Pair-copula constructions for non-Gaussian DAG models[*]

Alexander Bauer[†‡]   Claudia Czado[†]   Thomas Klein[†]

Abstract. We propose a new type of multivariate statistical model that permits non-Gaussian distributions as well as the inclusion of conditional independence assumptions specified by a directed acyclic graph. These models feature a specific factorisation of the likelihood that is based on pair-copula constructions and hence involves only univariate distributions and bivariate copulas, of which some may be conditional. We demonstrate maximum-likelihood estimation of the parameters of such models and compare them to various competing models from the literature. A simulation study investigates the effects of model misspecification and highlights the need for non-Gaussian conditional independence models. The proposed methods are finally applied to modeling financial return data.

Key words: Bayesian networks, conditional independence, copulas, graphical models, likelihood inference, regular vines.

1 Introduction

Graphical models are multivariate statistical models in which the corresponding joint distribution of a family of random variables is restricted by a list of conditional independence assumptions. This list is conveniently summarised in a graph whose vertices represent the variables and whose edges represent interrelations of these variables. Lauritzen (1996) and Cowell et al. (2003) are standard references on the theory of graphical models but are mainly limited to the assumption of joint normality as far as continuous variables are concerned. At the same time, it is well known from the literature on statistical models for financial markets that the assumption

---

of joint normality may lead to severe underestimation of certain risks and, in a more general sense, fails to yield suitable models in many applications, see, for instance, McNeil et al. (2005).

We hence propose a new type of statistical model based on generally non-Gaussian distributions which, by construction, satisfy conditional independence assumptions induced by a directed acyclic graph (DAG). This combination of features is achieved by using so-called pair-copula constructions (PCCs) in which a multivariate distribution is decomposed into bivariate, potentially conditional distributions based on iterated applications of Sklar's theorem on copulas. The basic idea of applying PCCs to distributions with certain conditional independence properties goes back to Hanea et al. (2006) and Kurowicka and Cooke (2006). We follow these authors' approach of utilising PCCs to specifically construct non-Gaussian distributions in order to capture features such as tail behaviour and non-linear, asymmetric dependence. This approach has various benefits: First, as is customary in applications of copula models, we may conveniently separate the tasks of modeling univariate margins and multivariate dependences. Second, since we need not limit ourselves to Gaussian margins and copulas and since univariate marginal distributions and bivariate copulas may be freely combined, we are able to capture certain distributional properties to be modeled, for instance, heavy-tailedness and tail dependence as observed in financial data. Third, the building blocks of PCCs are bivariate copulas, even though we model higher-dimensional distributions. In particular, we may draw from the rich literature on bivariate copula families, see, for example, Joe (1997). Hanea et al. (2006) rely on elicited expert knowledge to construct multivariate distributions using a PCC approach. By comparison our work is focused on data-driven parametric inference.

PCCs were first proposed by Joe (1996) and further extended by Bedford and Cooke (2001, 2002) who developed so-called regular vines as a graphical representation of a class of hierarchical PCC models. Aas et al. (2009) later recognized these models' aptitude for likelihood inference since densities of PCC distributions are easily obtainable in explicit analytical form. Applications to financial data have shown that these vine-PCC models outperform other multivariate copula models in predicting log-returns of equity portfolios, see Aas and Berg (2009), Chollete et al. (2009), Fischer et al. (2009), and Czado et al. (2011). Min and Czado (2010, 2011) demonstrate that vine-PCC models also lend themselves to Bayesian inference. Elidan (2010a,b) gives another copula decomposition of distributions associated with a DAG that is based on generally higher-variate copulas and therefore lacks the flexibility of the pair-copula approach. A concept crucial to PCCs are conditional copulas. Outside the PCC context conditional copulas have been used in finance applications, frequently with the aim of modeling time-varying dependence, see Cherubini et al. (2004, Section 5.9) and Patton (2006). Here time variation in the conditional copulas was captured through the inclusion of time-varying parameters. Hobæk Haff et al. (2010) investigate conditional bivariate distributions in which the conditioning values enter the conditional margins but not the conditional copula, a customary assumption in PCC modeling.

The flexibility of vine-PCC models comes at the price of exponential growth in the number of pair copulas to be specified as the number of variables increases. We show that by capturing conditional independences present in the data, our closely related DAG-PCC approach yields more parsimonious models in many settings.

The paper is organised as follows. In Section 2 we give a short review of PCCs and regular vines. Section 3 shows how PCCs can be used to obtain multivariate distributions with Markov properties given by a directed acyclic graph. Based on this idea we construct so-called DAG-PCC models whose aptitude for likelihood inference is explored in a simulation study in Section 4. Section 5 presents an application of DAG-PCC models to financial data, and the paper concludes with a brief discussion in Section 6.

2 Pair-copula constructions and regular vines

A copula is a multivariate cumulative distribution function (cdf) $C \colon [0,1]^d \to [0,1]$, $d \in \mathbb{N}$, such that all univariate marginals are uniform on the interval $[0,1]$. By Sklar's theorem (Sklar, 1959) every cdf $F \colon \mathbb{R}^d \to [0,1]$ with univariate marginals $F_1, \ldots, F_d$ may be written as

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big) \tag{2.1}$$

for some suitable copula C and all $x_1, \ldots, x_d \in \mathbb{R}$. If F is absolutely continuous and $F_1, \ldots, F_d$ are strictly increasing we can pass to its probability density function (pdf) and write

$$f(x_1, \ldots, x_d) = c\big(F_1(x_1), \ldots, F_d(x_d)\big) \prod_{i=1}^{d} f_i(x_i), \tag{2.2}$$

where the copula pdf c is uniquely determined. Equations (2.1) and (2.2) can be solved for C and c, respectively, using marginal quantile functions. Doing so we obtain

$$C(u_1, \ldots, u_d) = F\big(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\big)$$

and

$$c(u_1, \ldots, u_d) = \frac{f\big(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\big)}{\prod_{i=1}^{d} f_i\big(F_i^{-1}(u_i)\big)}$$

for all $u_1, \ldots, u_d \in [0,1]$. Various examples of copulas together with the underlying theory are presented in Joe (1997) and Nelsen (2006). We will restrict our considerations to cdfs with the above-mentioned properties.

