Computational Statistics and Data Analysis 155 (2021) 107109

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Outer power transformations of hierarchical Archimedean copulas: Construction, sampling and estimation

Jan Górecki [a,*], Marius Hofert [b], Ostap Okhrin [c]

[a] Silesian University in Opava, Univerzitní náměstí 1934/3, Karviná, Czech Republic
[b] University of Waterloo, 200 University Avenue West, Waterloo, Canada
[c] Technische Universität Dresden, Würzburger Straße 35, Dresden, Germany

ARTICLE INFO

Article history:
Received 26 March 2020
Received in revised form 24 September 2020
Accepted 25 September 2020
Available online 5 October 2020

Keywords:
Archimedean generator
Outer power transformation
Sampling
Estimation
Tail dependence coefficients
Value at risk

ABSTRACT

Outer power (OP) transformations of Archimedean generators are suggested to increase the modeling flexibility and statistical fitting capabilities of classical Archimedean copulas restricted to a single parameter. For OP-transformed Archimedean copulas, a formula for computing tail dependence coefficients is obtained, as well as two feasible OP Archimedean copula estimators are proposed and their properties studied by simulation. For hierarchical extensions of OP-transformed Archimedean copulas under the sufficient nesting condition, a new construction principle, efficient sampling and parameter estimation for models based on a single one-parameter Archimedean family are addressed. Special attention is paid to the case where the sufficient nesting condition simplifies to two types of restrictions on the corresponding parameters. By simulation, the convergence rate and standard errors of the proposed estimator are studied. Excellent tail fitting capabilities of OP-transformed hierarchical Archimedean copula models are demonstrated in a risk management application. The results show that the OP transformation is able to improve the statistical fit of exchangeable Archimedean copulas, particularly of those that cannot capture upper tail dependence or strong concordance, as well as the statistical fit of hierarchical Archimedean copulas, especially in terms of tail dependence and higher dimensions. Given how comparably simple it is to include OP transformations into existing exchangeable and hierarchical Archimedean copula models, OP transformations provide an attractive trade-off between computational effort and statistical improvement.

1. Introduction

Archimedean copulas (ACs) are dependence models frequently used in finance, insurance and risk management, e.g., for stress testing. In contrast to elliptical copulas such as the prominent Gaussian and t copulas, ACs allow for asymmetry in the joint tails, which is of particular interest, e.g., in risk management (McNeil et al., 2015, Chapter 5) or hydrology (Genest and Favre, 2007; Liu et al., 2018). ACs are also appreciated for their simple analytical form, for efficient sampling techniques and for likelihood-based inference; see, for example, Hofert (2011) and Hofert et al. (2013).

However, as follows from their construction, all multivariate margins of an AC are the same, which limits the applicability of ACs, particularly in high dimensions. Also, a vast majority of known ACs are one-parametric, which, on the one hand, allows for the aforementioned advantages, but, on the other hand, also causes limitations. The single parameter

* Corresponding author.
E-mail addresses: gorecki@opf.slu.cz (J. Górecki), marius.hofert@uwaterloo.ca (M. Hofert), ostap.okhrin@tu-dresden.de (O. Okhrin).

https://doi.org/10.1016/j.csda.2020.107109

determines all properties of an AC, and for many Archimedean families it is related in a one-to-one relationship to the strength of the dependence, e.g., expressed by Kendall's tau; see Table 1 in Genest and Rivest (1993). It is thus natural that this parameter is frequently estimated with a method-of-moments-like estimator such that the implied measure of association is close to its empirical counterpart. However, this often results in a model that fits well in its body, but not so much in its tails.

To alleviate these limitations, several approaches have been introduced in the literature. Our work particularly focuses on the following two:

1. Given a one-parameter family of ACs, a way to construct a two-parameter *outer power AC (OPAC) family* is proposed in Nelsen (2006, Theorem 4.5.1 (note that this reference uses the name *exterior* instead of *outer*)). With the additional parameter, one can fix, e.g., the model's Kendall's tau (τ) to a specific value to keep a good fit in the body while fine-tuning both parameters to get a good fit in one of the tails. Such a property is crucial, e.g., in risk management applications (Hofert et al., 2013; McNeil et al., 2015);
2. Joe (1997, pp. 87) proposed a way to construct *hierarchical* (or *nested*) *ACs* (*HACs*) by nesting several ACs into each other. This allows for different multivariate margins (so an asymmetric model) and extends the one-parameter model to allow for up to $(d-1)$-parameters. However, to this date, all contributions in the literature addressing HACs' estimation have been restricted to the case where all ACs nested in a HAC are one-parametric; see Okhrin et al. (2013a), Górecki et al. (2016, 2017b) to mention a few.

Our work merges these two approaches, resulting in *hierarchical outer power ACs* (*HOPACs*), which are copulas that allow for different multivariate margins with extra flexibility added by the outer power (OP) transformation; further motivation is provided in Appendix A.

This paper is organized as follows. Section 2 recalls basic concepts and notation concerning ACs and the OP transformation. Efficient sampling and estimation strategies for HOPACs are then developed in Section 3, and several HOPAC estimators are studied by means of simulations in Section 4. Excellent abilities of OP transformed AC models in tail dependence modeling are demonstrated in an application from risk management covered in Section 5. Section 6 provides concluding remarks and ideas for future research. In Appendix B.1, the relationship of the OP transformation to three measures of association is recalled. Also, a new result simplifying the computation of the tail dependence coefficients is proposed. An efficient strategy for sampling OPACs is recalled in Appendix B.2 and feasibility of two OPAC estimators is studied by simulation in Appendix B.3.

2. The exchangeable case

2.1. Archimedean copulas

An *Archimedean generator*, or simply *generator*, is a continuous, decreasing function $\psi : [0,\infty) \to [0,1]$ that is strictly decreasing on $[0, \inf\{t : \psi(t) = 0\}]$ and satisfies $\psi(0) = 1$ and $\lim_{t\to\infty} \psi(t) = 0$. If $(-1)^k \psi^{(k)}(t) \geq 0$ for all $k \in \mathbb{N}$, $t \in [0,\infty)$, then ψ is called *completely monotone* (*c.m.*). As follows from Kimberling (1974) or McNeil and Nešlehová (2009), given a c.m. generator ψ, the function $C_\psi : [0,1]^d \to [0,1]$ defined by

$$C_\psi(u_1,\dots,u_d) = \psi\{\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)\} \tag{1}$$

is a d-dimensional *Archimedean* copula (d-AC) for any $d \geq 2$, where ψ^{-1} is the generalized inverse of ψ given by $\psi^{-1}(s) = \inf\{t \in [0,\infty] \mid \psi(t) = s\}$, $s \in [0,1]$. By Bernstein (1929), c.m. generators are Laplace–Stieltjes transforms of distributions on the positive real line, a well-known fact exploited when sampling ACs via their frailties with the algorithm of Marshall and Olkin (1988). In what follows, we assume all appearing generators to be c.m., which allows us to conveniently construct and sample hierarchical versions of ACs considered later, independently of their dimensions.

Table 1 shows the popular Archimedean generators of Ali–Mikhail–Haq (A), Clayton (C), Frank (F), Gumbel (G) and Joe (J), which will serve as working examples throughout the paper. Also note that $\psi_{(a,\theta)}$ will denote the generator of a family a with a real parameter $\theta \in \Theta_a \subseteq [0,\infty)$.

2.2. Outer power transformation

Theorem 4.5.1 in Nelsen (2006) implies that for any $\beta \in [1,\infty)$ and any generator ψ of a 2-AC

$$\psi_\beta(t) = \psi(t^{1/\beta}) \tag{2}$$

generates again a proper 2-AC. Parametric families generated in this way are referred to as *outer power families*, where the unintuitive use of "outer" relates to the fact that they were originally named with reference to generator inverses. Given a one-parameter generator $\psi_{(a,\theta)}$, its OP transform with parameter $\beta \in [1,\infty)$ is denoted by $\psi_{(a,\theta,\beta)}$. A well-known example of an OPAC family is the *generalized Clayton copula* (denoted BB1 in Joe (2014, pp. 190)), which, as mentioned before, encompasses as special cases the three one-parameter families from Nelsen (2006, pp. 116–119) denoted 4.2.1 (Clayton), 4.2.12 and 4.2.14.

Table 1
Five popular families of c.m. one-parameter generators. For each family, the table shows its family label a, parameter range $\Theta_a \subseteq [0,\infty)$, form of $\psi_{(a,\theta)}$, and the lower- and upper-tail dependence coefficients $\lambda_l := \lim_{t\downarrow 0} C_\psi(t,t)/t$ and $\lambda_u := \lim_{t\downarrow 0}\{1-2t+C_\psi(t,t)\}/(1-t)$, where $\beta \in [1,\infty)$ is the OP transform parameter.

a	Θ_a	$\psi_{(a,\theta)}(t)$	λ_l	λ_u
Ali–Mikhail–Haq (A)	$[0,1)$	$(1-\theta)/(\mathrm{e}^t-\theta)$	0	$2-2^{1/\beta}$
Clayton (C)	$(0,\infty)$	$(1+t)^{-1/\theta}$	$2^{-1/(\theta\beta)}$	$2-2^{1/\beta}$
Frank (F)	$(0,\infty)$	$\frac{-\log\{1-(1-\mathrm{e}^{-\theta})\mathrm{e}^{-t}\}}{\theta}$	0	$2-2^{1/\beta}$
Gumbel (G)	$[1,\infty)$	$\mathrm{e}^{-t^{1/\theta}}$	0	$2-2^{1/(\theta\beta)}$
Joe (J)	$[1,\infty)$	$1-(1-\mathrm{e}^{-t})^{1/\theta}$	0	$2-2^{1/(\theta\beta)}$

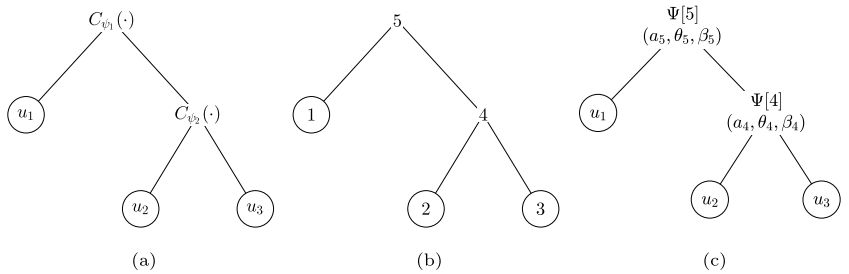


Fig. 1. (a) A tree-like representation of $C_{\psi_1,\psi_2}(u_1,u_2,u_3) = C_{\psi_1}\{u_1, C_{\psi_2}(u_2,u_3)\}$. (b) An undirected tree $(\mathcal{V},\mathcal{E})$, $\mathcal{V} = \{1,\dots,5\}$, $\mathcal{E} = \{\{1,5\},\{2,4\},\{3,4\},\{4,5\}\}$ derived for the tree in Fig. 1a. (c) Our representation of $C_{(\mathcal{V},\mathcal{E},\Psi)}(u_1,u_2,u_3) = C_{\Psi[5]}\{u_1, C_{\Psi[4]}(u_2,u_3)\}$, where $\Psi[4] = \psi_{(a_4,\theta_4,\beta_4)}$, and $\Psi[5] = \psi_{(a_5,\theta_5,\beta_5)}$ and $(\mathcal{V},\mathcal{E})$ is given by Fig. 1b.

Note that the Gumbel family has generator $\psi_{(\mathrm{G},\theta)} = \mathrm{e}^{-t^{1/\theta}}$, so OP Gumbel copulas are simply Gumbel copulas with parameter $\theta\beta$ instead of θ since

$$\psi_{(\mathrm{G},\theta,\beta)}(t) = \psi_{(\mathrm{G},\theta)}(t^{1/\beta}) = \mathrm{e}^{-(t^{1/\beta})^{1/\theta}} = \mathrm{e}^{-t^{1/(\theta\beta)}} = \psi_{(\mathrm{G},\theta\beta)},$$

where $\theta \in [1,\infty)$. For this reason, this family is not further considered for studying the properties of the OP transformation.

Appendix B.1 summarizes the impact of the OP transformation on three measures of association: Kendall's tau, the lower- and the upper-tail dependence coefficient. This includes a proposition for computing the tail dependence coefficients under regular variation. An efficient strategy for sampling OPACs is recalled in Appendix B.2 and feasibility of two OPAC estimators is studied via simulations in Appendix B.3.

3. The nested case

3.1. Hierarchical Archimedean copulas

To construct a *hierarchical Archimedean copula* (*HAC*), one replaces some arguments of an AC by other (H)ACs, see Joe (1997, pp. 87). One also needs to verify that a proper copula results from such a construction, typically by verifying a sufficient nesting condition. For example, given two 2-ACs C_{ψ_1} and C_{ψ_2}, a 3-variate HAC, denoted by C_{ψ_1,ψ_2}, can be constructed via

$$C_{\psi_1,\psi_2}(u_1,u_2,u_3) = C_{\psi_1}\{u_1, C_{\psi_2}(u_2,u_3)\}; \tag{3}$$

see Fig. 1a for a tree representation of such a construction.

Using the language of graph theory, an *undirected tree* $(\mathcal{V},\mathcal{E})$ related to this representation can be derived by enumerating all of its nodes. Here, $\mathcal{V}$ is a set of nodes $\{1,\dots,m\}$, $m \in \mathbb{N}$, and $\mathcal{E} \subset \mathcal{V}\times\mathcal{V}$. For example, in Fig. 1b, $\mathcal{V} = \{1,\dots,5\}$ and $\mathcal{E} = \{\{1,5\},\{2,4\},\{3,4\},\{4,5\}\}$. The *leaves* $\{1,2,3\}$ correspond to the HAC variables u_1,u_2 and u_3, whereas the non-leaf nodes $\{4,5\}$, called *forks*, correspond to the ACs (uniquely determined by the corresponding generators) nested in C_{ψ_1,ψ_2}. When deriving a particular (undirected) tree for the tree representation in Fig. 1a, we assume that

1. the leaves 1, 2 and 3 in Fig. 1b correspond to u_1,u_2 and u_3 in Fig. 1a, respectively, and that
2. the fork indices 4 and 5 are set arbitrarily (we could have also derived an undirected tree with fork indices 4 and 5 switched). In order to enumerate the forks uniquely, we set them according to the corresponding Kendall's tau, meaning that the root, which has always the lowest Kendall's tau, is numbered by m, the node with the second lowest Kendall's tau by $m-1$, etc.

As each fork corresponds to a generator, we represent this relationship using a labeling function Ψ, which maps the forks to the corresponding generators. In our example,

$$\Psi[4] = \psi_2 \text{ and } \Psi[5] = \psi_1. \tag{4}$$

Using this notation, (3) can be rewritten as

$$C_{\Psi[5]}\{u_1, C_{\Psi[4]}(u_2, u_3)\}. \tag{5}$$

Observe that the indices of the arguments of the inner copula $C_{\Psi[4]}$ correspond to the set of children of fork 4, i.e., to {2, 3}, and the indices of the arguments of the copula $C_{\Psi[5]}$ correspond to the set of children of fork 5, i.e., to {1, 4} where the number 4 represents the copula $C_{\Psi[4]}(u_2, u_3)$. This implies that one can express $C_{\psi_1,\psi_2}(u_1, u_2, u_3)$ in terms of the triplet $(\mathcal{V}, \mathcal{E}, \Psi)$. Following this observation, we denote a HAC in an arbitrary dimension by $C_{(\mathcal{V},\mathcal{E},\Psi)}$; see also Definition 3.1 in Górecki et al. (2017b). Finally, let $\Psi[i] = \psi_{(a_i,\theta_i,\beta_i)}$, $i \in \{4, 5\}$ be two OPAC generators. The graphical representation depicted in Fig. 1c fully determines the parametric HAC $C_{(\mathcal{V},\mathcal{E},\Psi)} = C_{\psi_1,\psi_2}$ given by (3) and (4), i.e., its structure, the families of its generators and its parameters.

To guarantee that a proper copula results from nesting ACs, we will use the *sufficient nesting condition* (SNC) proposed by Joe (1997, pp. 87) and McNeil (2008). It states that if for all parent–child pairs of forks (i, j) appearing in a nested construction $C_{(\mathcal{V},\mathcal{E},\Psi)}$ the first derivative of the composition $\Psi[i]^{-1} \circ \Psi[j]$ is c.m., then $C_{(\mathcal{V},\mathcal{E},\Psi)}$ is a copula. This SNC has three important practical advantages, which are that (1) its expression in terms of the corresponding parameters is known in most cases, (2) this expression does not depend on the copula dimension d for all pairs for which it is known, and, most importantly, (3) efficient sampling strategies based on a stochastic representation for HACs satisfying the SNC are known; see, for example, Hofert (2012). Note that there also exists a weaker sufficient condition, see Rezapour (2015), which however lacks those three advantages and is thus of limited practical use.

3.2. Constructing and sampling HOPACs

Starting with a simple trivariate nested copula structure, e.g., the one depicted in Fig. 1, we can sample from it according to the algorithm proposed by McNeil (2008). Note that one can easily apply the same strategy to the general d-variate HAC (d-HAC) case with $d \geq 3$. This sampling algorithm, extends the one of Marshall and Olkin (1988) for ACs, and, here, requires to sample the two following random variables:

1. $V_1 \sim \mathcal{LS}^{-1}[\mathring{\psi}_1]$ and
2. $V_{12} \sim \mathcal{LS}^{-1}[\exp(-V_1 \mathring{\psi}_1^{-1} \circ \mathring{\psi}_2)]$,

where the rings emphasize that the generators are OP transformed, say $\mathring{\psi}_1 = \psi_{(a_1,\theta_1,\beta_1)}$ and $\mathring{\psi}_2 = \psi_{(a_2,\theta_2,\beta_2)}$, where a_1 and a_2 are labels of c.m. families of one-parameter generators, $\theta_1 \in \Theta_{a_1}$, $\theta_2 \in \Theta_{a_2}$ and $\beta_1, \beta_2 \geq 1$.

A strategy for sampling from V_1 has been recalled in Appendix B.2. It is important to note that $V_1 > 0$ with probability 1 as a result of Bernstein's Theorem (Bernstein, 1929) and the fact that $\mathring{\psi}_1$ is a c.m. generator.

