*, 299–231.

Sraj, M., Bezak, N., & Brilly, M. (2014). Bivariate flood frequency analysis using the copula function: A case study of the Litija station on the Sava river. *Hydrological Processes*, *29*(2), 225–238.

Stöber, J., & Czado, C. (2012). Detecting regime switches in the dependence structure of high dimensional financial data. *arXiv Preprint*.

Su, C.-L., Nešlehová, J. G., & Wang, W. (2019). Modelling hierarchical clustered censored data with the hierarchical Kendall copula. *The Canadian Journal of Statistics*, *47*(2), 182–203.

Susam, S. O., & Ucer, B. H. (2018). Testing independence for Archimedean copula based on Bernstein estimate of Kendall distribution function. *Journal of Statistical Computation and Simulation*, *88*(13), 2589–2599.

Susam, S. O., & Ucer, B. H. (2020). A goodness-of-fit test based on Bezier curve estimation of Kendall distribution. *Journal of Statistical Computation and Simulation*, *90*(7), 1194–1215.

Torre, E., Marelli, S., Embrechts, P., & Sudret, B. (2019). A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas. *Probabilistic Engineering Mechanics*, *55*, 1–16.

Valle, D., & Kaplan, D. (2019). Quantifying the impacts of dams on riverine hydrology under non-stationary conditions using incomplete data and Gaussian copula models. *Science of the Total Environment*, *677*, 599–611.

Vettori, S., Huser, R., & Genton, M. G. (2017). A comparison of dependence function estimators in multivariate extremes. *Statistics and Computing*, *28*, 525–538.

Wang, W., & Wells, M. (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association*, *95*(449), 62–72.

Wieczorek, A., Wieser, M., Murezzan, D., and Roth, V. (2018). Learning sparse latent representations with the deep copula information bottleneck. In: International Conference on Learning Representations.

Wysocki, W. (2013). When a copula is Archimax. *Statistics and Probability Letters*, *83*(1), 37–45.

Zhang, L., & Baek, J. (2019). Mixtures of Gaussian copula factor analyzers for clustering high dimensional data. *Journal of the Korean Statistical Society*, *48*(3), 480–492.

Zhang, L., Lu, D., & Wang, X. (2020). Measuring and testing interdependence among random vectors based on Spearman's ρ and Kendall's τ. *Computational Statistics*, *35*, 1685–1713.

Zhang, M., & Beford, T. (2018). Vine copula approximation: A generic method for coping with conditional dependence. *Statistics and Computing*, *28*, 219–237.

Zhang, S., Geng, B., Varshney, P. K., and Rangaswamy, M. (2019). Fusion of deep neural networks for activity recognition: A regular vine copula based approach. In 2019 22th International Conference on Information FUSION (FUSION) (pp. 1–7).

Zhang, S., Okhrin, O., Zhou, Q. M., & Song, P. X.-K. (2016). Goodness-of-fit test for specification of semiparametric copula dependence models. *Journal of Econometrics*, *193*(1), 215–233.

Zhu, K., Kurowicka, D., & Nane, G. F. (2020). Common sampling orders of regular vines with application to model selection. *Computational Statistics and Data Analysis*, *142*, 106811.