While there is a plethora of literature on bivariate copula families (also called *pair-copula families*), the range of higher-variate copula families is rather limited, see Joe (1997, Chapter 4).

Many popular bivariate copulas have no straightforward multivariate extension. Based on work of Joe (1996), Bedford and Cooke (2001, 2002) therefore proposed a flexible way of constructing multivariate copulas that uses (conditional) pair copulas as building blocks only. The core of their approach is a graphical representation called a *regular vine* that consists of a sequence of trees, each edge of which is associated with a certain pair copula. We briefly review the idea behind such pair-copula constructions (PCCs) using the example of a D-vine, which is one of the most popular types of regular vines, see Figure 1.
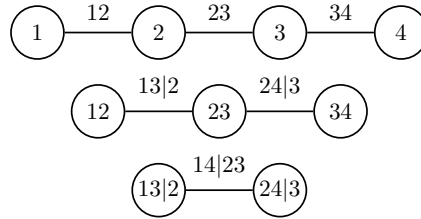

Figure 1: A four-variate D-vine specifying the pair copulas C_{12}, C_{23}, C_{34}, $C_{13|2}$, $C_{24|3}$, and $C_{14|23}$.

Let F be the cdf of a D-vine PCC on $\mathbb{R}^d$ and let $I = \{1\ldots,d\}$. The first tree (or *level*) of the D-vine comprises the d nodes $i \in I$ which represent the univariate margins F_i of F. These nodes are joined by the $d-1$ edges $(i-1, i)$, $i \in I \setminus \{1\}$, such that every node has at most two neighbours. The edge labels $(i-1, i)$ (displayed without parentheses and commas in Figure 1) represent the unconditional pair copulas $C_{(i-1),i}$ used in the PCC. Each subsequent tree of the D-vine is then derived from its predecessor by turning all edges into nodes and by introducing a new edge whenever two nodes share all but two indices. Those two indices form the conditioned set and the remaining ones the conditioning set of the associated pair copula, as denoted by the respective edge label. The edges of the second tree, for instance, denote the conditional pair copulas $C_{i,(i+2)|(i+1)}$, $i \in I \setminus \{d-1, d\}$. Altogether, the D-vine consists of $d-1$ trees with $\binom{d}{2}$ edges. As pointed out in Aas et al. (2009) the pdf of F is given by

$$f(\boldsymbol{x}) = \prod_{i=1}^{d-1}\prod_{j=1}^{d-i} c_{j,j+i|(j,j+i)}\big(F_{j|(j,j+i)}(x_j \,|\, \boldsymbol{x}_{(j,j+i)}), F_{j+i|(j,j+i)}(x_{j+i} \,|\, \boldsymbol{x}_{(j,j+i)}) \,\big|\, \boldsymbol{x}_{(j,j+i)}\big) \prod_{i=1}^{d} f_i(x_i).$$

Here we have written $\boldsymbol{x}_{(j,j+i)} := (x_{j+1}, \ldots, x_{j+i-1})$ for all $i \leq d-1$ and $j \leq d-i$. More generally we will write $\boldsymbol{x}_J := (x_j)_{j \in J}$ for all $J \subseteq I$. Also, we have denoted the conditional cdf of X_j given $\boldsymbol{X}_{(j,j+i)} = \boldsymbol{x}_{(j,j+i)}$ by $F_{j|(j,j+i)}(\,\cdot\,|\,\boldsymbol{x}_{(j,j+i)})$, where $\boldsymbol{X} = (X_1, \ldots, X_d)$ is distributed as F.

According to Joe (1996) the conditional cdfs $F_{j|K}$, $j \in I$, $K \subseteq I \setminus \{j\}$, may be computed using the recursive formula

$$F_{j|K}(x_j \,|\, \boldsymbol{x}_K) = \frac{\partial C_{jk|(K\setminus\{k\})}\big(F_{j|(K\setminus\{k\})}(x_j \,|\, \boldsymbol{x}_{(K\setminus\{k\})}), F_{k|(K\setminus\{k\})}(x_k \,|\, \boldsymbol{x}_{(K\setminus\{k\})}) \,\big|\, \boldsymbol{x}_{(K\setminus\{k\})}\big)}{\partial F_{k|(K\setminus\{k\})}(x_k \,|\, \boldsymbol{x}_{(K\setminus\{k\})})}$$

$$(2.3)$$

for some $k \in K$. We can thus iteratively compute the values of $F_{j|(j,j+i)}$ and $F_{j+i|(j,j+i)}$ in tree i of our D-vine by choosing $k = j + i - 1$ and $k = j + 1$, respectively. Note that the only copulas needed in this computation are the ones already specified in the preceding trees.

We can construct a multitude of multivariate D-vine copulas by selecting a number of pair copulas and by setting all univariate marginals to be uniform distributions on $[0, 1]$. To ease notational burden we then express conditional cdfs in terms of so-called h-functions defined as

$$h_{ij}(u_i, u_j) := F_{i|j}(u_i \,|\, u_j) = \frac{\partial C_{ij}(u_i, u_j)}{\partial u_j} \quad \text{and} \quad h_{i\underline{j}}(u_i, u_j) := F_{j|i}(u_j \,|\, u_i) = \frac{\partial C_{ij}(u_i, u_j)}{\partial u_i} \quad (2.4)$$

for all $i \neq j \in I$. Many popular copulas exhibit closed form expressions for these partial derivatives, see Aas et al. (2009). A more detailed exposition of D-vines can be found in Kurowicka and Cooke (2006, Section 4.4) and Kurowicka and Joe (2011).