To sample from V_{12} with Laplace–Stieltjes transform $\mathring{\psi}_{12}$, consider that for all $t \in [0, \infty)$,

$$\begin{aligned}
\mathring{\psi}_{12}(t; V_1) &:= \exp[-V_1 \mathring{\psi}_1^{-1}\{\mathring{\psi}_2(t)\}] \\
&= \exp[-V_1 \psi^{-1}_{(a_1,\theta_1,\beta_1)}\{\psi_{(a_2,\theta_2,\beta_2)}(t)\}] \\
&= \exp[-V_1 (\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}(t^{\frac{1}{\beta_2}})\})^{\beta_1}] \qquad (6) \\
&= \exp\big[-V_1 \big\{-\log\big(\exp[-\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}(t^{\frac{1}{\beta_2}})\}]^{\beta_1}\big)\big\}\big] \\
&= \bar{\psi}_1[-\log\{\bar{\psi}_{12}(t^{\frac{1}{\beta_2}})\}; V_1], \qquad (7)
\end{aligned}$$

where $\bar{\psi}_1(t; V_1) := \exp(-V_1 t^{\beta_1})$ and $\bar{\psi}_{12}(t) := \exp[-\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}(t)\}]$. Note that V_1 acts as a parameter for $\mathring{\psi}_{12}$. The following proposition provides an explicit way to sample from V_{12}.

Proposition 1. *Let $\beta_1 = 1$ and $[\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}\}]'$ be c.m. Then $\mathring{\psi}_{12}(t; V_1) = \exp[-V_1 \psi^{-1}_{(a_1,\theta_1,\beta_1)}\{\psi_{(a_2,\theta_2,\beta_2)}(t)\}]$ is c.m. for all $\beta_2 \in [1, \infty)$, $t \in [0, \infty)$ and $V_1 \in (0, \infty)$, and*

$$V_{12} := S \breve{V}_{12}^{\beta_2} \sim \mathcal{LS}^{-1}[\mathring{\psi}_{12}],$$

where $S \sim S(1/\beta_2, 1, \cos^{\beta_2}(\frac{\pi}{2\beta_2}), \mathbb{1}_{\{\beta_2=1\}}; 1)$ and $\breve{V}_{12} \sim \mathcal{LS}^{-1}[(\bar{\psi}_{12})^{V_1}]$.

Proof. $\beta_1 = 1$ implies that $\bar{\psi}_1$ is c.m. As $-\log\{\bar{\psi}_{12}(t)\} = \psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}(t)\}$, the assumptions and Proposition 2.1.5 (5) from Hofert (2010) imply that $(\bar{\psi}_{12})^{V_1}$ is c.m. for all $V_1 > 0$. For the special case $V_1 = 1$, Theorem 1 with $\beta = \beta_2$ further implies that also $\bar{\psi}_{12}(t^{\frac{1}{\beta_2}})$ is c.m. With Proposition 2.1.5 (5) from Hofert (2010), the first derivative of $\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}(t^{\frac{1}{\beta_2}})\}$

is also c.m. Finally, as $\bar{\psi}_1$ and $[-\log\{\bar{\psi}_{12}(t^{\frac{1}{\beta_2}})\}]'$ are c.m., Proposition 2.1.5 (2) from Hofert (2010) implies that also $\mathring{\psi}_{12}$ is c.m.

Further, starting from (6) and with $\beta_1 = 1$, $\mathring{\psi}_{12}(t; V_1)$ can be rewritten as

$$\mathring{\psi}_{12}(t; V_1) = \bar{\psi}_{12}(t^{\frac{1}{\beta_2}})^{V_1}. \tag{8}$$

Denoting by $\breve{V}_{12}$ a random variable distributed according to $\mathcal{LS}^{-1}[(\bar{\psi}_{12})^{V_1}]$ and applying Theorem 1 with $\psi(t) = \bar{\psi}_{12}(t)^{V_1}$ and $\beta = \beta_2$ based on (8) establishes the proof. □

Proposition 1 implies that:

1. *Any* OP family ($\beta_2 \geq 1$) can be nested into a non-OP family ($\beta_1 = 1$) just under the *one-parameter* SNC, i.e., if $[\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}\}]'$ is c.m., which simplifies to $\theta_1 \leq \theta_2$ for many families when $a_1 = a_2$; see the second column of Table 2.3 in Hofert (2010). To the best of our knowledge, this nesting case has not been considered in the literature.
2. Know-how of sampling $\mathcal{LS}^{-1}[\exp(-V_1\mathring{\psi}_1^{-1} \circ \mathring{\psi}_2)]$, under $\beta_1 = 1$, can be fully translated to know-how of sampling its non-OP transformed version, i.e., $\mathcal{LS}^{-1}[(\bar{\psi}_{12})^{V_1}]$, which is known for many families; see the third column of Table 2.3 in Hofert (2010). Also note that free implementations of sampling algorithms for $\mathcal{LS}^{-1}[(\bar{\psi}_{12})^{V_1}]$ are available, e.g., in the R package **copula** (Hofert et al., 2017) or in the **HACopula** toolbox for MATLAB and Octave (Górecki et al., 2020).

For $\beta_1 > 1$, $\bar{\psi}_1$ is not a c.m. generator and thus Proposition 1 does not provide a nesting condition. However, if $a_1 = a_2$ and $\theta_1 = \theta_2$, then

$$\mathring{\psi}_{12}(t; V_1) = \exp\big(-V_1[\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_1,\theta_1)}(t^{\frac{1}{\beta_2}})\}]^{\beta_1}\big) = \exp(-V_1 t^{\frac{\beta_1}{\beta_2}}),$$

which is a Gumbel generator (with inverse Laplace–Stieltjes transform $\mathrm{S}(\beta_1/\beta_2, 1, \{\cos(\frac{\beta_1}{\beta_2}\frac{\pi}{2})V_1\}^{\frac{\beta_2}{\beta_1}}, V_1\mathbb{1}_{\{\beta_1/\beta_2=1\}}; 1)$) provided that $\beta_1 \leq \beta_2$; see Hofert (2011).

Hence, we currently know of the following two scenarios in which the SNC holds and thus a proper HOPAC of the form $C_{\dot{\psi}_1}\{u_1, C_{\dot{\psi}_2}(u_2, u_3)\}$ results:

1. if $\beta_1 = 1$, then $[\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}\}]'$ is c.m and $\beta_2 \geq \beta_1$, or
2. if $\beta_1 > 1$, then $a_1 = a_2$, $\theta_1 = \theta_2$ and $\beta_2 \geq \beta_1$.

As mentioned above, these constraints can easily be translated to the general nesting case just by validating them for every parent–child pair of nodes in the nested copula structure.

Let us now consider the case when $a_1 = a_2$ and the condition on $[\psi^{-1}_{(a_1,\theta_1)}\{\psi_{(a_2,\theta_2)}\}]'$ to be c.m. simplifies to $\theta_1 \leq \theta_2$. Construction of HOPACs under the SNC is then similar to one-parameter HACs, where the parameters θ_i involved need to be increasing as one goes further down in a branch of the copula structure, as is shown in the following proposition.

Proposition 2. *Let a be a one-parameter c.m. Archimedean family such that the condition for $[\psi^{-1}_{(a,\theta_1)}\{\psi_{(a,\theta_2)}\}]'$ being c.m. holds if and only if $\theta_1 \leq \theta_2$ for all $\theta_1, \theta_2 \in \Theta_a$. Also, let $(\mathcal{V}, \mathcal{E})$ be an undirected tree with d leaves and f forks and $\Psi[i] = \psi_{(a,\theta_i,\beta_i)}$, $\theta_i \in \Theta_a$, $\beta_i \geq 1$ for $i \in \{d+1, \ldots, d+f\}$. If for each pair of forks (parent, child) in $(\mathcal{V}, \mathcal{E})$ it either holds that*

$$\begin{aligned} &(\mathcal{R}_1)\ \text{if } \beta_{parent} = 1,\ \text{then } \theta_{child} \geq \theta_{parent} \text{ and } \beta_{child} \geq \beta_{parent},\ \text{or} \\ &(\mathcal{R}_2)\ \text{if } \beta_{parent} > 1,\ \text{then } \theta_{child} = \theta_{parent} \text{ and } \beta_{child} \geq \beta_{parent}, \end{aligned} \tag{9}$$

then $C_{(\mathcal{V},\mathcal{E},\Psi)}$ is a HOPAC.

Proof. As follows from McNeil (2008), if $\{\Psi[parent]^{-1}(\Psi[child])\}'$ is c.m. for each pair of forks (*parent*, *child*) in $(\mathcal{V}, \mathcal{E})$, then $C_{(\mathcal{V},\mathcal{E},\Psi)}$ is a copula. As generators $\Psi[i]$, $i \in \{d+1, \ldots, d+f\}$ are OP generators, if $C_{(\mathcal{V},\mathcal{E},\Psi)}$ is a copula, then $C_{(\mathcal{V},\mathcal{E},\Psi)}$ is a HOPAC. We thus show that, given an arbitrary pair of forks (*parent*, *child*) in $(\mathcal{V}, \mathcal{E})$, if either $\mathcal{R}_1$ or $\mathcal{R}_2$ holds, then $[\psi^{-1}_{(a,\theta_{parent},\beta_{parent})}\{\psi_{(a,\theta_{child},\beta_{child})}\}]'$ is c.m. As the proof for $\mathcal{R}_1$ uses the argument of the proof for $\mathcal{R}_2$, the ordering is reversed.

($\mathcal{R}_2$) Let $\beta_{parent} > 1$ and $\theta_{child} = \theta_{parent}$ and $\beta_{child} \geq \beta_{parent}$. As $\psi^{-1}_{(a,\theta_{parent},\beta_{parent})}\{\psi_{(a,\theta_{child},\beta_{child})}(t)\} = \psi^{-1}_{(a,\theta_{parent})}\{\psi_{(a,\theta_{child})}(t^{\frac{1}{\beta_{child}}})\}^{\beta_{parent}} = t^{\frac{\beta_{parent}}{\beta_{child}}}$ and $\frac{\beta_{parent}}{\beta_{child}} \in (0, 1]$, considering that $(t^s)' = st^{s-1}$ is c.m. for all $s \in (0, 1)$, the proof is established.

($\mathcal{R}_1$) Let $\beta_{parent} = 1$. Then, under $\mathcal{R}_1$, two cases can be distinguished. (1) $\theta_{child} = \theta_{parent}$ and $\beta_{child} \geq \beta_{parent}$. Considering that $\frac{\beta_{parent}}{\beta_{child}} = \frac{1}{\beta_{child}} \in (0, 1]$, the same argument as for $\mathcal{R}_2$ can be used. (2) $\theta_{child} \geq \theta_{parent}$ and $\beta_{child} \geq \beta_{parent}$. Let $V_{parent} \sim \mathcal{LS}^{-1}[\psi_{(a,\theta_{parent})}]$ and consider that $V_{parent} > 0$ probability 1; see Bernstein's Theorem (Bernstein, 1929) and use the fact that $\psi_{(a,\theta_{parent})}$ is a c.m. generator. Applying Proposition 1 with $a_1 = a_2 = a$, $\beta_1 = 1$, $\beta_2 = \beta_{child}$,

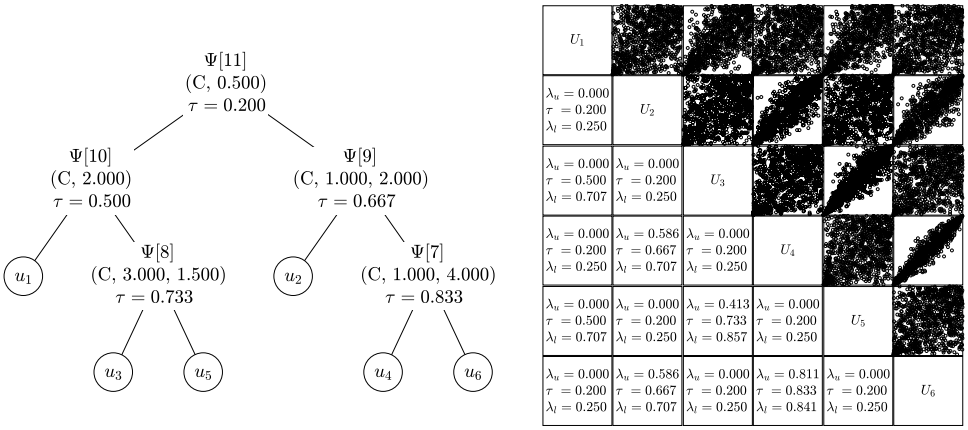


Fig. 2. (left) A 6-variate Clayton HOPAC satisfying the SNC. If $\beta = 1$ then the value of β is omitted (forks 10 and 11). (right) A pairwise plot of 1000 observations from this HOPAC, including the pairwise dependence measures λ_u, τ and λ_l computed from the HOPAC.

$\theta_1 = \theta_{parent}$ and $\theta_2 = \theta_{child}$ yields that $\exp[-V_{parent}\psi^{-1}_{(a,\theta_{parent})}\{\psi_{(a,\theta_{child},\beta_{child})}(t)\}]$ is c.m. for all $\beta_{child} \geq 1$, $t \in [0,\infty)$ and $V_{parent} \in (0,\infty)$. Using 2.1.5 (5) from Hofert (2010) implies that $\{-\log(\exp[\psi^{-1}_{(a,\theta_{parent})}\{\psi_{(a,\theta_{child},\beta_{child})}(t)\}])\}' = [\psi^{-1}_{(a,\theta_{parent})}\{\psi_{(a,\theta_{child},\beta_{child})}(t)\}]'$ is c.m. for all $\beta_{child} \geq 1$ and $t \in [0,\infty)$. □

Hence, in a HOPAC structure, the parameters also have to increase but one can choose which one of them. Let us illustrate this with the model depicted on the left-hand side of Fig. 2. One can observe that θ or β increases as one goes down in a branch of the copula structure. More precisely, if $\beta_{parent} = 1$, then $\theta_{child} \geq \theta_{parent}$ and of course $\beta_{child} \geq \beta_{parent}$, see, e.g., the pair of forks (10, 8) or (11, 9). This new flexible case is enabled by Proposition 1. Or, if $\beta_{parent} > 1$ then θ_{child} is fixed to θ_{parent}'s value and only β increases as one goes down in a branch, as the pair of forks (9, 7) indicates. This case has already been used in Hofert and Scherer (2011). For completeness, let us mention that the pair of forks (11, 10) represents nesting of two one-parameter ACs.

Remark 1. Before we move on to estimation, let us briefly remark that Archimax copulas with (hierarchical) logistic stable tail dependence functions are indeed also (H)OPACs.

Archimax copulas, see Capéraà et al. (2000) and Charpentier et al. (2014), are copulas of the form

$$C(\boldsymbol{u}) = \psi[\ell\{\psi^{-1}(u_1),\dots,\psi^{-1}(u_d)\}], \quad \boldsymbol{u} \in [0,1]^d, \tag{10}$$

where ψ is an Archimedean generator and $\ell : [0,\infty)^d \to [0,\infty)$ is a stable tail dependence function; see Ressel (2013) and Charpentier et al. (2014) for a characterization of stable tail dependence functions. *Logistic stable tail dependence functions* are given by $\ell(\boldsymbol{x}) = \ell_\beta(\boldsymbol{x}) = \left(\sum_{j=1}^d x_j^\beta\right)^{1/\beta}$, $\boldsymbol{x} \in [0,\infty)^d$, $\beta \in [1,\infty)$, and the resulting Archimax copulas (10) can immediately be identified as OPACs. Similarly, *hierarchical logistic stable tail dependence functions* are given by nested logistic stable tail dependence functions, for example

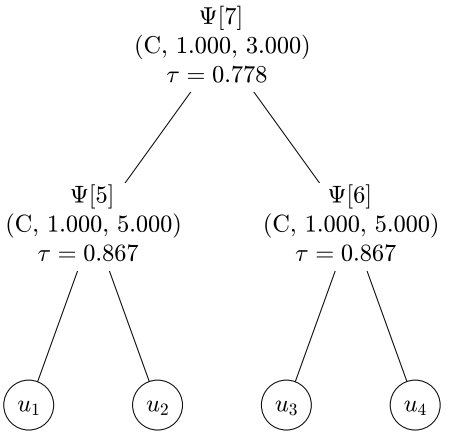
$$\ell(\boldsymbol{x}) = \ell_{\beta_0}(\ell_{\beta_1}(\boldsymbol{x}_1),\dots,\ell_{\beta_S}(\boldsymbol{x}_S)) = \left(\sum_{s=1}^S\left(\sum_{j=1}^{d_s} x_{sj}^{\beta_s}\right)^{\frac{\beta_0}{\beta_s}}\right)^{\frac{1}{\beta_0}},$$

where $\boldsymbol{x} = (\boldsymbol{x}_1,\dots,\boldsymbol{x}_S) \in [0,\infty)^d$ and $\beta_0,\dots,\beta_S \in [1,\infty)$ for which the resulting Archimax copula (10) can easily be shown to be a HOPAC. As such, our findings about (H)OPACs will also apply to Archimax copulas with (hierarchical) logistic stable tail dependence functions, one of the most tractable and thus widely used stable tail dependence functions of Archimax copulas.

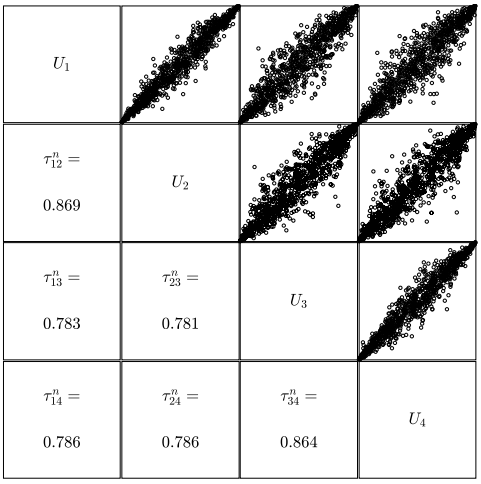
3.3. Estimating HOPACs

In this section, we develop estimators under the SNC for HOPACs under restrictions $\mathcal{R}_1$ and $\mathcal{R}_2$ given by (9), i.e., assuming the same one-parameter AC family a for all generators and that the condition on $[\psi^{-1}_{(a,\theta_{parent})}\{\psi_{(a,\theta_{child})}\}]'$ to be c.m. can be simplified to $\theta_{parent} \leq \theta_{child}$.

The literature provides different HAC estimation methods; see Okhrin et al. (2013a) or Górecki et al. (2016, 2017b) for those concerning estimation of both structure and parameters. However, as already mentioned in Section 1, all of them only consider HACs involving just one-parameter ACs. On the one hand, these methods cannot be directly used for HOPAC estimation, on the other hand they serve as a natural starting point for the development of such estimators.



(a) A HOPAC model built under the parametric SNC in the form $\beta_{parent} \leq \beta_{child}$, if $\theta_{parent} = \theta_{child}$.