Although very convenient a model, the flexibility of regular vines comes at a price. The construction of a d-variate vine copula requires the specification of $\binom{d}{2}$ pair copulas, a number increasing quadratically in d. The actual number of decisions to make in practical applications might, however, be lower if the analysed data exhibit conditional independences. In that case the corresponding pair copulas are nothing but product copulas with pdf equal to one.

Instead of starting one's analysis with a set of regular vines it may therefore be more fruitful to look for conditional independences first. Finding a vine copula that satisfies a given set of conditional independence assumptions is in general, however, a hard problem. One class of models tailor-made for this task are Bayesian networks. By applying the pair-copula concept to graphical models, Hanea et al. (2006) provided an opportunity to exploit the advantages of both worlds. However, their analysis is restricted to pair copula families with the property that zero rank correlation implies independence. We will review PCCs for directed graphs in the next section.

3 Pair-copula constructions and directed acyclic graphs

Let $V \neq \emptyset$ be a finite set and let E be a subset of $\{(v, w) \in V^2 \,|\, v \neq w\}$ such that $(w, v) \notin E$ whenever $(v, w) \in E$. Then $D = (V, E)$ is a directed graph with vertex set V and edge set E. We denote a pair $(v, w) \in E$ by an arrow $v \to w$. A path in D is a sequence $v_1, \ldots, v_n \in V$, $n \geq 2$, such that D contains all arrows $v_i \to v_{i+1}$, $i \leq n - 1$. In the special case $v_1 = v_n$ the path $v_1, \ldots, v_n$ is a cycle. If there are no cycles in D, then D is called a *directed acyclic graph*

(DAG). We set

$$\text{pa}(v) := \{w \in V \mid D \text{ contains } w \to v\} \qquad \text{(set of \textit{parents} of } v),$$
$$\text{de}(v) := \{w \in V \mid D \text{ contains a path from } v \text{ to } w\} \qquad \text{(set of \textit{descendants} of } v), \text{ and}$$
$$\text{nd}(v) := V \setminus \big(\text{de}(v) \cup \{v\}\big) \qquad \text{(set of \textit{non-descendants} of } v)$$

for all $v \in V$.

Now let P be a probability measure on $\mathbb{R}^{|V|}$ and let P_I denote its I-margin for all non-empty $I \subseteq V$. Furthermore, let $\boldsymbol{X}$ be an $\mathbb{R}^{|V|}$-valued random variable distributed as P. Then clearly $\boldsymbol{X}_I \sim P_I$. We will only consider those probability measures P that can be associated with some DAG $D = (V, E)$ via certain conditional independence properties. Let us therefore write $\boldsymbol{X}_I \perp\!\!\!\perp \boldsymbol{X}_J \mid \boldsymbol{X}_K$ whenever $\boldsymbol{X}_I$ and $\boldsymbol{X}_J$ are conditionally independent given $\boldsymbol{X}_K$ for pairwise disjoint sets $I, J, K \subseteq V$.

P is said to be *D-Markovian* or, equivalently, to possess the *local D-Markov property* if

$$X_v \perp\!\!\!\perp \boldsymbol{X}_{\text{nd}(v) \setminus \text{pa}(v)} \mid \boldsymbol{X}_{\text{pa}(v)} \quad \text{for all } v \in V. \tag{3.1}$$

Note that P might exhibit further conditional independence properties. If P has a Lebesgue-density f, then the D-Markov property is equivalent to f admitting a *D-recursive factorisation* of the form

$$f(\boldsymbol{x}) = \prod_{v \in V} f_{v|\text{pa}(v)} \big(x_v \mid \boldsymbol{x}_{\text{pa}(v)}\big) \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^{|V|}, \tag{3.2}$$

where $f_{v|\text{pa}(v)}$ is the pdf of $P_{v|\text{pa}(v)}$. A *graphical model* based on D is a family of D-Markovian probability measures. For instance, the family of all regular D-Markovian normal distributions on $\mathbb{R}^{|V|}$ is called the *Gaussian graphical model* based on D. A comprehensive introduction to graphical models is found in Lauritzen (1996). Directed graphical models are also known as *Bayesian networks*, see Cowell et al. (2003, Section 2.10). Applications of Bayesian networks range from artificial intelligence, decision support systems, and engineering to genetics, geology, medicine, and finance, see also Pourret et al. (2008).

As an example consider the DAG from Figure 2 and an absolutely continuous probability measure P on $\mathbb{R}^4$ possessing the respective local D-Markov property. Straightforward evaluation of condition (3.1) yields the restrictions $X_1 \perp\!\!\!\perp X_\emptyset \mid X_\emptyset$ (for $v = 1$), $X_2 \perp\!\!\!\perp X_3 \mid X_1$ (both for $v = 2$ and $v = 3$), and $X_1 \perp\!\!\!\perp X_4 \mid \boldsymbol{X}_{23}$ (for $v = 4$), of which the first is vacuous. There are no other implicit conditional independence properties in this example. As for the pdf of P, equation (3.2) yields the representation

$$f(\boldsymbol{x}) = f_1(x_1)\, f_{2|1}(x_2 \mid x_1)\, f_{3|1}(x_3 \mid x_1)\, f_{4|23}(x_4 \mid \boldsymbol{x}_{23}) \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^4.$$
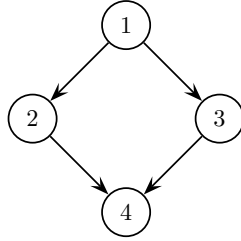
Figure 2: A DAG $D = (V, E)$ specifying the conditional independence properties $X_2 \perp\!\!\!\perp X_3 \mid X_1$ and $X_1 \perp\!\!\!\perp X_4 \mid \boldsymbol{X}_{23}$.

We shall mention that verifying whether a given empirical distribution on $\mathbb{R}^{|V|}$ can be assumed to originate from a D-Markovian probability measure is hard. One approach is to apply structure learning algorithms like the PC algorithm (Spirtes et al., 2000, Section 5.4.2) to the given data, see also Neapolitan (2003, Chapters 8 to 11), Koller and Friedman (2009, Chapter 18), and the discussion in Section 6. As an alternative approach expert knowledge is frequently exploited to define the graph D specifying the Markov structure, see Kurowicka and Cooke (2006, Chapter 5). However, both approaches are mainly confined to discrete or Gaussian modeling.

Straightforward application of Sklar's theorem to equation (3.2) yields a copula decomposition for the pdf f of a D-Markovian probability measure P on $\mathbb{R}^{|V|}$, see Elidan (2010a,b). This decomposition, however, consists of generally higher-variate copulas and hence leads to statistical models hampered by the difficulties described in Section 2.

We aim to derive a pair-copula decomposition for the pdf f of P. For every $v \in V$ we therefore order the elements of $\mathrm{pa}(v)$ increasingly (with respect to some strict total order $<_v$ on $\mathrm{pa}(v)$) and set

$$\mathrm{pa}(v; w) \coloneqq \left\{ u \in \mathrm{pa}(v) \,\middle|\, u <_v w \right\}, \quad w \in \mathrm{pa}(v).$$

Contrary to the vine-based approach of Hanea et al. (2006), the proof of our theorem relies on graph theoretical considerations only.

Theorem 3.1. *Let $D = (V, E)$ be a DAG and let P be an absolutely continuous D-Markovian probability measure whose univariate marginal cdfs are strictly increasing. Then P is uniquely determined by its univariate margins P_v, $v \in V$, and its (conditional) pair copulas $C_{vw|\mathrm{pa}(v;w)}$, $v \in V$, $w \in \mathrm{pa}(v)$.*

Proof. (Induction on the cardinality of V.) Since P is D-Markovian, its pdf f admits a D-recursive factorisation of the form (3.2). The claim is trivial for $|V| = 1$. Now let $|V| \geq 2$. Since D is acyclic we may choose some maximal vertex of D, that is, some $m \in V$ with $\mathrm{de}(m) = \emptyset$. Let $V' = V \setminus \{m\}$ and $E' = E \cap (V' \times V')$. Then above-mentioned factorisation can be written

as

$$f(\boldsymbol{x}) = f_{m|\mathrm{pa}(m)}\left(x_m \,\big|\, \boldsymbol{x}_{\mathrm{pa}(m)}\right) \prod_{v \in V'} f_{v|\mathrm{pa}(v)}\left(x_v \,\big|\, \boldsymbol{x}_{\mathrm{pa}(v)}\right), \quad \boldsymbol{x} \in \mathbb{R}^{|V|}. \tag{3.3}$$

By the choice of m the sets $\mathrm{pa}(v)$ and $\mathrm{nd}(v)$, $v \in V'$, remain unaffected by a transition from D to the subgraph $D' = (V', E')$. Hence, $P_{V'}$ is D'-Markovian and the product $\prod_{v \in V'} f_{v|\mathrm{pa}(v)}(\ldots)$ on the right hand side of (3.3) is the D'-recursive factorisation of $f_{V'}$. We may assume inductively that $P_{V'}$ and thus $f_{V'}$ are uniquely determined by the univariate margins P_v, $v \in V'$, and by the (conditional) pair copulas $C_{vw|\mathrm{pa}(v;w)}$, $v \in V'$, $w \in \mathrm{pa}(v)$. It remains to show that the same property holds for $f_{m|\mathrm{pa}(m)}$ if we include P_m and $C_{mw|\mathrm{pa}(m;w)}$, $w \in \mathrm{pa}(m)$, in our analysis.

We prove the latter claim by induction on $n = |\mathrm{pa}(m)|$. The claim is trivial for $n = 0$. In the case $n \geq 1$ let $w_1 <_m \ldots <_m w_n$ denote the elements of $\mathrm{pa}(m)$ and let $S := \mathrm{pa}(m; w_n)$. By Sklar's theorem,

$$f_{m|\mathrm{pa}(m)}\left(x_m \,\big|\, \boldsymbol{x}_{\mathrm{pa}(m)}\right) = c_{mw_n|S}\big(F_{m|S}(x_m \,|\, \boldsymbol{x}_S), F_{w_n|S}(x_{w_n} \,|\, \boldsymbol{x}_S) \,\big|\, \boldsymbol{x}_S\big)\, f_{m|S}\left(x_m \,|\, \boldsymbol{x}_S\right)$$

for all $x_m \in \mathbb{R}$ and $\boldsymbol{x}_{\mathrm{pa}(m)} \in \mathbb{R}^n$. Since $\mathrm{pa}(m) \subseteq V'$, the conditional cdf $F_{w_n|\mathrm{pa}(m;w_n)}$ is completely determined by $P_{V'}$ and therefore by the quantities specified in the theorem's claim. Observing that $\mathrm{pa}(m; w_n) = \mathrm{pa}(m) \setminus \{w_n\}$ we may conclude by induction that $F_{m|\mathrm{pa}(m;w_n)}$ and $f_{m|\mathrm{pa}(m;w_n)}$ exhibit the claimed property, too. This establishes the claim. $\qquad\square$