(b) A sample of $n = 1000$ observations from the HOPAC of Figure 3a. Note that τ_{ij}^n denotes the pairwise Kendall's tau for observations of (U_i, U_j).

Fig. 3. An example for the Bottom-Up estimator.

In general, three ingredients are necessary to get a HAC estimated, (1) its structure, (2) the AC families and (3) their parameters. (1) typically encompasses a kind of agglomerative clustering, where the structure finally results from clustering of variables under concern according to some measure of similarity between pairs of the variables, e.g., according to Kendall's tau. As, in contrast to Okhrin et al. (2013a), the structure estimator in Górecki et al. (2016, 2017b) does not require any assumptions on the family of the nested ACs, it is immediately feasible also for HOPAC estimation. Moreover, there exist theoretical justifications for this estimator. Given a HAC, Okhrin et al. (2013b) show that its structure can be uniquely recovered from all its bivariate margins, and Theorem 2 in Górecki et al. (2017a) shows that this identification is possible just from all pairwise Kendall's taus, so the Kendall's tau matrix. Also, Corollary 6.2 in Górecki et al. (2017b) implies that, given any input Kendall's tau matrix, the estimator returns a tree in which Kendall's tau increases when going from the root down towards the leaves. As addressed there in Section 6.2, this follows from the fact that the clustering method implemented by the estimator always produces a *monotonic dendrogram* (Batagelj, 1981). Finally, as this estimator, formalized by Algorithm 1 in Górecki et al. (2017b) and recalled in Algorithm 1 here, showed the best results in the ratio of successfully estimated true HAC structures on the basis of simulation studies, see Górecki et al. (2016) or Uyttendaele (2017), we adopt it to our HOPAC estimation approach.

To estimate the families, the mentioned references provide two different approaches. The one used in Okhrin et al. (2013a) and Górecki et al. (2016) arbitrarily assumes the same family for all nested ACs, whereas then one in Górecki et al. (2017b) involves an extra model selection step that chooses for each AC the best fitting family out of some predetermined pool of families, and thus allows the families in the estimated HAC to differ. The first approach substantially simplifies the parametric constraints following from the SNC (to the condition $\theta_{parent} \leq \theta_{child}$ in most cases), whereas latter requires an extensive analysis of the input family pool as not all families can be mixed together or since the parameter ranges of the families in the pool have to be adjusted before the estimation process. Hence, when the OP transformation comes into play (which makes the estimation process substantially more complex even under the assumption of the same family for all nested ACs), the approach allowing for different nested families becomes computationally challenging. In this work, we thus focus on the case of having the same family for all ACs nested in a HOPAC. Note that this case can still cover estimation of HACs with one-parameter ACs from different families, for example, the families of Clayton, 12 and 14, as stated in Section 1, where the latter two family numbers correspond to the notation used in Nelsen (2006, pp. 116–119).

For parameter estimation, we use a mixture of existing step-wise procedures. This follows from the existence of a close relationship between bivariate margins of a HAC and the location of ACs nested in this HAC. To clarify, according to Proposition 3 in Górecki et al. (2016), given two random variables U_i and U_j from the vector $(U_1, \ldots, U_d)$ distributed according to a d-variate HAC, the copula of (U_i, U_j) is the AC allocated in the HAC structure at the fork where the two branches (from each of the leaves i and j to the root) meet for the first time. This AC (fork) is called *youngest common ancestor* of leaves i and j. In Fig. 3a, the youngest common ancestor of leaves 1 and 2 is fork 5 (AC $C_{\Psi[5]}$), whereas for leaves 1 and 3 it is fork 7 (AC $C_{\Psi[7]}$). It follows that the parameters of this copula can be estimated from the observations of U_i and U_j only. To this end, we use the AC maximum-likelihood (ML) estimator as in Okhrin et al. (2013a). There are

often multiple pairs with the same youngest common ancestor, which we utilize by estimating the parameter(s) of the youngest common ancestor from the observations of *all* pairs of random variables corresponding to this ancestor, and then averaging these estimates; aggregation by mean performed best in a simulation study involving the mean, maximum and minimum aggregation statistics implemented in Step 2 of Algorithm 3 in Górecki et al. (2016).

The concept of our HOPAC estimation approach is summarized in Algorithm 2, which requires two inputs: (1) a one-parameter Archimedean family a, e.g., from Table 1, and (2) realizations $\boldsymbol{u} = (u_{ij}) \in [0,1]^{n\times d}$ of the *pseudo-observations* $(U_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,d\}}$ given by

$$U_{ij} = \frac{n}{n+1}\hat{F}_{n,j}(X_{ij}) = \frac{R_{ij}}{n+1}, \tag{11}$$

where $\hat{F}_{n,j}$ denotes the *empirical distribution function* corresponding to the jth margin, R_{ij} denotes the *rank* of X_{ij} among $X_{1j},\dots,X_{nj}$, and $(X_{i1},\dots,X_{id})$, $i\in\{1,\dots,n\}$ are i.i.d. random vectors distributed according to a joint distribution function with continuous margins F_j, $j\in\{1,\dots,d\}$, and a HOPAC C. The algorithm returns a HOPAC from family a with estimated structure and parameters.

As an example, let C be the Clayton HOPAC from Fig. 3a, with $d=4$, and let $(\bar{u}_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,4\}}$ be a sample from it; see Fig. 3b. Now assume C to be unknown and let us estimate it based on the pseudo-observations $(u_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,4\}}$ of $(\bar{u}_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,4\}}$. Note that in practice, we mostly do not observe $\bar{u}_{ij}$ but rather the pseudo-observations u_{ij}.

The algorithm estimates the structure in its first two steps using Algorithm 1, which returns a binary tree $(\hat{\mathcal{V}},\hat{\mathcal{E}})$ and estimates $\hat{\tau}_5$, $\hat{\tau}_6$ and $\hat{\tau}_7$ of the Kendall's tau corresponding to each fork in that tree; see Fig. 4a. As described before, the families of nested ACs are assumed to be from some pool of available families, e.g., a pool implemented by the software toolbox we use. Such an assumption implies that the user, when deciding which family suits best for the considered data, should repeat the whole estimation process for each of the available families and then perform some extra evaluation of their fit. For simplicity, assume that the Clayton family was selected for all the nested ACs. Recall that the HOPAC estimates will be built based on the SNC under $\mathcal{R}_1$ and $\mathcal{R}_2$ stated in Proposition 2. Let us consider two possible estimators under $\mathcal{R}_1$ and $\mathcal{R}_2$ following the concept in Algorithm 2.

3.3.1. Bottom-Up estimator

The idea of the *Bottom-Up* estimator lies in traversing through the forks in the estimated structure in the way that one starts at the bottom of the structure and then continues up until the root is reached, similar to Okhrin et al. (2013a) and Górecki et al. (2016, 2017b). One way to achieve this utilizes Kendall's tau estimates returned by Algorithm 1, starting with the fork with the highest value, then processing the fork with the second highest value, etc., until one gets to the fork with the lowest value, the root.

Algorithm 1 HACs structure estimation (Górecki et al., 2017b, Algorithm 1)

Input:

1) (τ_{ij}^n) – the sample version of a Kendall's tau matrix

The estimation:

1. $\hat{\mathcal{V}} := \{1,\dots,2d-1\}$, $\hat{\mathcal{E}} := \emptyset$, $\mathcal{I} := \{1,\dots,d\}$
(recall that $\downarrow(i) = \{i\}$ for $i\in\{1,\dots,d\}$, where $\downarrow(i)$ denotes the set of all descendant leaves of the fork i for $i\in\{d+1,\dots,2d-1\}$)

for $k = 1,\dots,d-1$ **do**

2. find two nodes from $\mathcal{I}$ to join, i.e,
$(i,j) := \underset{\tilde{i}<\tilde{j},\ \tilde{i}\in\mathcal{I},\ \tilde{j}\in\mathcal{I}}{\operatorname{argmax}} \operatorname{avg}((\tau_{\tilde{i}\tilde{j}}^n)_{(\tilde{i},\tilde{j})\in\downarrow(\tilde{i})\times\downarrow(\tilde{j})})$
3. let the children of the fork $d+k$ (denoted by $\wedge(d+k)$) be the set $\{i,j\}$, i.e., $\hat{\mathcal{E}} := \hat{\mathcal{E}}\cup\{\{i,d+k\},\{j,d+k\}\}$, which implies that $\wedge(d+k)=\{i,j\}$ and $\downarrow(d+k) = \downarrow(i)\cup\downarrow(j)$
4. remove the nodes i and j from the clustering process (as they have been already joined) and add the fork $d+k$ to be considered for joining in the following steps, i.e.,
$\mathcal{I} := \mathcal{I}\cup\{d+k\}\setminus\{i,j\}$
5. estimate the Kendall's tau corresponding to the fork $d+k$, i.e.,
$\hat{\tau}_{d+k} := \operatorname{avg}((\tau_{\tilde{i}\tilde{j}}^n)_{(\tilde{i},\tilde{j})\in\downarrow(i)\times\downarrow(j)})$

end for

Output:

$(\hat{\mathcal{V}},\hat{\mathcal{E}},(\hat{\tau}_k)_{d+1}^{2d-1})$

Algorithm 2 The HOPAC estimation concept

1: compute the matrix of pairwise Kendall's tau (τ_{ij}) from $\boldsymbol{u}$
2: estimate the structure using Algorithm 1 with (τ_{ij})
3: **for** each fork k in the structure **do**
4: **for** each pair of leaves (i_k, j_k) in the structure such that k is the youngest common ancestor of i_k and j_k **do**
5: use ML to estimate the OPAC parameters θ and β based on (u_{m,i_k}, u_{m,j_k}), $m = 1, \ldots, n$
6: **end for**
7: aggregate the estimated θs and βs by computing their means $\hat{\theta}_k$ and $\hat{\beta}_k$
8: let $C_{\psi_{(a,\hat{\theta}_k,\hat{\beta}_k)}}$ be the OPAC estimate corresponding to the fork k in the estimated HOPAC structure
9: **end for**

In Fig. 3a, fork 5 corresponds to the bottom of the structure. As it is the youngest common ancestor of leaves u_1 and u_2, we compute the ML estimator for θ_5 and β_5 according to (15):

$$(\hat{\theta}_5, \hat{\beta}_5) = \underset{(\theta\in\Theta_C, \beta\in[1,\infty))}{\operatorname{argmax}} \sum_{i=1}^{n} \log c_{\psi_{(C,\theta,\beta)}}(u_{i1}, u_{i2}). \tag{12}$$

As node 6 is the youngest common ancestor of u_3 and u_4, the corresponding parameters θ_6 and β_6 can be estimated according to (12) with u_{i1} and u_{i2} being replaced by u_{i3} and u_{i4}, respectively. The estimated parameters are shown in Fig. 4b. Having the bottom level estimated, we continue to the upper levels until we reach to the root. Here the SNC comes into play.

As $\hat{\theta}_5 \neq \hat{\theta}_6$, it is clear that $\mathcal{R}_2$ is violated because it is impossible to satisfy both $\hat{\theta}_7 = \hat{\theta}_5$ and $\hat{\theta}_7 = \hat{\theta}_6$ for our example. The restriction $\mathcal{R}_2$ was, however, a constraint under which the model was built; see Fig. 3a. It follows that turning to restriction $\mathcal{R}_1$ prevents a good fit for node 7, as $\mathcal{R}_1$ requires that $\hat{\beta}_7 = 1$ and that $\hat{\theta}_7$ has to be trimmed to the closest value allowed by $\mathcal{R}_1$, i.e., to 0.844; see Fig. 4c. A sample of 1000 observations from this HOPAC estimate is shown in Fig. 4d. It is evident that choosing $\mathcal{R}_1$ instead of $\mathcal{R}_2$ substantially affects the fit; compare the distributions of the pairs (U_1, U_3), (U_2, U_3), (U_2, U_4) and (U_1, U_4) shown in Fig. 3b (which correspond to the true HOPAC) with the corresponding ones from the estimated model in Fig. 4d. A way to cope with this problem could be to test if the parameters θ of the children are all equal to some aggregated value like their mean. With m children of a given fork, this would however require additional $\binom{m}{2}$ tests (and there is also the problem of multiple testing), making the computation substantially more involved. An efficient solution requiring at most one test for each fork is presented in the following section.

3.3.2. Top-Down estimator

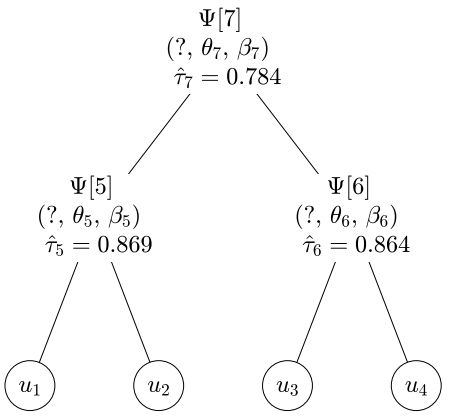
A solution to those problems consists of reversing the way in which the structure is traversed during the estimation, meaning starting at the root of the copula structure and then using the depth-first approach to go through all forks. This approach already appeared in connection to other hierarchical copula structures; see Zhu et al. (2016) or recently Cossette et al. (2019). Before considering our example with the *Top-Down* estimator, we first present it in terms of pseudo-code in Algorithm 3.

Let $\boldsymbol{u} = (u_{ij}) \in [0, 1]^{n\times d}$ be realizations of $(U_{ij})_{i\in\{1,\ldots,n\}}^{j\in\{1,\ldots,d\}}$ given by (11). As for the Bottom-Up estimator, these are first used to estimate the copula structure $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$ with Algorithm 1 based on the matrix of pairwise sample versions of Kendall's tau.

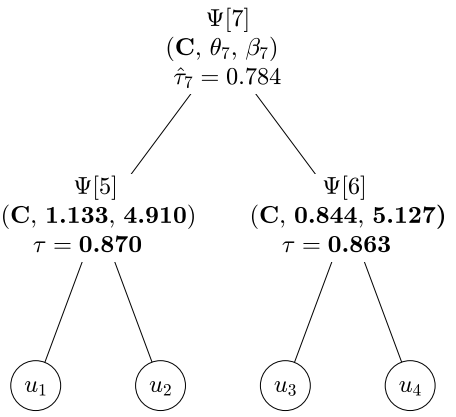
Estimation of the parameters is then performed by calling the function TopDown$\{\boldsymbol{u}, a, (\hat{\mathcal{V}}, \hat{\mathcal{E}}), 2d-1, \Theta_a, [1,\infty), \beta_{\mathcal{R}}\}$ presented in Algorithm 3, where a is an Archimedean family; $2d-1$ denotes the root in the binary tree $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$; Θ_a (see Table 1) and $[1, \infty)$ are ranges for the parameters θ and β of the OPAC estimate $C_{\hat{\psi}_{[2d-1]}}$ corresponding to the root; and $\beta_{\mathcal{R}} \in [1, \infty)$ is an upper bound for accepting $\beta_{parent} = 1$ in $\mathcal{R}_1$. Recall that *descendants* of a given node refer to the set of the children of that node, the children of these children, etc.

Several aspects of Algorithm 3 are worth being addressed:

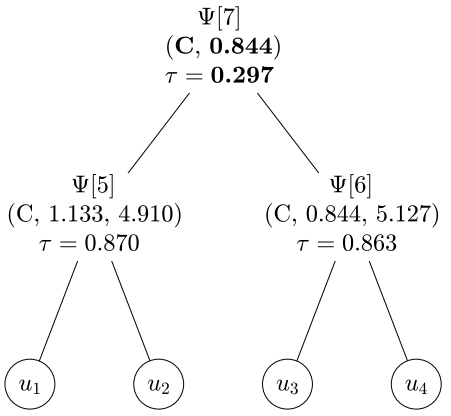
- The function TopDown recursively goes through all the forks of the tree $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$ in the depth-first search manner, which can be seen from steps 2, 4, 13 and 14. It is also important to note that the tree $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$ estimated via Algorithm 1 is binary. This restriction is, however, not an issue since a multivariate AC can be represented by a nested binary-structured AC with the same generator. Moreover, for accessing non-binary structures, several *collapsing* approaches have been proposed; see Segers and Uyttendaele (2014), Górecki et al. (2017b, Section 6.1) or Uyttendaele (2017).
- The assumption of $i < j$ in Step 4 is without loss of generality as in all remaining steps of the algorithm exchanging i and j does not affect the results.
- According to Remark 2 in Górecki et al. (2017a), all pairs from the Cartesian product of l_i and l_j defined in Steps 5 and 6 have the same youngest common ancestor k. It follows from Proposition 3 in Górecki et al. (2016) that bivariate margins corresponding to these pairs share the same copula, which motivates the mean aggregated ML estimator used in Step 7. Note that such an aggregation approach in a one-parameter version has already been successfully



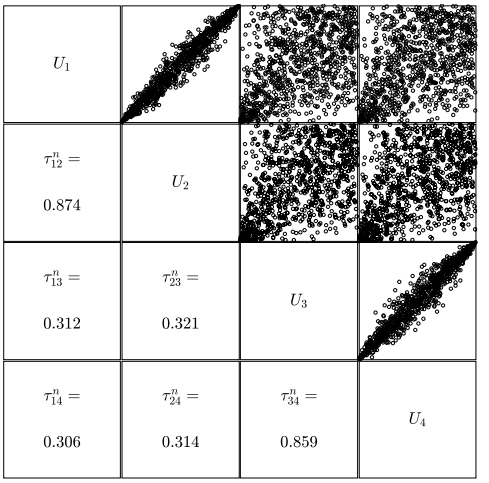
(a) The binary tree $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$ and estimates of Kendall's tau $(\hat{\tau}_5, \hat{\tau}_6, \hat{\tau}_7)$ for the sample from Figure 3b.



(b) An update of the estimate from Figure 4a after assuming the Clayton family and estimating the bottom-level using ML estimation.



(c) A final update of the estimate from Figure 4b after estimating the root under $\mathcal{R}_1$.



(d) A sample of $n = 1000$ observations from the HOPAC in Figure 4c.

Fig. 4. Evolution of a HOPAC model during Bottom-Up estimation. Note that the values in bold show what has been updated compared to a previous model.

used in Algorithm 3 in Górecki et al. (2016), see Step 2 therein. Also note that the viability of such an aggregation approach is studied in Section 4.

- As the result of the argmax in Step 7 is the vector of two components $(\theta_{\widetilde{ij}}, \beta_{\widetilde{ij}})$, all the operators to its left (the two sums and the division) are considered component-wise.
- The sets l_i and l_j are always disjoint, which follows from the fact that node i and node j do not lie at the same branch of $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$. This fact avoids that the same pair appears twice (first as (i, j) and then as (j, i)) in the first two sums in Step 7.
- The **if** statement in Step 8 decides which one of the restrictions $\mathcal{R}_1$ and $\mathcal{R}_2$ applies for the children i and j of node k at the recursive Steps 13 and 14. As the parameters $\hat{\theta}$ and $\hat{\beta}$ are estimated by ML, we have asymptotic normality and the variances of these estimates. As θ_{child} appearing in $\mathcal{R}_2$ is not yet available (it is estimated in further steps that depend on the decision made in Step 8), it is thus convenient to test for $\mathcal{R}_1$ as it requires only the value of β_{parent}. In practice, however, testing the hypothesis $\beta_{parent} = 1$ would slow down the computations, therefore we decided to only check whether $\hat{\beta}$ lies in the prescribed interval $[1, \beta_{\mathcal{R}}]$. The involved parameter $\beta_{\mathcal{R}}$ also allows us to emphasize one of the restrictions — we emphasize $\mathcal{R}_1$ with larger values of $\beta_{\mathcal{R}}$, whereas $\mathcal{R}_2$ with smaller ones. In the following, we use $\beta_{\mathcal{R}} = 1.05$ as it turned out to provide a convenient balance between $\mathcal{R}_1$ and $\mathcal{R}_2$.
- The output-triplet $(\hat{\mathcal{V}}, \hat{\mathcal{E}}, \hat{\Psi})$ contains the structure in $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$, the family and the parameters of the generators in $\hat{\Psi}$ of the HOPAC estimate.

Algorithm 3 The Top-Down HOPAC estimator

Inputs:

$\boldsymbol{u} = (u_{ij}) \in [0,1]^{n\times d}$: realizations of $(U_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,d\}}$ given by (11)
a : a one-parameter Archimedean family, e.g., from Table 1
$(\hat{\mathcal{V}}, \hat{\mathcal{E}})$: a binary tree (copula structure)
$k \in \{d+1, \dots, 2d-1\}$: a fork to estimate its parameters
$r_\theta \subset \mathbb{R}$: a range for parameter θ
$r_\beta \subseteq [1,\infty)$: a range for parameter β
$\beta_{\mathcal{R}} \in [1,\infty)$: an upper bound for accepting $\beta_{parent} = 1$ in $\mathcal{R}_1$

Output:

$(\hat{\mathcal{V}}, \hat{\mathcal{E}}, \hat{\Psi})$

Function TopDown$\{\boldsymbol{u}, a, (\hat{\mathcal{V}}, \hat{\mathcal{E}}), k, r_\theta, r_\beta, \beta_{\mathcal{R}}\}$

1: **if** $k \in \{1, \dots, d\}$ **then**
2: **return** – k is a leaf, so stop the recursion
3: **end if**
4: $\{i, j\} \leftarrow$ the children of k in $(\hat{\mathcal{V}}, \hat{\mathcal{E}})$ – assume $i < j$
5: $l_i \leftarrow$ the set of descendant leaves of i if i is a fork, otherwise $l_i \leftarrow \{i\}$
6: $l_j \leftarrow$ analogous to l_i (by replacing i by j)
7: perform MLE $\#l_i\#l_j$-times and aggregate the resulting parameter estimates (component-wise) by mean, i.e.,

$$(\hat{\theta}, \hat{\beta}) \leftarrow \frac{1}{\#l_i\#l_j} \sum_{\tilde{i}\in l_i} \sum_{\tilde{j}\in l_j} \operatorname*{argmax}_{(\theta_{\tilde{i}\tilde{j}}, \beta_{\tilde{i}\tilde{j}})\in\Theta_a\times[1,\infty)} \sum_{m=1}^{n} \log c_{\psi_{(a,\theta_{\tilde{i}\tilde{j}},\beta_{\tilde{i}\tilde{j}})}}(u_{m\tilde{i}}, u_{m\tilde{j}})$$

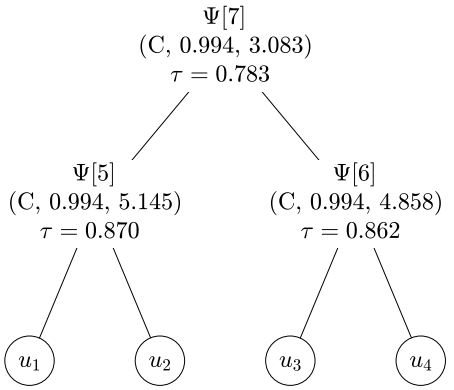
8: **if** $\hat{\beta} \in [1, \beta_{\mathcal{R}}]$ **then**
9: $\tilde{r}_\theta \leftarrow r_\theta \cap [\hat{\theta}, \infty)$ and $\tilde{r}_\beta \leftarrow r_\beta$ – apply restriction $\mathcal{R}_1$ for the children
10: **else**
11: $\tilde{r}_\theta \leftarrow [\hat{\theta}, \hat{\theta}]$ and $\tilde{r}_\beta \leftarrow [\hat{\beta}, \infty)$ – apply restriction $\mathcal{R}_2$ for the children
12: **end if**
13: TopDown$\{\boldsymbol{u}, a, (\hat{\mathcal{V}}, \hat{\mathcal{E}}), i, \tilde{r}_\theta, \tilde{r}_\beta, \beta_{\mathcal{R}}\}$ – depth-first traversing
14: TopDown$\{\boldsymbol{u}, a, (\hat{\mathcal{V}}, \hat{\mathcal{E}}), j, \tilde{r}_\theta, \tilde{r}_\beta, \beta_{\mathcal{R}}\}$ – depth-first traversing
15: $\hat{\Psi}[k] \leftarrow \psi_{(a,\hat{\theta},\hat{\beta})}$ – store the estimated parameters to the output

To illustrate the HOPAC estimation with the Top-Down approach, let $\boldsymbol{u} = (u_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,d\}}$ be the pseudo-observations of the sample from Fig. 3b, i.e., $d = 4$ and $n = 1000$. To estimate the structure, apply Algorithm 1, resulting in the tree $(\hat{\mathcal{V}}, \hat{\mathcal{E}}) = (\{1, \dots, 7\}, \{\{1,5\}, \{2,5\}, \{3,6\}, \{4,6\}, \{5,7\}, \{6,7\}\})$, which corresponds to the tree depicted in Fig. 3a. Let a be again the Clayton family, and recall that $\Theta_{\mathrm{C}} = (0,\infty)$; see Table 1. Finally, to obtain the parameter estimates, call TopDown$(\boldsymbol{u}, \mathrm{C}, (\hat{\mathcal{V}}, \hat{\mathcal{E}}), 7, \Theta_{\mathrm{C}}, [1,\infty), 1.05)$.

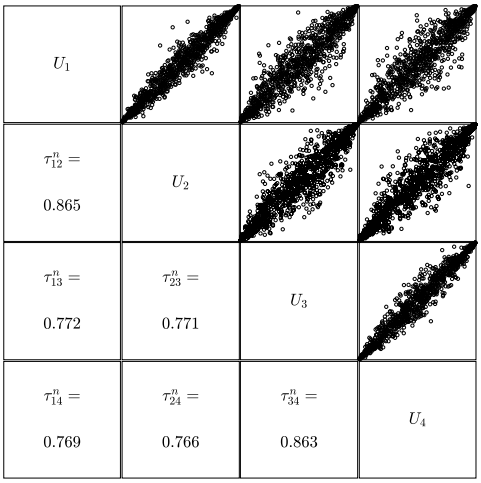
In Step 4, $i \leftarrow 5$ and $j \leftarrow 6$. In the next two steps, $l_5 \leftarrow \{1,2\}$ and $l_6 \leftarrow \{3,4\}$. Step 7 computes the argmax for $(\tilde{i}, \tilde{j}) \in ((1,3), (2,3), (1,4), (2,4))$. The four pairs of $(\theta_{\tilde{i}\tilde{j}}, \beta_{\tilde{i}\tilde{j}})$ are (0.99, 3.106), (0.94, 3.124), (1.07, 3.006) and (0.976, 3.095). Using the component-wise mean results in $(\hat{\theta}, \hat{\beta}) \leftarrow (0.994, 3.083)$. In the next step (as $\hat{\beta} > \beta_{\mathcal{R}}$) the restriction $\mathcal{R}_2$ is applied, resulting in $r_\theta \leftarrow [0.994, 0.994]$ and $r_\beta \leftarrow [3.083, \infty)$.

As the recursive Steps 13 and 14 involve the estimation of bivariate OPACs for forks 5 and 6 ($l_i \leftarrow \{i\}$ and $l_j \leftarrow \{j\}$ in both of the nested calls of TopDown), we just show the results of Step 15, which are $\hat{\Psi}[5] \leftarrow \psi_{(C,0.994,5.145)}$ and $\hat{\Psi}[6] \leftarrow \psi_{(C,0.994,4.858)}$. Finally, Step 15 results in $\hat{\Psi}[7] \leftarrow \psi_{(C,0.994,3.083)}$. The resulting estimated HOPAC $C_{(\hat{\mathcal{V}},\hat{\mathcal{E}},\hat{\Psi})}$ is depicted in Fig. 5a. We observe that the parameters are relatively close to the true parameters (shown in Fig. 3a), particularly in comparison to the Bottom-Up approach, for β of fork 7. This is further reflected via distributions of samples from these HOPACs; compare the distributions and particularly the strength of the correlation in the pairs (U_1, U_3), (U_2, U_3), (U_2, U_4) and (U_1, U_4) shown in Figs. 3b, 4c and 5b corresponding to the true copula, the Bottom-Up and Top-Down estimate, respectively. In contrast to the Bottom-Up estimate, the Top-Down estimate closely follows the true distribution.

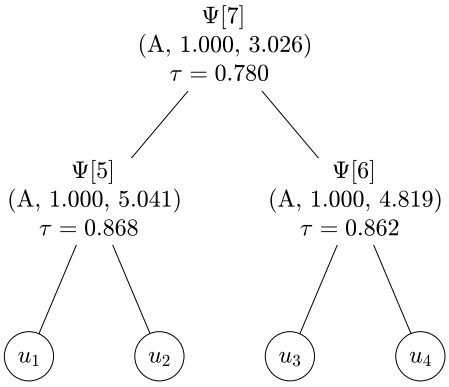
It is clear that the arbitrary assumption of the Clayton family needs extra attention. As suggested above, an extra criterion should be used to evaluate feasibility of such an assumption. To this end, the goodness-of-fit test statistic S_n used in the estimator defined in (16) can be used; see Genest et al. (2009). Evaluating this statistic for the sample from Fig. 3b and the Top-Down HOPAC estimate shown in Fig. 5a, we obtain $S_n = 0.0148$. For $a = \mathrm{A}$ we get $S_n = 0.0148$, for $a = \mathrm{F}$ we observe $S_n = 0.0694$ and for $a = \mathrm{J}$ we receive $S_n = 0.1516$. It is not surprising that S_n for the true family is minimal. What might be surprising is that the minimum is also obtained for the Ali–Mikhail–Haq family. However, looking at page 117 in Nelsen (2006), Table 4.1 shows that for $\theta = 1$ the copulas from Clayton and Ali–Mikhail–Haq (there denoted 4.2.1



(a) A HOPAC estimate from sample in Figure 3b under assumption of Clayton family.



(b) A sample of $n = 1000$ observations from the HOPAC in Figure 5a.



(c) A HOPAC estimate from the sample in Figure 3b based on the Ali-Mikhail-Haq family. Note that the values shown are rounded to 3 digits, hence, e.g., even if all θs are < 1, the rounded values can be 1.000.

Fig. 5. Top-Down estimates.

and 4.2.3, respectively) are both equal to $C(u, v) = uv/(u + v - uv)$. Looking further at the resulting Top-Down estimate for Ali–Mikhail–Haq shown in Fig. 5c, and considering that the parameters θ for both families are relatively close to 1, this result rather confirms that the presented framework works correctly.

4. Simulation study

To evaluate the HOPAC estimator presented in Section 3.3.2, $N = 500$ repetitions of the following routine is performed for each of the families Ali–Mikhail–Haq, Clayton, Frank and Joe. This routine first randomly generates a HOPAC model, then samples from it, computes several estimates based on the sample, and finally measures certain types of discrepancy between the model and the estimate, and eventually between the sample and the estimate. More precisely, the setup is as follows:

1. Given a dimension $d \in \{5, 10, 20\}$, randomly generate a correlation matrix of size $d \times d$ according to an algorithm based on computing a lower triangular matrix from angles sampled from the distribution scaled from $\sin^k(x)$; see Makalic and Schmidt (2018).
2. This matrix is then passed to Algorithm 1, which returns a binary tree with d leaves that serves as the *structure* of the randomly generated HOPAC model. The algorithm also returns the estimates $\hat{\tau}_{d+1}, \ldots, \hat{\tau}_{2d-1}$ corresponding to the forks in that tree, which are used, in the next step, to generate the *parameters* of the HOPAC model.

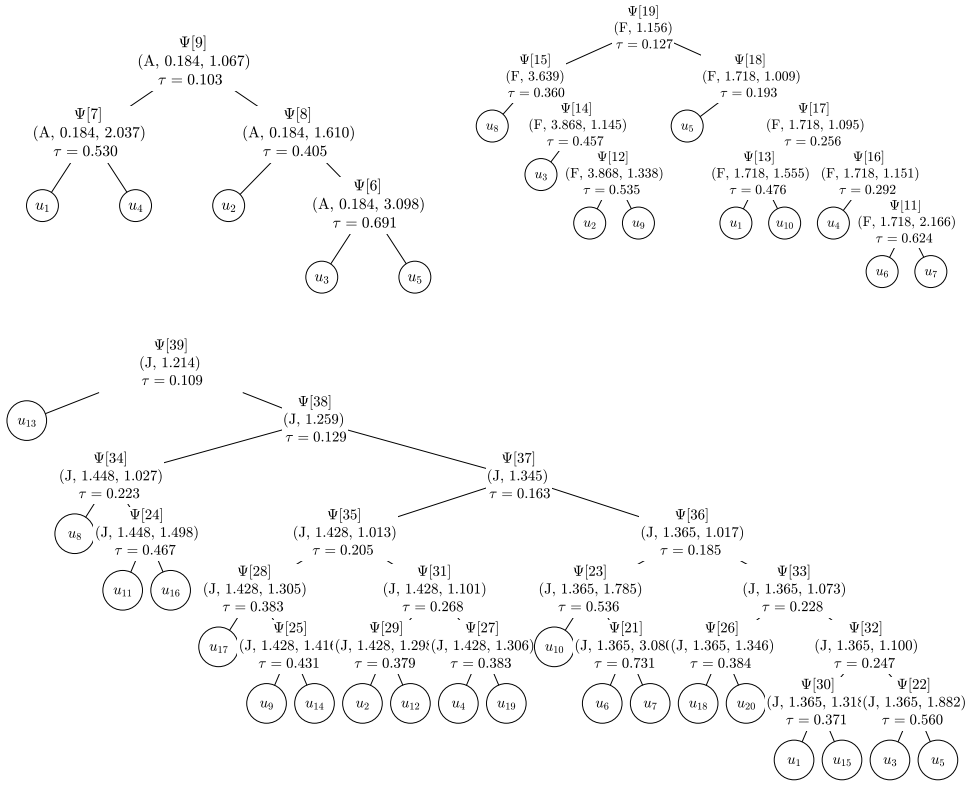


Fig. 6. Three randomly generated HOPAC models.

3. The forks in the structure are traversed depth-first starting from the root ($k = 2d - 1$) and for each given fork $k \in \{d + 1, \dots, 2d - 1\}$, the parameters θ and β are set as follows:
   (a) if the parent's β is 1 (assumed true for the root), then the fork's β is set equal to 1, i.e., to the case corresponding to $\mathcal{R}_1$, with probability of 50%. Hence, if β is 1, the parameter θ is just adjusted in a way that Kendall's tau of this fork remains equal to $\hat{\tau}_k$. For the case corresponding to $\mathcal{R}_2$ (the remaining 50%), the parameter θ is first generated randomly and then β is adjusted to keep Kendall's tau equal to $\hat{\tau}_k$.
   (b) if the parent's β is greater than 1, then the fork's θ is set equal to the parent's θ and the fork's β is adjusted to keep Kendall's tau equal to $\hat{\tau}_k$.

   For examples, see Fig. 6.
4. Assume the same family $a \in \{\text{A, C, F, J}\}$ for each OPAC nested in the HOPAC model and sample from it with sizes $n \in \{200, 400, \dots, 1000\}$.
5. Based on these samples, compute realizations of the following estimators:
   (a) The OPAC estimator (denoted **OPAC**) defined by

$$\frac{1}{\binom{d}{2}} \sum_{i=1}^{d} \sum_{j=i+1}^{d} \operatorname*{argmax}_{(\theta_{ij},\beta_{ij}) \in \Theta_a \times [1,\infty)} \sum_{m=1}^{n} \log c_{\psi_{(a,\theta_{ij},\beta_{ij})}}(u_{mi}, u_{mj}). \tag{13}$$

   We include this estimator in our study in order to show to which level OPACs are (un)able to fit HOPACs, in other words to assess, how important the presence of hierarchy/structure in the copula model is. Also note that accessing the density c_ψ can be challenging due to the need of differentiating the cumulative distribution function d-times. The estimator given by (13) is thus used instead of the standard (non-aggregated) generalization of (15). It is, however, worth noting here that this simple (OPAC) estimator shows an excellent improvement/complexity trade-off in the tail-dependence modeling application reported in Section 5, which hints at feasibility of such an aggregation approach in general.
   (b) The HAC estimator (**HAC**) based on one-parameter generators given by Algorithm 3, where the OP transformation is avoided simply by setting the input argument r_β equal to [1, 1]. Note that this estimator is included in our study in order to stress the importance of having the OP transformation in the copula model.

(c) The HOPAC Top-Down estimator (**HOPAC-miss-fam**) given in Algorithm 3 with misspecified family a in order to study the robustness of the estimator against family misspecification. The input a of Algorithm 3 is set C if the true family is A (shortly C instead A), F instead C, J instead F and finally A instead J. Note that all the other estimators assume the true family for the input a.