By Theorem 3.1 we can decompose f into

$$f(\boldsymbol{x}) = \prod_{v \in V} f_v(x_v) \prod_{w \in \mathrm{pa}(v)} c_{vw|\mathrm{pa}(v;w)}\big(F_{v|\mathrm{pa}(v;w)}(x_v \,|\, \boldsymbol{x}_{\mathrm{pa}(v;w)}), F_{w|\mathrm{pa}(v;w)}(x_w \,|\, \boldsymbol{x}_{\mathrm{pa}(v;w)}) \,\big|\, \boldsymbol{x}_{\mathrm{pa}(v;w)}\big),$$

$$\tag{3.4}$$

and thus again distinguish the pair copulas involved by the number of conditioning variables. Similarly to regular vines, the number of levels ranges from 0 to possibly $|V| - 1$. Although the graphical representations of the two models look fairly similar, the concepts behind are completely different. While regular vines illustrate the required pair copulas only, the arrows of a DAG specify conditional independence conditions. In both cases, however, these representations are visual aids only and can be omitted.

The remaining question is whether DAGs actually extend the set of pair-copula decompositions beyond regular vines. This question is closely related to the computation of the conditional cdfs $F_{v|\mathrm{pa}(v;w)}$ and $F_{w|\mathrm{pa}(v;w)}$ in equation (3.4). In contrast to vines, DAGs allow for the specification of conditional cdfs that cannot be computed by simply plugging in results from preceding levels. Hence the values of these functions have to be computed via integration of other margins. In view of their application for statistical inference, however, DAG PCCs may be more parsimonious than vine PCCs, see Section 4.

The lowest dimensional DAG models exhibiting this characteristic are the ones having the same Markov properties as the DAG in Figure 2. Using the ordering $2 <_4 3$ of the vertices $v \in \mathrm{pa}(4)$ we obtain $\mathrm{pa}(1; \emptyset) = \emptyset$, $\mathrm{pa}(2; 1) = \emptyset$, $\mathrm{pa}(3; 1) = \emptyset$, $\mathrm{pa}(4; 2) = \emptyset$, and $\mathrm{pa}(4; 3) = \{2\}$. Equation (3.4) therefore yields

$$
\begin{aligned}
f(\boldsymbol{x}) = {} & f_1(x_1) \cdots f_4(x_4) \cdot c_{12}\big(F_1(x_1), F_2(x_2)\big) \cdot c_{13}\big(F_1(x_1), F_3(x_3)\big) \cdot c_{24}\big(F_2(x_2), F_4(x_4)\big) \\
& \cdot c_{34|2}\big(F_{3|2}(x_3 \,|\, x_2), F_{4|2}(x_4 \,|\, x_2) \,|\, x_2\big), \quad \boldsymbol{x} \in \mathbb{R}^4,
\end{aligned} \tag{3.5}
$$

where by equation (2.3)

$$
F_{4|2}(x_4 \,|\, x_2) = \frac{\partial C_{24}\big(F_2(x_2), F_4(x_4)\big)}{\partial F_2(x_2)}.
$$

Since the copula C_{23} is not available in the decomposition of f we exploit the conditional independence property $X_2 \perp\!\!\!\perp X_3 \mid X_1$ to get

$$
\begin{aligned}
F_{3|2}(x_3 \,|\, x_2) &= \int_{-\infty}^{x_3} f_{3|2}(y_3 \,|\, x_2)\, \mathrm{d}y_3 \\
&= f_2^{-1}(x_2) \int_{-\infty}^{x_3} \int_{-\infty}^{\infty} f_{123}(y_1, x_2, y_3)\, \mathrm{d}y_1\, \mathrm{d}y_3 \\
&\stackrel{(3.2)}{=} f_2^{-1}(x_2) \int_{-\infty}^{\infty} f_1(y_1)\, f_{2|1}(x_2 \,|\, y_1) \int_{-\infty}^{x_3} f_{3|1}(y_3 \,|\, y_1)\, \mathrm{d}y_3\, \mathrm{d}y_1 \\
&= \int_{-\infty}^{\infty} f_1^{-1}(y_1)\, f_2^{-1}(x_2)\, f_{12}(y_1, x_2)\, F_{3|1}(x_3 \,|\, y_1)\, \mathrm{d}F_1(y_1) \\
&\stackrel{(2.3)}{=} \int_{-\infty}^{\infty} c_{12}\big(F_1(y_1), F_2(x_2)\big) \frac{\partial C_{13}\big(F_1(y_1), F_3(x_3)\big)}{\partial F_1(y_1)}\, \mathrm{d}F_1(y_1) \\
&= \int_0^1 c_{12}\big(u_1, F_2(x_2)\big) \frac{\partial C_{13}\big(u_1, F_3(x_3)\big)}{\partial u_1}\, \mathrm{d}u_1 \\
&\stackrel{(2.4)}{=} \int_0^1 c_{12}\big(u_1, F_2(x_2)\big)\, h_{1\underline{3}}\big(u_1, F_3(x_3)\big)\, \mathrm{d}u_1.
\end{aligned}
$$

There is no general closed-form solution for the last integral.

4 Likelihood inference: A simulation study

An appealing feature of DAG-PCC models is their flexibility gained from using bivariate copulas as building blocks only. In particular, these models can accomodate distributions other than the multivariate Gaussian, which is a desirable property in many statistical applications, see for example McNeil et al. (2005, Section 3.1.4). We now investigate the tractability of likelihood inference in DAG-PCC models. With reference to Sklar's theorem we restrict our considerations to *DAG copulas* in this section, that is, DAG distributions with uniform $[0, 1]$ univariate margins.