(d) The HOPAC Top-Down estimator (**HOPAC-indep-lkh**) given in Algorithm 3 with the OPAC ML estimator in Step 7 being replaced by an alternative estimator given by

$$\operatorname*{argmax}_{(\theta,\beta)\in\Theta_a\times[1,\infty)} \sum_{m=1}^{n} \sum_{\tilde{i}\in I_i} \sum_{\tilde{j}\in I_j} \log c_{\psi(a,\theta,\beta)}(u_{m\tilde{i}}, u_{m\tilde{j}}),$$

which approaches the aggregation via assuming independence of the pairs $(U_{\tilde{i}}, U_{\tilde{j}})$ and thus considering sums of their log-likelihoods, which allows one to perform the argmax optimization only once.

(e) The HOPAC Top-Down ML estimator (**HOPAC**) given exactly according to Algorithm 3.

6. For each sample (eventually replaced by the model) and estimator, evaluate the following six measures concerning their discrepancy in the distribution, Kendall's tau, upper-tail dependence coefficient and parameters. These measures can be divided into the following two groups:

(i) **True versus estimate**. This group includes the three measures given at the top of Figs. 7 and 8, where (θ_i, β_i), τ_i and λ_i^u correspond to the fork i in the copula model whereas $(\hat{\theta}_i, \hat{\beta}_i)$, $\hat{\tau}_i$ and $\hat{\lambda}_i^u$ correspond to the fork i in the copula estimate; the lower-tail dependence coefficient is not considered as it equals 0 for all families considered except Clayton, see Table 1. As the first measure (left columns) involves comparison based on the parameter θ, which shares different scale for different families, the HOPAC-miss-fam estimator is removed there. Note that these measures require that the structure of the model and the estimate match and are thus evaluated only for the HOPAC models. To generate $N = 500$ estimates matching the true structure, a new sample according to the model is generated in each out of $N = 500$ repetitions until the structure returned by Algorithm 1 equals the true structure. The ratio of true structures returned out of $N = 500$ trials is depicted in Fig. 11a.

(ii) **Sample versus estimate**. This group includes the three measures given at the top of Figs. 9 and 10, where $\hat{C}$ and C_n denote the estimated and empirical copulas, respectively, $\hat{\tau}_{ij}$ and $\hat{\lambda}_{ij}^u$ denote the Kendall's tau and upper-tail dependence coefficient corresponding to the youngest common ancestor of leaves i and j in the estimated structure, respectively, τ_{ij}^n denotes the sample version of Kendall's tau corresponding to variables U_i and U_j, and $\lambda_{ij}^{u,n,1\%}$ denotes the non-parametric estimate of the upper-tail dependence coefficient for variables U_i and U_j at the level $k/n = 0.01$ according to (13) in Schmidt and Stadtmüller (2006), where k is fixed so that $k/n = 0.01$. Note that for larger values of k/n, we have observed in some cases that the estimator produces a bias oriented towards the values estimated by the HOPAC-miss-fam estimator, resulting for some families in undesirable behavior of the measure in which the HOPAC-miss-fam estimator outperformed the (true-family-based) HOPAC estimator. The results are presented using the log-scale.

It can be observed that:

- As n increases, all measures decrease (converge to 0) for all HOPAC estimators.
- The estimators HOPAC and HOPAC-indep-lkh show lowest means and standard errors in most of the cases. The results for these two estimators are the most similar among all considered ones; we observe only slightly larger variances in parameter estimation (left columns, true versus estimate) for HOPAC-indep-lkh.
- The estimators without hierarchy (OPAC) or with no OP transformation available (HAC) are unable to model HOPAC data, as is clear from Figs. 9 and 10.
- All the previous observations do not depend on the underlying family.
- Misspecification of the family (reflected by the HOPAC-miss-fam estimator) mostly provides results that are, on the one hand, slightly worse than those of the HOPAC estimator, but on the other hand better than the results for the OPAC and the HAC estimator.
- As already mentioned in Remark 1, Archimax copulas with (hierarchical) logistic stable tail dependence functions are also (H)OPACs. As such, our findings immediately apply to such type of tractable Archimax copulas.

The quality of the structure estimator (Algorithm 1) is also evaluated, see Fig. 11. Note that each bar in Fig. 11a shows the value $N/m \times 100$, where m is the number of sampling repetitions until $N = 500$ true structures have been recovered. As such, an equal-or-not criterion is too strict as it does not take into account *how much* the estimated structure differs from the true structure. An extra proportional criterion based on a trivariate decomposition of the full structure according to Segers and Uyttendaele (2014) is also evaluated, see Fig. 11b. There, each bar shows the value $r/m \times 100$ for $r = \sum_{i=1}^{m} r_j$, where r_j is the ratio expressing how close is the estimated structure to the true one (1 if and only if these two are equal) computed as follows:

1. Decompose the true structure to $\binom{d}{3}$ trivariate true structures according to Segers and Uyttendaele (2014). Analogously for the estimated structure.

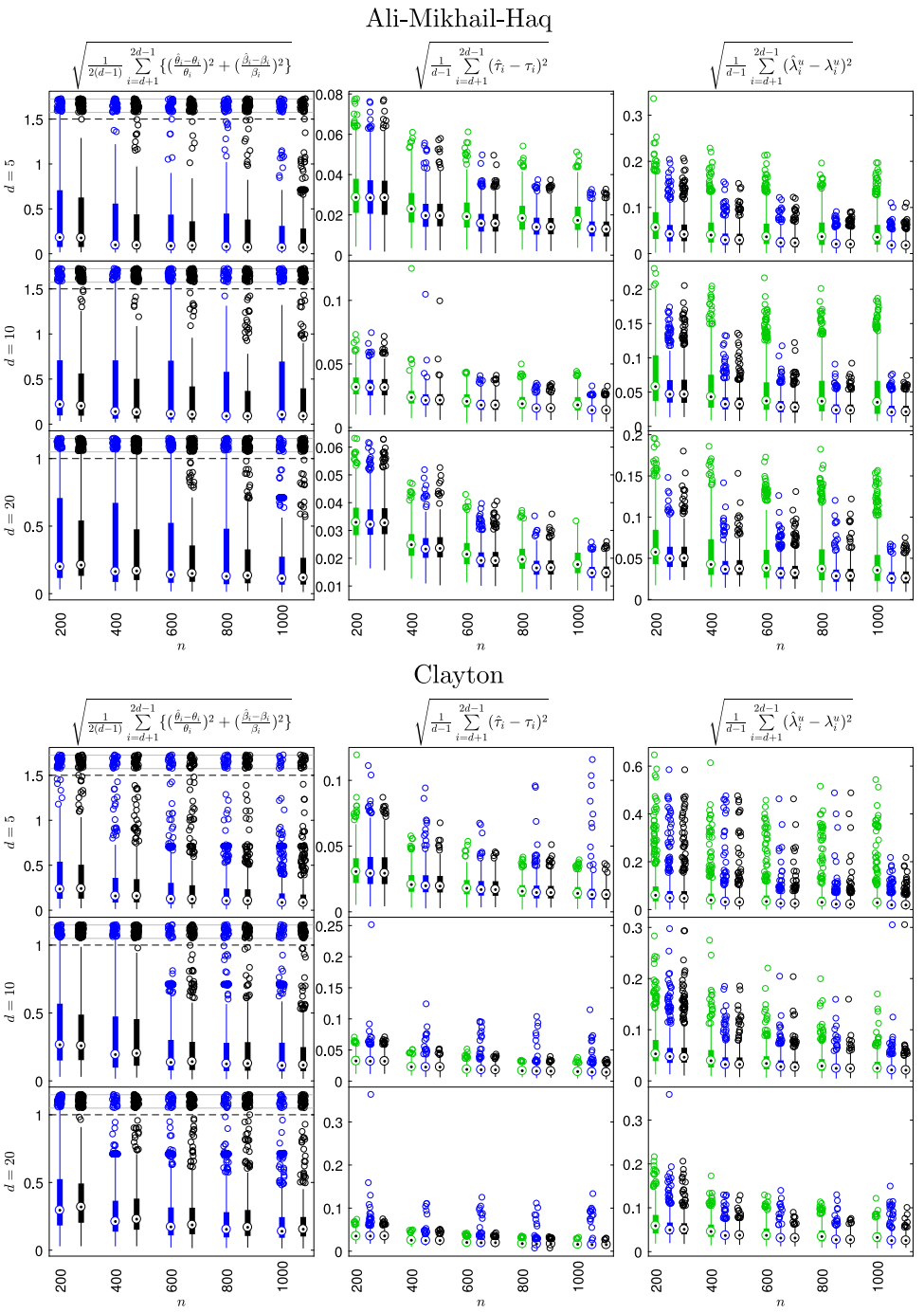


Fig. 7. Realizations of the **true versus estimate** measures for Ali–Mikhail–Haq and Clayton copulas estimated by HOPAC-miss-fam, HOPAC-indep-lkh and HOPAC.

2. Compare each trivariate true structure to the corresponding trivariate estimated structure and compute how many times these two match, say $s \in \{0, \dots, \binom{d}{3}\}$.
3. $r_j = s/\binom{d}{3}$.

Note that such a criterion has already been used, e.g., in Uyttendaele (2017). As can be observed, the ratio of estimated true structures is:

- Independent of the true family used for sampling.
- Increasing with n.

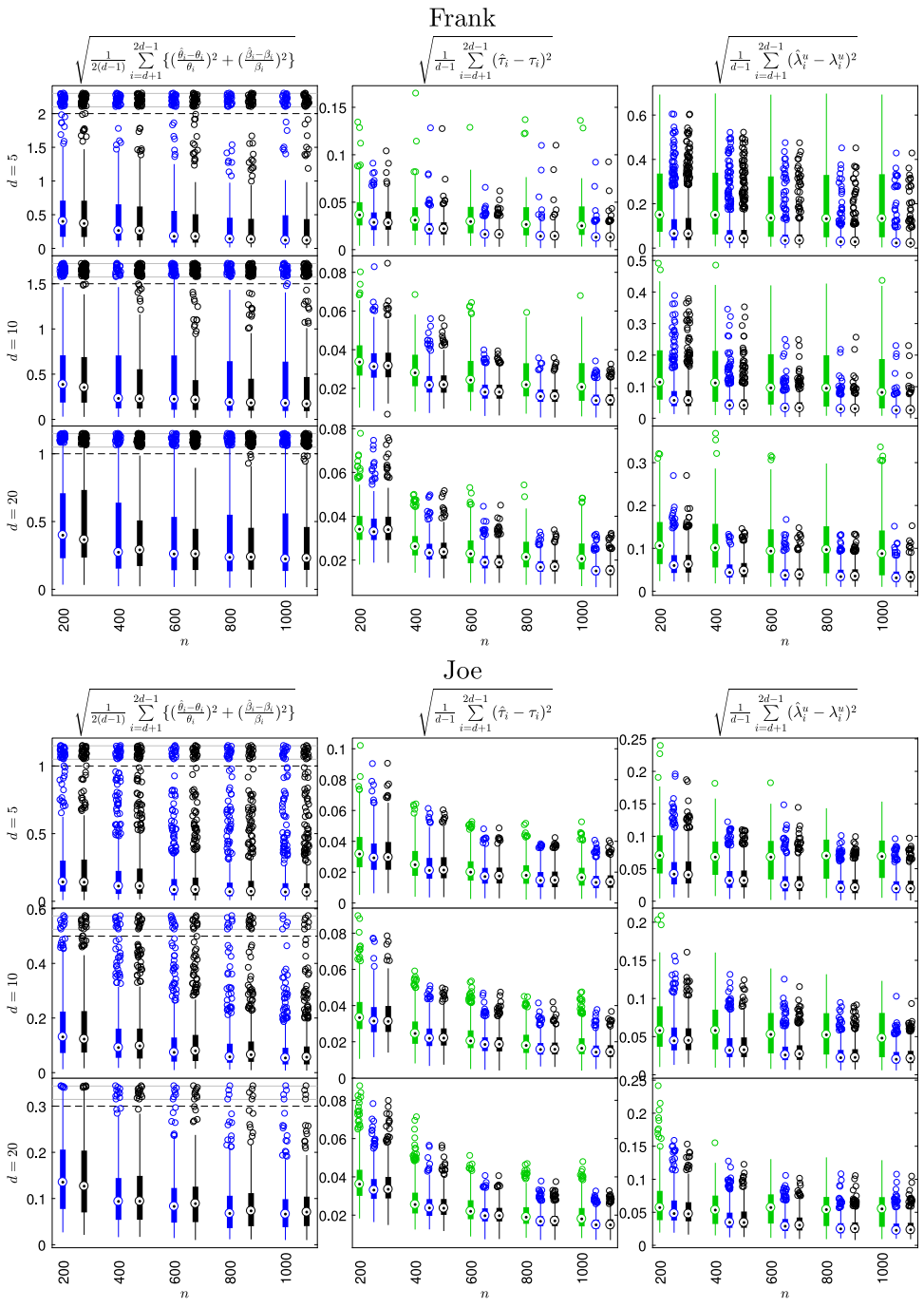


Fig. 8. Realizations of the **true versus estimate** measures for Frank and Joe copulas estimated by HOPAC-miss-fam, HOPAC-indep-lkh and HOPAC.

- Converging to 100 (in n); this convergence is slower as d increases. The impact of increasing d is substantially lower for the proportional equal-or-not criterion.

Finally note that other classes of copulas, e.g., elliptical or vine (Czado, 2010; Joe and Kurowicka, 2011), could be included in this simulation study. These were, however, not included as:

1. It follows directly from their theoretical construction that radially symmetric elliptical copulas cannot fit asymmetric HOPACs;

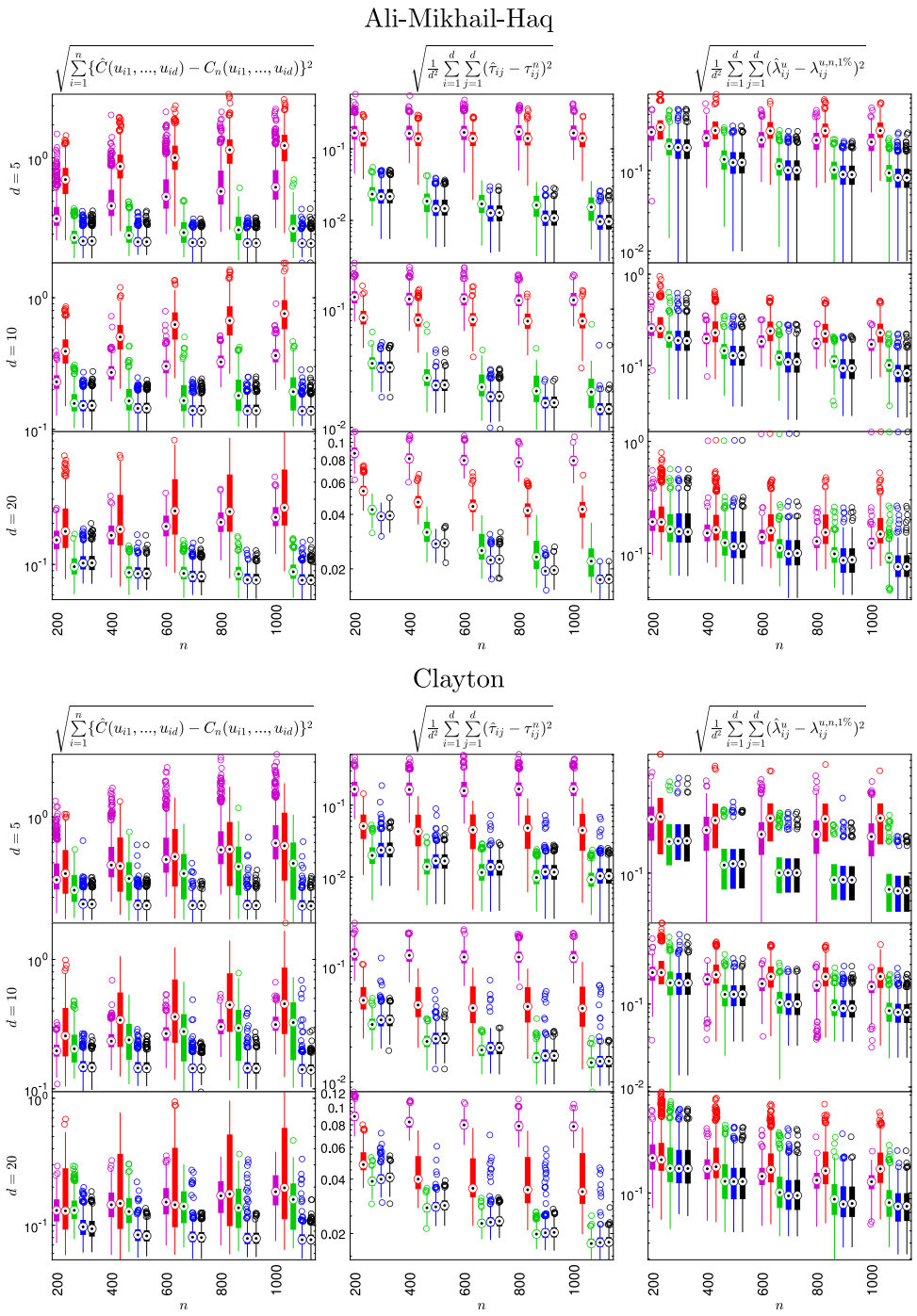


Fig. 9. Realizations of the **sample versus estimate** measures for Ali–Mikhail–Haq and Clayton copulas estimated by OPAC, HAC, HOPAC-miss-fam, HOPAC-indep-lkh and HOPAC.

2. For vine copulas, realizations of the "sample versus estimate" measures require either the cumulative distribution function ($\hat{C}$), which is computationally demanding already for $d = 5$, or to access all bivariate margins (to get $\hat{\tau}_{ij}$ or $\hat{\lambda}_{ij}^u$), which is, in general, not possible.

Nevertheless, these two important classes of copulas are included as benchmarks in the application reported in Section 5, where they are compared to HOPACs in their ability of tail dependence modeling.