Let $D = (V, E)$ be a DAG and let $\{P_{\boldsymbol{\theta}} \,|\, \boldsymbol{\theta} \in \Theta\}$ be a family of D-Markovian probability measures

on $[0,1]^{|V|}$ satisfying the usual assumptions. Given a realisation $\boldsymbol{u} = (\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n) \in ([0,1]^{|V|})^n$, $n \in \mathbb{N}$, of a sample of i.i.d. random variables $\boldsymbol{U}^1, \ldots, \boldsymbol{U}^n$, equation (3.4) yields the log-likelihood function

$$L(\boldsymbol{\theta}; \boldsymbol{u}) = \sum_{k=1}^{n} \sum_{v \in V} \sum_{w \in \mathrm{pa}(v)} \log c_{vw|\mathrm{pa}(v;w)} \left(F_{v|\mathrm{pa}(v;w)} \left(u_v^k \,\Big|\, \boldsymbol{u}_{\mathrm{pa}(v;w)}^k ; \boldsymbol{\theta} \right), F_{w|\mathrm{pa}(v;w)} \left(u_w^k \,\Big|\, \boldsymbol{u}_{\mathrm{pa}(v;w)}^k ; \boldsymbol{\theta} \right); \boldsymbol{\theta} \right)$$

$$(4.1)$$

for all $\boldsymbol{\theta} = (\boldsymbol{\theta}_{vw|\mathrm{pa}(v;w)})_{v \in V, w \in \mathrm{pa}(v)} \in \Theta$. Note that the components $\boldsymbol{\theta}_{vw|\mathrm{pa}(v;w)}$ are potentially vector-valued. On the right hand side we have omitted parameter subscripts as well as the values of the conditioning variables in the pair-copula pdfs $c_{vw|\mathrm{pa}(v;w)}$. The latter omission is based on the assumption of constant copula parameters $\boldsymbol{\theta}_{vw|\mathrm{pa}(v;w)}$ for all observations $\boldsymbol{u}^k$, $k \leq n$. By that assumption the copula pdfs involved are treated as if they were unconditional. The cdfs $F_{v|\mathrm{pa}(v;w)}$ and $F_{w|\mathrm{pa}(v;w)}$, however, remain conditional. This assumption reduces model complexity while still encompassing a rich class of DAG copulas and has become common practice in likelihood inference for PCCs, see Aas et al. (2009) and Hobæk Haff et al. (2010).

Note that equation (3.4) also allows the inclusion of univariate marginal distributions other than the uniform $[0, 1]$ distribution. The joint estimation of the parameters of the marginal distributions and the copula can, however, become computationally intensive, especially in higher dimensions. Joe and Xu (1996) therefore proposed the *inference functions for margins* (IFM) method in which the estimation procedure is split into two steps. First, the marginal parameters are estimated and second, given the estimates of the marginal parameters, the copula parameters are inferred, see also the example in Section 5. Alternatively, Genest et al. (1995) suggested a semiparametric approach in which the univariate marginals are transformed to uniform $[0, 1]$ distributions using the empirical cdf before estimating the parameters of the copula model. See Kim et al. (2007) for a comparison of the joint, IFM, and semiparametric inference methods. Hofmann and Czado (2010) conducted a comparison of the joint and IFM estimation methods using D-vine-copula models with GARCH margins (Bollerslev, 1986) in a Bayesian framework. The results of their analysis show that the benefit of applying a joint estimator is negligible compared to the increase in computational effort that comes along with it. It is thus reasonable to focus attention on DAG copulas in this section.

We examined the practical performance of the DAG-copula model by a simulation study. To this end we drew samples from two families of DAG copulas with conditional independence properties given by the DAG D in Figure 2. More precisely, these families of DAG copulas emerge from two different choices of pair-copula families involved. In either setting we considered six different parameter configurations, resulting in the twelve simulation scenarios described in Table 1. The copula families used (Clayton, Gumbel, Gaussian, and Student's t, see Aas (2004))

exhibit notable differences in tail behaviour as captured by the lower and upper tail dependence coefficients (TDCs)

$$
\begin{aligned}
\lambda_{\mathrm{L}}(U_1, U_2) &= \lim_{u \to 0} \mathbb{P}(U_2 \le u \,|\, U_1 \le u) = \lim_{u \to 0} \frac{C_{12}(u, u)}{u}, \\
\lambda_{\mathrm{U}}(U_1, U_2) &= \lim_{u \to 1} \mathbb{P}(U_2 > u \,|\, U_1 > u) = \lim_{u \to 1} \frac{1 - 2u + C_{12}(u, u)}{1 - u},
\end{aligned}
\tag{4.2}
$$

where C_{12} is the copula of (U_1, U_2), see Joe (1997, Section 2.1.10). Within each copula family a range of values of Kendall's τ as a dependence measure may be obtained by suitable parameter choices. We considered two scenarios in which all selected copulas have $\tau = 0.25$, and two scenarios with $\tau = 0.75$. In each of the eight remaining scenarios one copula has $\tau = 0.75$, while all other copulas have $\tau = 0.25$. See Table 2 for the relations between copula parameters, TDCs, and Kendall's τ.

Scenario			1	2	3	4	5	6
C_{12}	Clayton	δ	0.67 (0.25)	6.00 (0.75)	6.00 (0.75)	0.67 (0.25)	0.67 (0.25)	0.67 (0.25)
C_{13}	Gumbel	δ	1.33 (0.25)	4.00 (0.75)	1.33 (0.25)	4.00 (0.75)	1.33 (0.25)	1.33 (0.25)
C_{24}	Student	ρ, ν	0.38, 5 (0.25)	0.92, 5 (0.75)	0.38, 5 (0.25)	0.38, 5 (0.25)	0.92, 5 (0.75)	0.38, 5 (0.25)
$C_{34|2}$	Gauss	ρ	0.38 (0.25)	0.92 (0.75)	0.38 (0.25)	0.38 (0.25)	0.38 (0.25)	0.92 (0.75)

Scenario			7	8	9	10	11	12
C_{12}	Clayton	δ	0.67 (0.25)	6.00 (0.75)	6.00 (0.75)	0.67 (0.25)	0.67 (0.25)	0.67 (0.25)
C_{13}	Clayton	δ	0.67 (0.25)	6.00 (0.75)	0.67 (0.25)	6.00 (0.75)	0.67 (0.25)	0.67 (0.25)
C_{24}	Clayton	δ	0.67 (0.25)	6.00 (0.75)	0.67 (0.25)	0.67 (0.25)	6.00 (0.75)	0.67 (0.25)
$C_{34|2}$	Clayton	δ	0.67 (0.25)	6.00 (0.75)	0.67 (0.25)	0.67 (0.25)	0.67 (0.25)	6.00 (0.75)

Table 1: Twelve simulation scenarios given by various choices of copula families and parameters for C_{12}, C_{13}, C_{24}, $C_{34|2}$, with notation as in Aas (2004). Parameters were chosen with regard to Kendall's τ as a measure of dependence. The values of τ for each scenario are given in parentheses. See Table 2 for further details on the copula families used.

Copula	Kendall's τ	Lower TDC λ_{L}	Upper TDC λ_{U}
Clayton	$\frac{\delta}{\delta+2} \in [0, 1]$	$2^{-\frac{1}{\delta}}$	0
Gumbel	$1 - \frac{1}{\delta} \in [0, 1]$	0	$2 - 2^{\frac{1}{\delta}}$
Gauss	$\frac{2}{\pi} \arcsin(\rho) \in [-1, 1]$	0	0
Student	$\frac{2}{\pi} \arcsin(\rho) \in [-1, 1]$	$2\,\mathrm{t}_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$	$2\,\mathrm{t}_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$

Table 2: Kendall's τ and tail-dependence coefficients (TDCs) of the selected copulas. $\mathrm{t}_{\nu+1}$ denotes the cdf of a univariate Student's t distribution with $\nu + 1$ degrees of freedom.

In a given simulation scenario with fully specified copulas C_{12}, C_{13}, C_{24}, and $C_{34|2}$, samples $\boldsymbol{u} = (u_1, \ldots, u_4) \in [0, 1]^4$ are obtained by simulating four independent uniform $[0, 1]$ variables

$w_1, \dots, w_4$ and applying the quantile transformations

$u_1 = w_1,$

$u_2 = F_{2|1}^{-1}(w_2 \,|\, u_1; \boldsymbol{\theta}) = h_{1\underline{2}}^{-1}(u_1, w_2; \boldsymbol{\theta}_{12}),$

$u_3 = F_{3|12}^{-1}(w_3 \,|\, \boldsymbol{u}_{12}; \boldsymbol{\theta}) = h_{1\underline{3}}^{-1}(u_1, w_3; \boldsymbol{\theta}_{13}),$

$u_4 = F_{4|123}^{-1}(w_4 \,|\, \boldsymbol{u}_{123}; \boldsymbol{\theta}) = h_{2\underline{4}}^{-1}\left(u_2, h_{3\underline{4}|2}^{-1}\left(\int_0^1 c_{12}(v_1, u_2; \boldsymbol{\theta}_{12})\, h_{1\underline{3}}(v_1, u_3; \boldsymbol{\theta}_{13})\, \mathrm{d}v_1, w_4; \boldsymbol{\theta}_{34|2}\right); \boldsymbol{\theta}_{24}\right),$

cf. equations (2.4) and (3.5). In each simulation run we generated $n = 1000$ i.i.d. observations by that method, and in each scenario we performed $N = 100$ such runs.

For each of the 1200 runs we numerically calculated maximum-likelihood (ML) estimates of the parameters of the underlying DAG-copula model (as specified in Table 1) and compared them to the respective true parameters. Due to the complex structure of the log-likelihood function

$$L(\boldsymbol{\theta}; \boldsymbol{u}) = \sum_{k=1}^{n} \log c_{12}\big(u_1^k, u_2^k; \boldsymbol{\theta}_{12}\big) + \log c_{13}\big(u_1^k, u_3^k; \boldsymbol{\theta}_{13}\big) + \log c_{24}\big(u_2^k, u_4^k; \boldsymbol{\theta}_{24}\big)$$
$$+ \log c_{34|2}\left(\int_0^1 c_{12}\big(v_1, u_2^k; \boldsymbol{\theta}_{12}\big)\, h_{1\underline{3}}\big(v_1, u_3^k; \boldsymbol{\theta}_{13}\big)\, \mathrm{d}v_1, h_{2\underline{4}}\big(u_2^k, u_4^k; \boldsymbol{\theta}_{24}\big); \boldsymbol{\theta}_{34|2}\right) \tag{4.3}$$

(cf. equations (3.5) and (4.1)) our ML estimation procedure follows a stepwise approach as described in Aas et al. (2009) and Hobæk Haff (2010) for vine-PCC models. In a first step we compute ML estimates of the parameters of each pair-copula family separately. Since the evaluation of the copula pdf $c_{34|2}$ requires parameter estimates for the pair copulas C_{12}, C_{13}, and C_{24} on the first level, the estimation of $\boldsymbol{\theta}_{34|2}$ has to be performed last. The parameter estimates obtained in this first step are therefore called *sequential ML estimates*. In a second step the full log-likelihood function is maximised jointly using the sequential ML estimates as starting values, resulting in the so-called *joint ML estimates* $\widehat{\boldsymbol{\theta}}_{vw|\mathrm{pa}(v;w)}$, $v \in V$, $w \in \mathrm{pa}(v)$. This stepwise procedure is performed in each simulation run. In order to summarise ML estimates within each scenario, we used a common scale by converting all parameter estimates to estimates for Kendall's τ as described in Table 2. An overview of our results in terms of empirical bias and mean squared error (MSE) with respect to the true value of Kendall's τ is given in Figure 5.