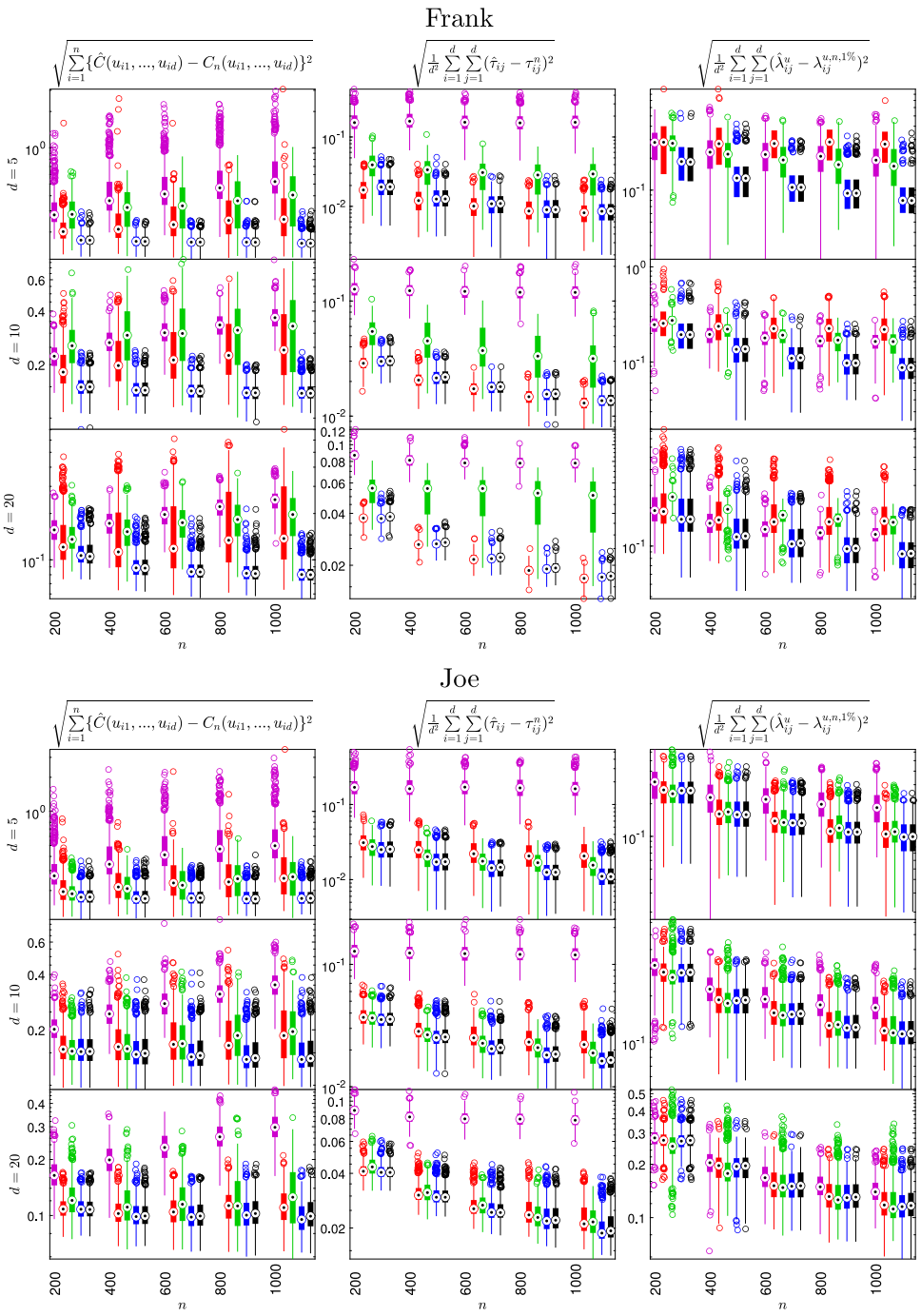
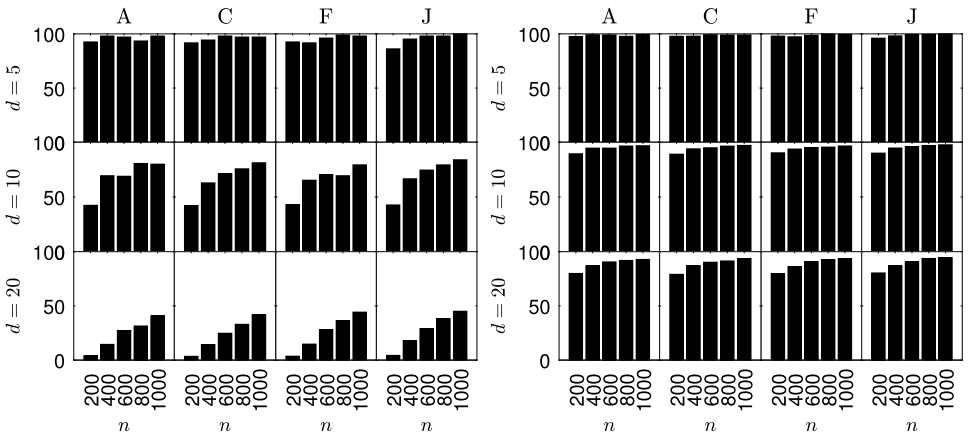


Fig. 10. Realizations of the **sample versus estimate** measures for Frank and Joe copulas estimated by OPAC, HAC, HOPAC-miss-fam, HOPAC-indep-lkh and HOPAC.

5. Empirical study

Value-at-Risk (VaR) is an important risk measure in quantitative risk management. In this section, we consider two different datasets of daily stock prices downloaded from Alpha Vantage.[1] The first one contains the five time series of stock prices of ADI (Analog Devices, Inc.), AVB (Avalonbay Communities Inc.), EQR (Equity Residential), LLY (Eli Lilly and

[1] www.alphavantage.co.



(a) Equal-or-not criterion for the full structure.

(b) Proportional equal-or-not criterion based on a trivariate decomposition of the full structure (Segers and Uyttendaele, 2014).

Fig. 11. The y-axis shows the ratio of estimated true structures according to a selected criterion.

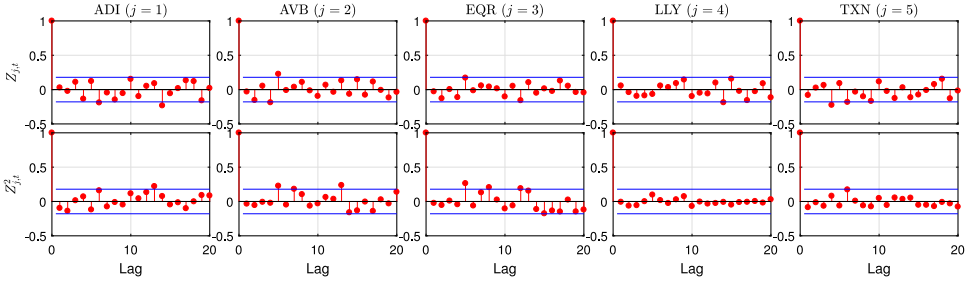


Fig. 12. Partial autocorrelation function of the $Z_{j,t}$ (first row) and $Z_{j,t}^2$ (second row) from the GARCH(1, 1) models fitted to the last half of a year of the considered time period for the five-dimensional portfolio.

Company) and TXN (Texas Instruments Inc.). It is important here that the clustering of the companies is given by their industry sector: ADI and TXN belong to the high-tech industry, AVB and EQR to the real estate industry and LLY to the pharmaceutical industry. Therefore, we would expect that the structure of the HOPAC or HAC used in the study will resemble these groupings and thus that the copula structure will play an important role. The second dataset contains the first 10 time series of daily stock prices from the S&P500 according to their market capitalization. For both datasets we use the time span 2002-02-01–2019-02-01.

The negative profit-and-loss random variable of the portfolio is defined as $L_{t+1} = \sum_{j=1}^{d} b_j P_{j,t}(1 - e^{R_{j,t+1}})$, where $P_{j,t}$ and $R_{j,t}$ are the price and log-return respectively of asset j at time point t, d is the dimension of the portfolio (5 for the first dataset, 10 for the second). Weights of the assets in the portfolio are denoted by b_j, $j = 1, \ldots, d$, with $\sum_{j=1}^{d} b_j = 1$. As the study aims at proving the general power of the HOPAC model and not the comparison between different portfolio allocation schemes we consider only the equally weighted portfolio $b_j = 1/d$, $j = 1, \ldots, d$, advocated by DeMiguel et al. (2009). Let F_L denote the distribution function of L_{t+1}. This leads to $\mathrm{VaR}(\alpha) = F_L^{-1}(\alpha)$ as VaR of the portfolio at level α. We focus on $\alpha = \{0.95, 0.99\}$ in this study. The distribution function F_L is estimated by simulating paths of the asset returns from the underlying multivariate process estimated in a rolling window fashion on windows of widths $w = \{126, 252, 504\}$. This corresponds to half a year, one year and two years of data. The underlying temporal dependency is modeled by marginal GARCH(1, 1) models (deGARCHing) with t-distributed innovations:

$$R_{j,t} = \mu_{j,t} + \sigma_{j,t} Z_{j,t} \quad \text{with} \quad \sigma_{j,t}^2 = \omega_j + \alpha_j \sigma_{j,t-1}^2 + \beta_j (R_{j,t-1} - \mu_{j,t-1})^2,$$

and $\omega > 0, \alpha_j \geq 0, \beta_j \geq 0, \alpha_j + \beta_j < 1$. Partial autocorrelation functions computed for the (squared) standardized residuals of the GARCH(1, 1) models for the five-dimensional portfolio on the last 126 points from the dataset are provided in Fig. 12. They indicate that most of the temporal dependence has been removed and the standardized residuals can be considered stationary. It seems difficult to provide a more in-depth analysis and selection of the time series models as they are applied on each moving window. Afterwards, various copula models are estimated from the pseudo observations of the standardized residuals $Z_{j,t}$, $t = 1, \ldots, T = 4279$, $j = 1, \ldots, d$. Thus, the estimated $\widehat{\mathrm{VaR}}_{t,w}(\alpha)$ at a given time

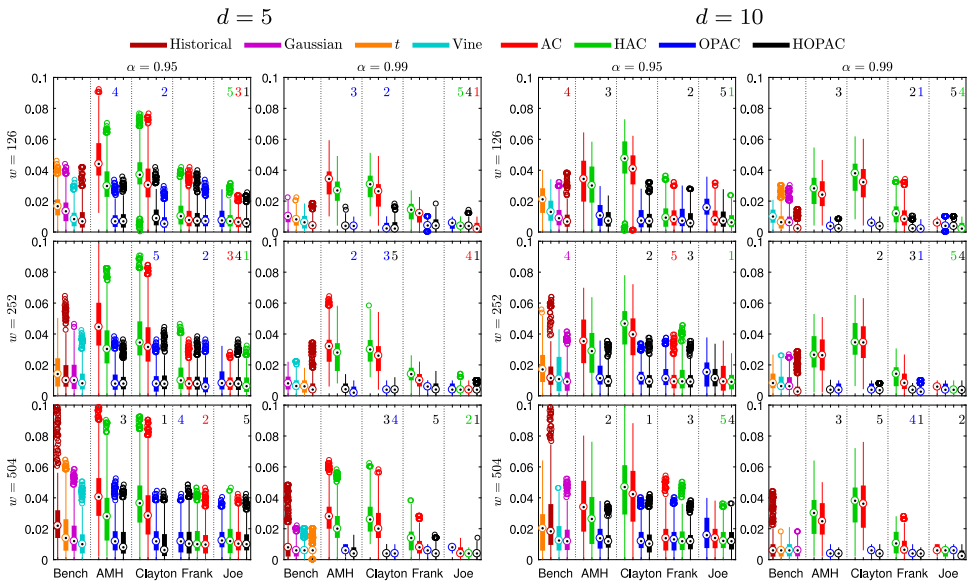


Fig. 13. Comparison of the models on the basis of $|\hat{\alpha}_t - \alpha|$ for $d = \{5, 10\}$, $\alpha = \{0.95, 0.99\}$, $w = \{126, 252, 504\}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

point t, window width w and level α is computed as follows: (a) we estimate the GARCH(1, 1) for all univariate marginal time series of log-returns in the time interval $(t - w - 1, t - 1]$; (b) extract standardized residuals, build their pseudo-observations and estimate the copula; (c) simulate a sample of size $n = 1000$ from the estimated copula, plug them into the estimated GARCH equation in order to obtain 1000 predictions of log-returns for the time point t and compute the corresponding 1000 predictions of negative profit-and-loss realizations; and (d) compute empirical quantiles at level α of the 1000 predicted negative profit-and-loss realizations. The number of the simulated trajectories, $n = 1000$, should be sufficient, as it leads to ten exceedances on average for our most extreme case of $\alpha = 0.99$. Furthermore, the evaluation of each model is made on the basis of the *VaR violation ratio* in the rolling window approach with the window width being 500 forecasts and a step of five days to allow for some uncertainty quantification

$$\hat{\alpha}_t = \frac{1}{500} \sum_{s=t}^{t+499} \mathbb{1}_{\{L_s > \widehat{\mathrm{VaR}}_{s,w}(\alpha)\}},$$

where $t \in \{505, 510, 515, \ldots, 3780\}$, i.e., 656 values. The closer $\hat{\alpha}_t$ is to the theoretical level α, the better the model. We thus compare the absolute deviations $|\hat{\alpha}_t - \alpha|$ for the different models. We also considered various tests in the spirit of Kupiec (1995) but as they did not give any new insights visible from pure deviations we decided not to present them in the paper.

All in all, our study considers 20 models: ACs, OPACs, HACs and HOPACs for the families of Ali–Mikhail–Haq, Clayton, Frank and Joe; Gaussian and t-copulas; R-Vine copulas, see Coblenz (2019), and the quantile-based historical estimator (denoted "Historical") which is computed directly on the true profit-and-loss variates without any underlying time-series model.

The results are summed up in Fig. 13, where the first two columns of panels correspond to the five-dimensional portfolio and the second two columns to the 10-dimensional portfolio. For each portfolio the first column depicts the results for $\alpha = 0.95$ and the second for $\alpha = 0.99$. The rows show the results for the different moving window widths $w = \{126, 252, 504\}$. Each panel represents box-plots of the deviations $|\hat{\alpha}_t - \alpha|$ for all the models. In each panel we have, separated by the vertical dotted lines, five groups of box-plots: 1. Benchmark models, 2. Ali–Mikhail–Haq, 3. Clayton, 4. Frank and 5. Joe families. Box-plots in each group are ordered in decreasing order based on their median values. The benchmark models in the first four box-plots in each panel are Historical (brown), Gaussian (violet), t (orange) and vine (navy) models. Red, green, blue and black box-plots represent the AC, HAC, OPAC and HOPAC models, respectively. Numbers on the top of each panel show the top-five ranks of the 20 models according to the average of $|\hat{\alpha}_t - \alpha|$, $t \in \{505, 510, 515, \ldots, 3780\}$ for each model in the corresponding panel and the color of this number corresponds to the respective box-plot.

We clearly see from Fig. 13 that the OPAC and HOPAC estimators outperform all the remaining estimators in almost all cases, almost independently of the type of the generator family. Moreover, this implies that the OP transformation consistently improves the non-OP (H)AC estimators. In particular, the OP transformation is crucial for families that are unable to model upper-tail dependence, such as Ali–Mikhail–Haq, Clayton or Frank, or stronger forms of concordance (e.g., Ali–Mikhail–Haq). For the Joe family, we observe good and robust results also for the non-OP estimators. The

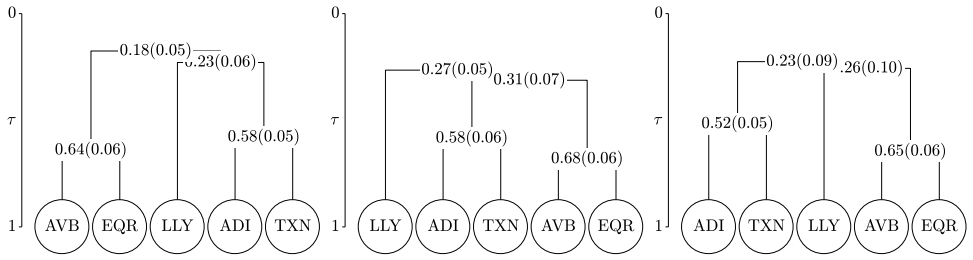


Fig. 14. All structures of the hierarchical models obtained for $w = 504$. The most often found one (left) with averaged Kendall's tau (standard deviation) over all the models with this structure, second most often (middle) and the least frequent (right).

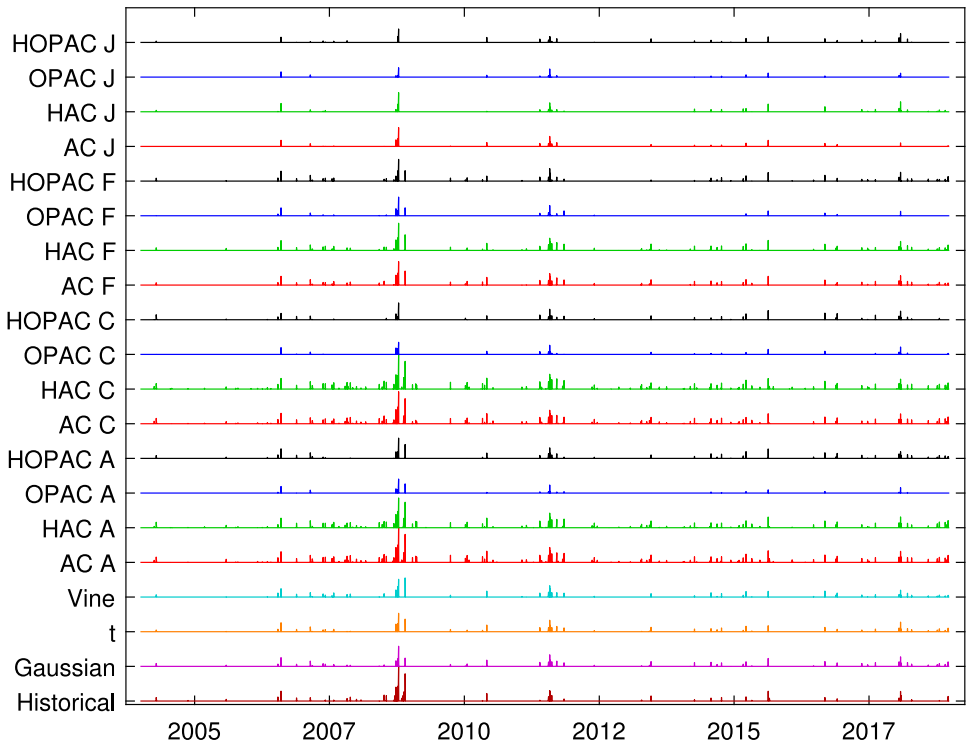


Fig. 15. Normalized violations of all the models considered for the five-dimensional portfolio with $w = 252$.

OP-based estimators more frequently outperform the benchmark estimators for bigger αs, e.g. there is no benchmark model in the top-five ranking for $\alpha = 0.99$. An improvement given by considering hierarchies (i.e., from OPAC to HOPAC) is observed mainly for $d = 10$ which may be explained by the fact that it becomes more important to model hierarchies in higher dimensions.

Structures of the hierarchical models fully support the industry clustering. There were only three structures for the five-dimensional portfolio found across all the windows with $w = 504$ and they are depicted in Fig. 14 according to their frequency order. The parameters provided in the structures are the averaged Kendall's tau values corresponding to each node over all the cases of the particular structure with the empirical standard deviations in parentheses. In all the three structures, ADI and TXN from the high-tech industry are together, AVB and EQR from the real estate industry are always joined and LLY from the pharmaceutical industry is connected slightly differently, but always separately from the other groups. Almost identical results are obtained for $w = 126$ and $w = 252$. For the 10-dimensional portfolio results are similar with a larger set of structures due to the higher dimensionality. Surprisingly, for $d = 5$ where the stocks possess strong hierarchical dependency structure and $\alpha = 0.95$, exchangeable OPACs provide better results than HOPACs. For the AC-based estimators, no substantial influence of the size of the time window (w) is observed which can be explained by the relative robustness of these models over time.

In-line with the findings from Figs. 13–14 we investigated the number of violations. In Fig. 15, we present violations for the five-dimensional portfolio normalized over all the models with the rolling window width $w = 252$. Note that the result for the 10-dimensional portfolio is very similar. All the models share comparably big violations around the global

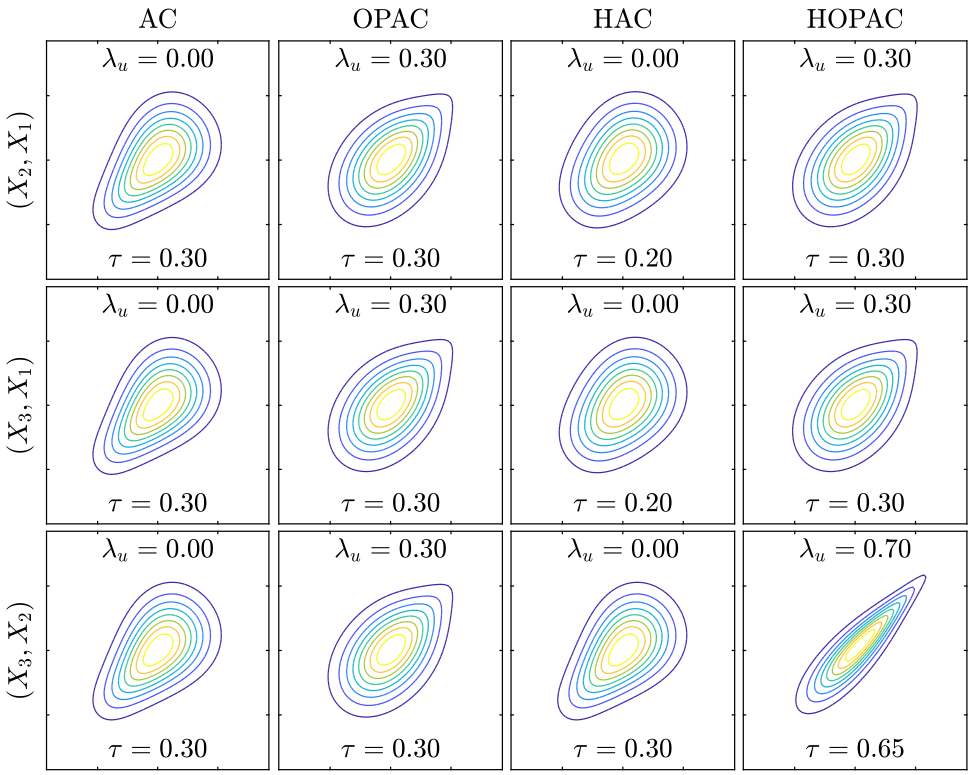


Fig. 16. Contour plots of the bivariate marginal densities of a random vector (X_1, X_2, X_3) with N(0, 1) margins and varying (across columns) Ali–Mikhail–Haq copulas. Note that τ and λ_u denote the corresponding pairwise Kendall's tau and upper-tail dependence coefficient, respectively.

financial crisis of 2008 and hit of the financial markets due to the earthquake and tsunamis in Japan in 2011. HAC and AC models show in general more violations not only during but also around the critical time periods. Similar to Fig. 13, all the models based on the Joe family have only few outliers and only in the economically critical times as 2008 and 2011.

6. Conclusion

We demonstrated the improvements OP transformations can bring to exchangeable ACs and hierarchical ACs. For the exchangeable case, a simplified way to compute the tail dependence coefficients was proposed. Also, the statistical properties of two feasible OPAC estimators were investigated by simulation. Furthermore, a new construction method, an efficient sampling strategy and an estimator were provided for HOPACs, including a simulation study confirming the feasibility of the proposed estimator. Excellent abilities of the (H)OPAC models were finally demonstrated in an application from risk management.

Note that there also exist other, more general transformations for ACs, e.g., the *tilted* OP transformation given by $\mathring{\psi}(t) = \psi\{(c^\beta + t)^{1/\beta} - c\}$, where $c \in [0, \infty)$, see Hofert (2011), or the *regularly varying transformed generator*, where the transform is given by a whole function, see Bernardino and Rulliére (2016). Their interpretation and tractability is, however, less clear, and the same applies to the hierarchical case. Also, all presented results based on Algorithm 3 use the same value 1.05 for $\beta_{\mathcal{R}}$, based on which the decision between restrictions $\mathcal{R}_1$ and $\mathcal{R}_2$ is made. This issue, however, needs more attention as it is not yet clear how much wrongly accepting $\mathcal{R}_1$ or alternatively $\mathcal{R}_2$ affects the overall fit of the HOPAC. The asymptotic treatment of the presented estimators is also an interesting topic for future research.

Acknowledgments

The authors would like to thank Maximilian Coblenz for providing them early versions of the MATVines toolbox. Also, the authors are grateful to the project No. CZ.02.2.69/0.0/0.0/16_027/0008521 "Support of International Mobility of Researchers at SU" which supports international cooperation.

Appendix A. Motivation

To illustrate the flexibility of the AC-based models, Fig. 16 shows contour plots of bivariate marginal densities of trivariate examples of ACs, OPACs, HACs and HOPACs based on Ali–Mikhail–Haq copulas with N(0, 1) margins.

In the first column, the parameter of the trivariate Ali–Mikhail–Haq copula is chosen such that Kendall's tau of all bivariate margins is 0.3. Such a choice provides a good fit in the model's body if we *observe* Kendall's tau close to this value in data of interest. Following the nature of Ali–Mikhail–Haq copulas, the dependence is clearly tail-asymmetric since for small joint values (found in the bottom-left corners of each contour plot) it is stronger than for large values (top-right corners). However, if we observe stronger dependence for large values in our data, Ali–Mikhail–Haq copulas do not allow to adjust to such a situation and thus leave to be discarded for modeling purposes. Involving the OP transformation, this adjustment is possible, as illustrated in the second column, where the extra parameter allows one to fine-tune the upper-tail dependence coefficient, e.g., to 0.3, while still *keeping* the same value of Kendall's tau. This allows one to build dependence models that fit well *both* in the body and tail, and, as we demonstrate in Section 5, just this transition from ACs to OPACs alone can already yield substantial improvements in tail-dependence modeling for risk management applications.

In contrast to ACs and OPACs (which are exchangeable and thus all their multivariate margins are the same), HACs and HOPACs provide asymmetry in the bivariate margins. For example, it can be seen from the last two columns that the margin corresponding to (X_2, X_3) differs (in τ) from the remaining two. Moreover, as Ali–Mikhail–Haq ACs are limited in τ to $[0, 1/3)$, the OP transformation allows one to attain any $\tau \in [0, 1)$, which makes (H)OPACs based on this family more suitable for modeling purposes. An Ali–Mikhail–Haq OPAC with $\tau \geq 1/3$ is, e.g., the bivariate margin corresponding to (X_2, X_3) in the fourth column. Similarly, as Ali–Mikhail–Haq ACs are tail independent, i.e., $\lambda_u = 0$ for all parameter values, their OP transformation can attain all λ_u in $[0, 1)$, which is illustrated by the same example. This extended flexibility in tail dependence modeling enabled the Ali–Mikhail–Haq HOPACs to gain the best results in several scenarios out of all copula models considered in a risk management application provided in Section 5.

Finally, note that an OP transformation establishes a connection between several one-parameter AC families from the list in Nelsen (2006, pp. 116–119). For example, the three families denoted 4.2.1 (Clayton), 4.2.12 and 4.2.14 are special cases of the OP Clayton family. To have these families in a single HAC, it was necessary to develop sampling and estimation strategies allowing to nest *different* families, which was done in Hofert (2011) and Górecki et al. (2017b), respectively. With regard to OP transformations, sampling and estimation of HACs involving these different one-parameter families can be addressed by considering hierarchical models involving just *one* but *OP-transformed* AC family, i.e., Clayton HOPACs.

Appendix B. Dependence measures, sampling and estimation of outer power Archimedean copulas

B.1. Dependence measures of outer power Archimedean copulas

As can be observed from Fig. 16, the OP transformation can have an impact on measures of association such as Kendall's tau (e.g., reaching beyond $\frac{1}{3}$ for the Ali–Mikhail–Haq family) or the tail dependence coefficients (reaching $\lambda_u > 0$ for the upper-tail independent Ali–Mikhail–Haq family). We now consider these measures of association in more detail for OPACs.

Given a one-parameter 2-AC $C_{\psi_{(a,\theta)}}$, there exists a functional relationship between the parameter θ and Kendall's tau which can sometimes be expressed in closed form, e.g., $\tau_{(\mathrm{C})}(\theta) = \theta/(\theta + 2)$ for the Clayton family. This relationship can easily be extended to OPACs. As follows from Proposition 3.7 in Hofert (2011), given a 2-OPAC $C_{\psi_{(a,\theta,\beta)}}$, its corresponding Kendall's tau $\tau_{(a)}(\theta, \beta)$ is

$$\tau_{(a)}(\theta, \beta) = 1 - \{1 - \tau_{(a)}(\theta)\}/\beta; \tag{14}$$

in particular $\tau_{(a)}(\theta, \beta) \geq \tau_{(a)}(\theta)$. We thus see how Kendall's tau of Ali–Mikhail–Haq copulas can cover the whole $[0, 1)$, while $\tau_{(\mathrm{A})}(\theta)$ only covers $\left[0, \frac{1}{3}\right)$. A similar result can be derived for the coefficients of tail dependence $\lambda_l(C)$ and $\lambda_u(C)$ of a 2-AC C under additional assumptions on $\psi_{(a,\theta)}$ (or $\psi'_{(a,\theta)}$) such as regular variation.

Proposition 3. *Let ψ be a generator of a 2-AC C_ψ and $\psi_\beta(t) = \psi(t^{\frac{1}{\beta}})$ for all $t \in [0, \infty)$ and $\beta \in [1, \infty)$. Then:*

1. *If ψ is regularly varying at infinity with index $\alpha_\infty \in \mathbb{R}$, i.e., $\lim_{t\to\infty} \frac{\psi(ct)}{\psi(t)} = c^{\alpha_\infty}$ for all $c \in (0, \infty)$, then $\lambda_l(C_{\psi_\beta}) = 2^{\frac{\alpha_\infty}{\beta}}$.*
2. *If ψ' is regularly varying at zero with index $\alpha_0 \in \mathbb{R}$, i.e., $\lim_{t\downarrow 0} \frac{\psi'(ct)}{\psi'(t)} = c^{\alpha_0}$ for all $c \in (0, \infty)$ then $\lambda_u(C_{\psi_\beta}) = 2 - 2^{\frac{\alpha_0+1}{\beta}}$.*

Proof.

1. Using (2.11) from Hofert (2010), $\lambda_l(C_{\psi_\beta}) = \lim_{t\to\infty} \frac{\psi_\beta(2t)}{\psi_\beta(t)} = \lim_{t\to\infty} \frac{\psi(2^{1/\beta}t^{1/\beta})}{\psi(t^{1/\beta})} = \lim_{s\to\infty} \frac{\psi(2^{1/\beta}s)}{\psi(s)}$, where $s = t^{1/\beta}$. If ψ is regularly varying at infinity with index α_∞, $c = 2^{1/\beta}$ establishes the proof.
2. Using (2.12) from Hofert (2010), $\lambda_u(C_{\psi_\beta}) = 2 - 2\lim_{t\downarrow 0} \frac{1-\psi_\beta(2t)}{1-\psi_\beta(t)} = 2 - \lim_{t\downarrow 0} \frac{1-\psi(2^{1/\beta}t^{1/\beta})}{\psi(t^{1/\beta})} = 2 - \lim_{s\downarrow 0} \frac{1-\psi(2^{1/\beta}s)}{1-\psi(s)}$, where $s = t^{1/\beta}$. Applying l'Hôpital's rule, $2 - \lim_{s\downarrow 0} \frac{1-\psi(2^{1/\beta}s)}{1-\psi(s)} = 2 - 2^{1/\beta} \lim_{s\downarrow 0} \frac{\psi'(2^{1/\beta}s)}{\psi'(s)}$. If ψ' is regularly varying at zero with index α_0, using $c = 2^{1/\beta}$ implies that $2 - 2^{1/\beta} \lim_{s\downarrow 0} \frac{\psi'(2^{1/\beta}s)}{\psi'(s)} = 2 - 2^{1/\beta}(2^{1/\beta})^{\alpha_0} = 2 - 2^{\frac{\alpha_0+1}{\beta}}$. □

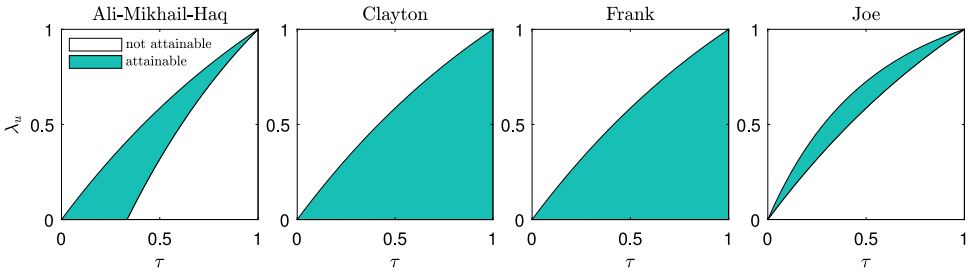


Fig. 17. Attainable pairs of τ and λ_u for four OPAC families.

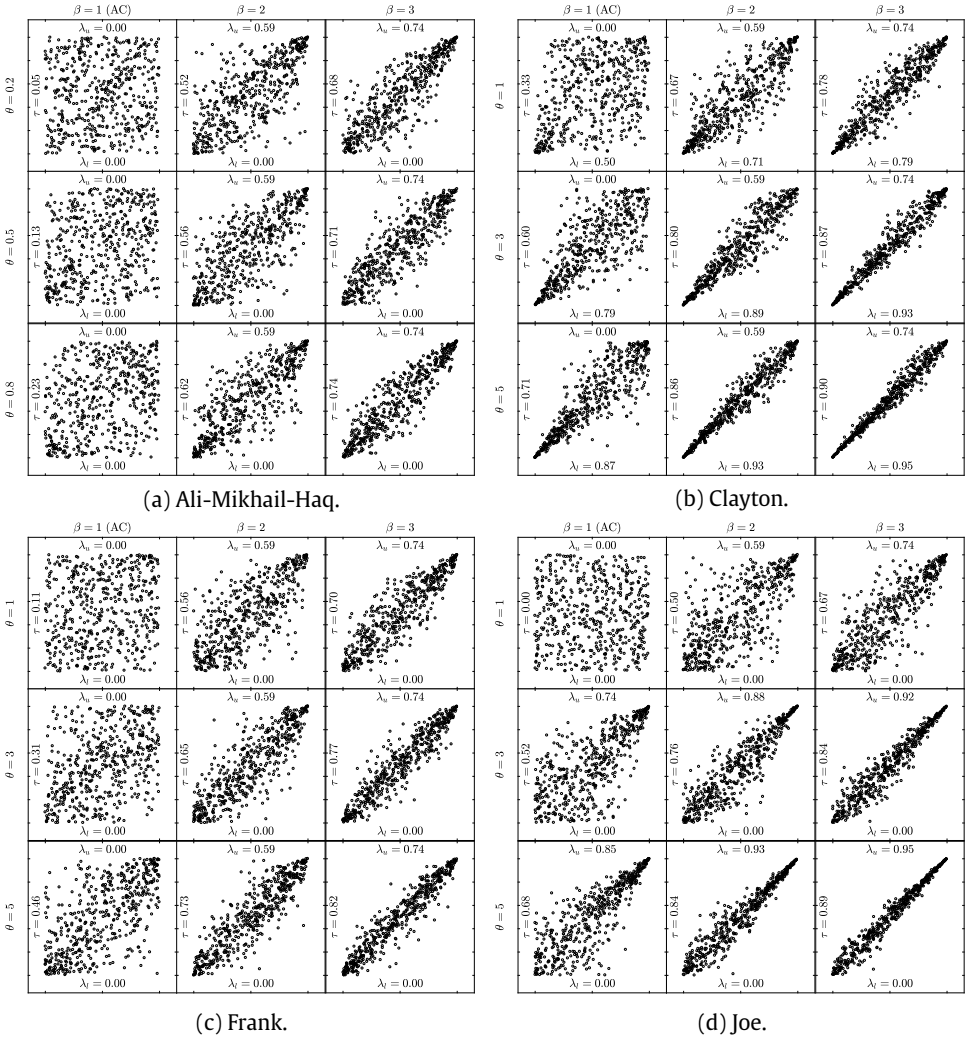
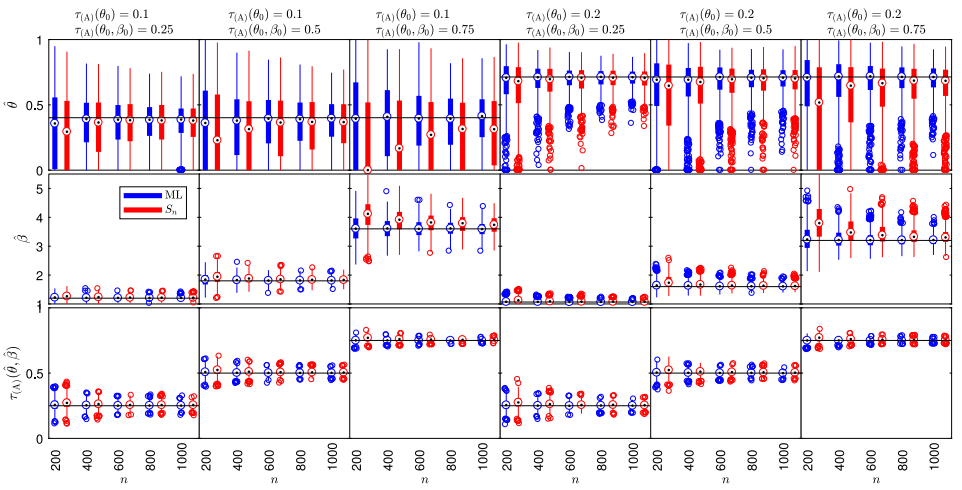
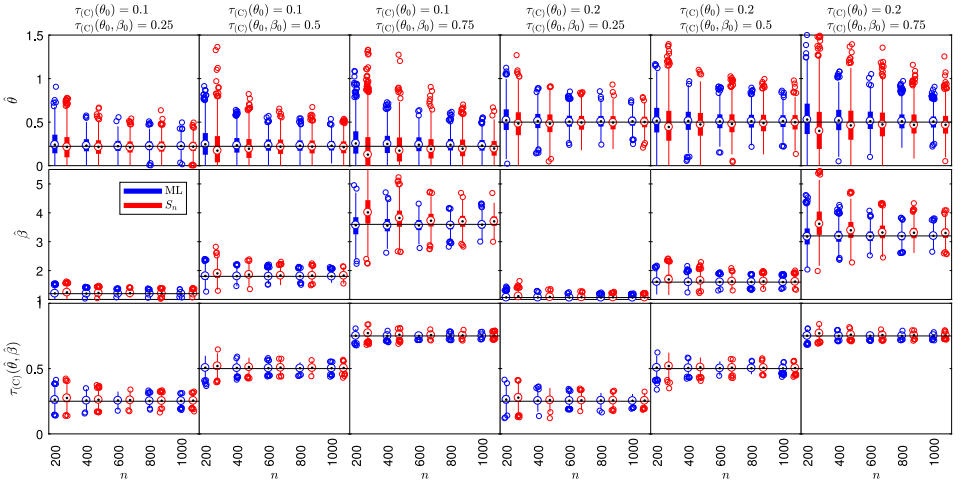


Fig. 18. Samples from OPACs $C_{\psi_{(a,\theta,\beta)}}$ with $n = 500$.

Having the index α_∞ (α_0) for a one-parameter family, Proposition 3 often provides λ_l and λ_u for its OP family. It can easily be verified that $\alpha_\infty = -\theta^{-1}$ for $a = \mathrm{C}$, whereas $\psi_{(a,\theta)}$ is not regularly varying at ∞ for $a \in \{\mathrm{A}, \mathrm{F}, \mathrm{J}\}$, and that $\alpha_0 = 1$ for $a \in \{\mathrm{A}, \mathrm{C}, \mathrm{F}\}$, whereas $\alpha_0 = \theta^{-1} - 1$ for $a = \mathrm{J}$; see also the last two columns of Table 1. These two columns also reveal that β influences λ_u for all listed families, whereas λ_l only for the Clayton family. From this point of view, β plays an important role particularly for upper-tail dependence modeling. Given a bivariate OPAC $C_{\psi_{(a,\theta,\beta)}}$ and also considering its Kendall's tau via (14), an interesting question is if, given $(\tau, \lambda_u) \in [0, 1]^2$, there exist values of θ and β such that



(a) Ali-Mikhail-Haq.



(b) Clayton.

Fig. 19. Results of the simulation study for the estimators ML and S_n of the Archimedean families A and C. The black line in each plot shows θ_0, β_0 or $\tau_a(\theta_0, \beta_0)$, respectively.

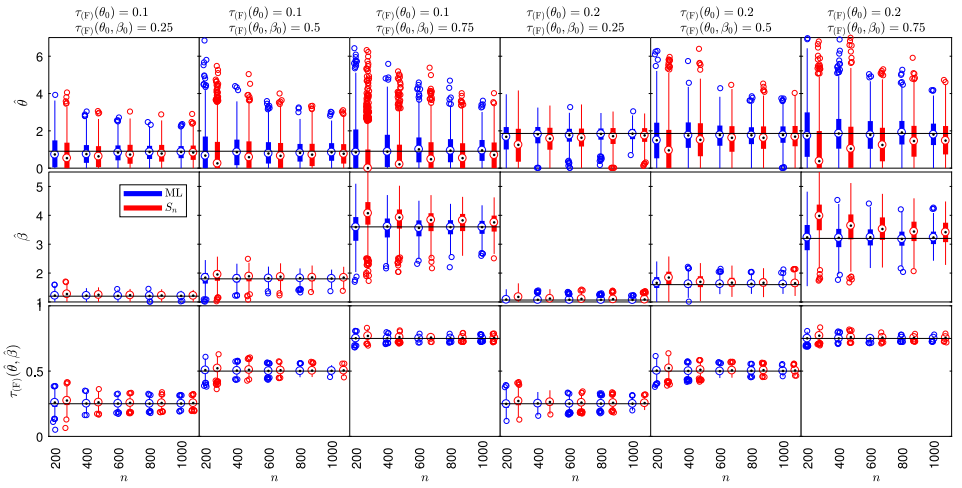
$\tau_{(a)}(\theta, \beta) = \tau$ and $\lambda_u(C_{\psi_{(a,\theta,\beta)}}) = \lambda_u$. Fig. 17 visually highlights the pairs of τ and λ_u for which such θ and β exist, so which τ and λ_u are *attainable*. We observe a larger region of attainable pairs for the upper-tail independent AC families of Ali–Mikhail–Haq, Clayton and Frank; for the Joe family, $\lambda_u > 0$ even with $\beta = 1$. Hence, fixing τ to a desired value in order to obtain a good fit in an OPAC's body, the most flexibility in the upper tail is provided by, somewhat unexpected, the upper-tail *independent* AC families.

Finally note that all measures of association considered are monotone with respect to β, which, for λ_l and λ_u, follows from Proposition 3.7 in Hofert (2011). This is also clearly visible in Fig. 18, where samples of size $n = 500$ from a bivariate OPAC with different values of θ and β are shown for each of our working families considered.

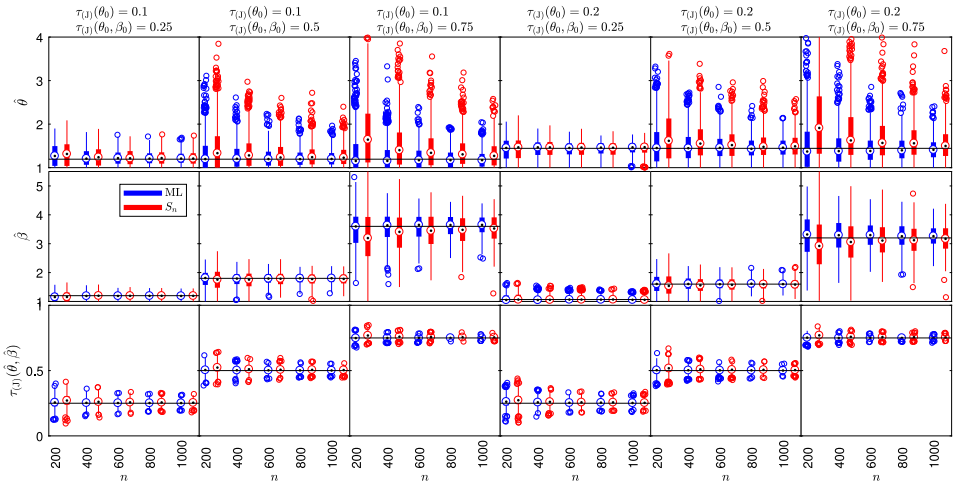
B.2. Sampling from OPACs

The samples in Fig. 18 were obtained using the Marshall–Olkin algorithm together with Theorem 3.6 from Hofert (2011); for the sake of completeness, we recall the latter below as it is needed later in Section 3.2. Note that $\mathcal{LS}^{-1}[\psi]$ denotes the inverse Laplace–Stieltjes transform of ψ.

Theorem 1 (*Hofert, 2011*)**.** *Let $\beta \in [1, \infty)$, ψ be a c.m. generator and ψ_β be its OP transformation given by (2). Then $\mathring{V} := SV^\beta \sim \mathring{F} := \mathcal{LS}^{-1}[\psi_\beta]$, where $V \sim \mathcal{LS}^{-1}[\psi]$ and $S \sim S\{1/\beta, 1, \cos^\beta(\pi/2\beta), \mathbb{1}_{\{\beta=1\}}; 1\}$ ($1/\beta$-stable distribution).*



(a) Frank.



(b) Joe.

Fig. 20. Results of the simulation study for the estimators ML and S_n of the Archimedean families A and C. The black line in each plot shows θ_0, β_0 or $\tau_a(\theta_0, \beta_0)$, respectively.

As sampling from S is a standard routine, sampling from an OPAC only requires a sampling strategy for $\mathcal{LS}^{-1}[\psi]$, which is known for many one-parameter AC families; see Hofert (2010, Table 2.1).

B.3. Estimating OPACs

For the one-parameter case, two AC estimators are particularly popular, the 1) maximum likelihood (ML) estimator and the 2) Kendall's tau inverse estimator; see Genest and Rivest (1993). The latter can be viewed as a method-of-moments-like estimator and is statistically not as efficient as ML. In contrast, the ML estimator naturally extends to any parameter dimension and is also feasible for estimating OPACs. To complement this estimator, we also consider a distance-based estimator (based on the goodness-of-fit test statistic S_n of Genest and Favre (2007)) in what follows.

The ML estimator for the parameters θ and β of a d-OPAC $C_{\psi(a,\theta,\beta)}$ with family a is defined by

$$(\hat{\theta}_{\mathrm{ML}}, \hat{\beta}_{\mathrm{ML}}) = \operatorname*{argmax}_{(\theta \in \Theta_a, \beta \geq 1)} \sum_{i=1}^{n} \log c_{\psi(a,\theta,\beta)}(\boldsymbol{u}_i), \tag{15}$$

and a distance-based estimator

$$(\hat{\theta}_{S_n}, \hat{\beta}_{S_n}) = \operatorname*{argmin}_{(\theta \in \Theta_a, \beta \geq 1)} \sum_{i=1}^{n} \{C_{\psi(a,\theta,\beta)}(\boldsymbol{u}_i) - C_n(\boldsymbol{u}_i)\}^2, \tag{16}$$

where $c_{\psi_{(a,\theta,\beta)}}$ is the density of $C_{\psi_{(a,\theta,\beta)}}$, $C_n(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{\boldsymbol{u}_i \le \boldsymbol{u}\}}$ is the empirical copula of a sample of (pseudo-)observations $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{id}) \in [0,1]^d$, $i \in \{1, \ldots, n\}$, and $n \in \mathbb{N}$ for all $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$.

In the following simulation study we compare these two estimators, where all estimates are replicated $N = 1000$ times for sample sizes $n = 200, 400, \ldots, 1000$, and 6 OPAC models $C_{\psi_{(a,\theta_0,\beta_0)}}$ with

- θ_0 chosen such that $\tau_{(a)}(\theta_0) \in \{0.1, 0.2\}$, and
- given θ_0, β_0 is set such that $\tau_{(a)}(\theta_0, \beta_0) \in \{0.25, 0.5, 0.75\}$.

Based on the results for $a \in \{A, C, F, J\}$ shown in Figs. 19 and 20, we conclude that:

1. Both estimators converge to the true values (θ_0 and β_0) with increasing n;
2. The ML estimator is unbiased and more efficient than S_n, which is expected from classical statistical estimation theory;
3. The standard errors of $\hat{\theta}$ and $\hat{\beta}$ increase with $\tau_{(a)}(\theta_0, \beta_0)$, whereas for $\tau_{(a)}(\hat{\theta}, \hat{\beta})$ they decrease; and
4. Conclusions 1.-3. are independent of the family a considered.

To summarize, both estimators are viable for OPAC estimation. We thus use these estimators also for HOPAC estimation considered in Sections 3.3 and 4.

References

Batagelj, V., 1981. Note on ultrametric hierarchical clustering algorithms. Psychometrika 46 (3), 351–352. http://dx.doi.org/10.1007/BF02293743.

Bernardino, E.D., Rulliére, D., 2016. On tail dependence coefficients of transformed multivariate archimedean copulas. Fuzzy Sets and Systems 284, 89–112. http://dx.doi.org/10.1016/j.fss.2015.08.030, Theme: Uncertainty and Copulas.

Bernstein, S., 1929. Sur les fonctions absolument monotones. Acta Math. 52 (1), 1–66.

Capéraà, P., Fougères, A.-L., Genest, C., 2000. Bivariate distributions with given extreme value distributions. J. Multivariate Anal. 72, 30–49.

Charpentier, A., Fougères, A.-L., Genest, C., Nešlehová, J.G., 2014. Multivariate archimax copulas. J. Multivariate Anal. 126, 118–136.

Coblenz, M., 2019. MATVines: A vine copula toolbox for MATLAB. URL: https://sites.google.com/view/maximiliancoblenz/startseite.

Cossette, H., Gadoury, S.-P., Marceau, E., Robert, C.Y., 2019. Composite likelihood estimation method for hierarchical archimedean copulas defined with multivariate compound distributions. J. Multivariate Anal. 172, 59–83. http://dx.doi.org/10.1016/j.jmva.2019.03.008, Dependence Models.

Czado, C., 2010. Pair-copula constructions of multivariate copulas. In: Jaworski, P., Durante, F., Hardle, W.K., Rychlik, T. (Eds.), Copula Theory and Its Applications. In: Lecture Notes in Statistics, vol. 198, Springer Berlin Heidelberg, pp. 93–109.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?. Rev. Financ. Stud. 22 (5), 1915–1953. http://dx.doi.org/10.1093/rfs/hhm075.

Genest, C., Favre, A., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. Hydrol. Eng. 12, 347–368.

Genest, C., Rémillard, B., Beaudoin, D., 2009. Goodness-of-fit tests for copulas: A review and a power study. Insurance Math. Econom. 44 (2), 199–213. http://dx.doi.org/10.1016/j.insmatheco.2007.10.005.

Genest, C., Rivest, L.-P., 1993. Statistical inference procedures for bivariate archimedean copulas. J. Amer. Statist. Assoc. 88 (423), 1034–1043.

Górecki, J., Hofert, M., Holeňa, M., 2016. An approach to structure determination and estimation of hierarchical archimedean copulas and its application to Bayesian classification. J. Intell. Inf. Syst. 46 (1), 21–59. http://dx.doi.org/10.1007/s10844-014-0350-3.

Górecki, J., Hofert, M., Holeňa, M., 2017a. Kendall's tau and agglomerative clustering for structure determination of hierarchical archimedean copulas. Depend. Model. 5 (1), 75–87. http://dx.doi.org/10.1515/demo-2017-0005.

Górecki, J., Hofert, M., Holeňa, M., 2017b. On structure, family and parameter estimation of hierarchical archimedean copulas. J. Stat. Comput. Simul. 87 (17), 3261–3324. http://dx.doi.org/10.1080/00949655.2017.1365148.

Górecki, J., Hofert, M., Holeňa, M., 2020. Hierarchical archimedean copulas for MATLAB and OCTAVE: The **HACopula** toolbox. J. Stat. Softw. 93 (10), 1–36.

Hofert, M., 2010. Sampling Nested Archimedean Copulas with Applications to CDO Pricing. Universität Ulm, http://dx.doi.org/10.18725/OPARU-1787.

Hofert, M., 2011. Efficiently sampling nested archimedean copulas. Comput. Statist. Data Anal. 55 (1), 57–70. http://dx.doi.org/10.1016/j.csda.2010.04.025.

Hofert, M., 2012. A stochastic representation and sampling algorithm for nested archimedean copulas. J. Stat. Comput. Simul. 82 (9), 1239–1255. http://dx.doi.org/10.1080/00949655.2011.574632.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2017. Copula: Multivariate dependence with copulas. URL: https://CRAN.R-project.org/package=copula R package version 0.999-16.

Hofert, M., Mächler, M., McNeil, A.J., 2013. Archimedean copulas in high dimensions: Estimators and numerical challenges motivated by financial applications. J.Soc. Franç. Statist. 154 (1), 25–63.

Hofert, M., Scherer, M., 2011. CDO pricing with nested archimedean copulas. Quant. Finance 11 (5), 775–787.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London.

Joe, H., 2014. Dependence Modeling with Copulas. CRC Press.

Joe, H., Kurowicka, D., 2011. Dependence Modeling: Vine Copula Handbook. World Scientific.

Kimberling, C.H., 1974. A probabilistic interpretation of complete monotonicity. Aequationes Math. 10, 152–164. http://dx.doi.org/10.1007/BF01832852.

Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. J. Deriv. 3, 73–84.

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

Makalic, E., Schmidt, D.F., 2018. An efficient algorithm for sampling from $sim^k(x)$ for generating random correlation matrices. arXiv preprint arXiv:1809.05212.

Marshall, A.W., Olkin, I., 1988. Families of multivariate distributions. J. Amer. Statist. Assoc. 83 (403), 834–841.

McNeil, A.J., 2008. Sampling nested archimedean copulas. J. Stat. Comput. Simul. 78 (6), 567–581. http://dx.doi.org/10.1080/00949650701255834.

McNeil, A., Frey, R., Embrechts, P., 2015. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton university press.

McNeil, A.J., Nešlehová, J., 2009. Multivariate archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. Ann. Statist. 37, 3059–3097.

Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer-Verlag.

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

Okhrin, O., Okhrin, Y., Schmid, W., 2013b. Properties of hierarchical archimedean copulas. Statist. Risk Model. 30 (1), 21–54. http://dx.doi.org/10.1524/strm.2013.1071.
Ressel, P., 2013. Homogeneous distributions – and a spectral representation of classical mean values and stable tail dependence functions. J. Multivariate Anal. 117, 246–256.
Rezapour, M., 2015. On the construction of nested archimedean copulas for d-monotone generators. Statist. Probab. Lett. 101, 21–32. http://dx.doi.org/10.1016/j.spl.2015.03.001.
Schmidt, R., Stadtmüller, U., 2006. Non-parametric estimation of tail dependence. Scand. J. Stat. 33 (2), 307–335.
Segers, J., Uyttendaele, N., 2014. Nonparametric estimation of the tree structure of a nested archimedean copula. Comput. Statist. Data Anal. 72, 190–204.
Uyttendaele, N., 2017. On the estimation of nested archimedean copulas: a theoretical and an experimental comparison. Comput. Statist. http://dx.doi.org/10.1007/s00180-017-0743-1.
Zhu, W., Wang, C.-W., Tan, K.S., 2016. Structure and estimation of Lévy subordinated hierarchical archimedean copulas (LSHAC): Theory and empirical tests. J. Bank. Financ. 69, 20–36.