DAG-PCC models may be viewed as generalisations of Gaussian graphical models as presented in Lauritzen (1996, Chapter 4). As their name implies, the latter models are based on the assumption of a joint normal distribution. ML estimation in these models reduces to an estimation of the correlation matrix, observing the restrictions captured in the DAG specifying the model, see Cox and Wermuth (1996, Chapter 3). By choosing all univariate margins as well as all pair-copula families in a DAG-PCC model based on a DAG D to be Gaussian we obtain the

Gaussian graphical model based on D. The Gaussian DAG copula corresponding to the DAG of our simulation study, for instance, is a four-variate Gaussian copula with correlation matrix

$$\boldsymbol{R} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & R_{14} \\ \rho_{12} & 1 & \rho_{12}\,\rho_{13} & \rho_{24} \\ \rho_{13} & \rho_{12}\,\rho_{13} & 1 & R_{34} \\ R_{14} & \rho_{24} & R_{34} & 1 \end{bmatrix},$$

where ρ_{12}, ρ_{13}, and ρ_{24} are the correlations implied by the pair copulas C_{12}, C_{13}, and C_{24}. Using the conditional correlation $\rho_{34|2}$ implied by the copula $C_{34|2}$ we may represent the correlations R_{14} and R_{34} as

$$R_{14} = \rho_{12}\,\rho_{24} + \frac{\rho_{13}\,\rho_{34|2}\,(1 - \rho_{12}^2)\sqrt{1 - \rho_{24}^2}}{\sqrt{1 - \rho_{12}^2\,\rho_{13}^2}},$$

$$R_{34} = \rho_{12}\,\rho_{13}\,\rho_{24} + \rho_{34|2}\sqrt{1 - \rho_{12}^2\,\rho_{13}^2}\sqrt{1 - \rho_{24}^2}.$$

This representation is based on the conditional independence properties specified by the DAG D in Figure 2 and the iterative formula for partial correlations given in Kurowicka and Cooke (2006, Section 3.3). The latter is applicable since partial and conditional correlations coincide for normal distributions, see Whittaker (1990, Section 6.2). Fitting the Gaussian DAG-copula model to our simulated data sets allows us to compare the estimated pair-copula parameters to the true parameters of the generating models. To ensure comparability, we again transform estimated parameters to estimates of Kendall's τ. Error estimates in this *Gaussian DAG-copula model* are interpreted as in the true model, which we will henceforth refer to as the *non-Gaussian DAG-copula model*. An overview of these error estimates is given in Figure 5.

As stated in Section 3 the DAG-copula model corresponding to the DAG D of our simulation study cannot be represented by a regular vine. The D-vine featuring the same first level of pair copulas C_{12}, C_{13}, and C_{24} as the DAG PCC is given in Figure 3 and specifies a D-vine-copula model that approximates our DAG-copula model. On the second and third level this D-vine comprises the conditional pair copulas $C_{23|1}$, $C_{14|2}$, and $C_{34|12}$. In order to study how well this D-vine-copula model approximates the given DAG-copula model we performed the following procedure in each of the 1200 runs of our simulation study. First we selected pair-copula families for C_{12}, C_{13}, C_{24}, $C_{23|1}$, $C_{14|2}$, and $C_{34|12}$, choosing from the four copula families described in Table 2 and the product copula. More precisely, we first computed sequential ML estimates of the parameters $\boldsymbol{\theta}_{12}$, $\boldsymbol{\theta}_{13}$, $\boldsymbol{\theta}_{24}$, $\boldsymbol{\theta}_{14|2}$, $\boldsymbol{\theta}_{23|1}$, and $\boldsymbol{\theta}_{34|12}$ for the Clayton, Gumbel, Gaussian, and Student's t copula, respectively, and then used Akaike's information criterion (AIC) (Akaike, 1974) to identify the most appropriate copula families. We included C_{12}, C_{13}, and C_{24}, of which the true families are known, in this procedure to be able to judge the reliability of our selection

criterion. Table 3 gives an overview of how often each copula family was selected. In almost all simulation runs the families of C_{12}, C_{13}, and C_{24} were identified correctly. This is in line with results of a simulation study conducted by Brechmann (2010) who concluded that AIC is a reliable selection criterion for bivariate copula families. If $C_{23|1}$ is the product copula in this model then the D-vine in Figure 3 yields a PCC satisfying $U_2 \perp\!\!\!\perp U_3 \mid U_1$ but, in general, fails to satisfy $U_1 \perp\!\!\!\perp U_4 \mid \mathbf{U}_{23}$ as specified by D. If the product copula does not appear in the D-vine we obtain the same log-likelihood as in the model specified by the complete DAG in Figure 4, which is hence an example of a DAG-copula model that can be represented by a D-vine. Note, however, that interest with DAG models lies in capturing conditional independence assumptions and therefore rather in DAGs with missing edges than in complete DAGs. The approximating D-vine-copula model can be viewed as structurally misspecified model for the given data. Given a choice of pair-copula families we then computed joint ML parameter estimates in the D-vine-copula model and compared the maximised log-likelihoods to those obtained in the Gaussian and the non-Gaussian DAG-copula model. The results are given in Figure 6.
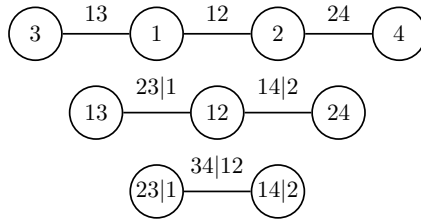

Figure 3: A four-variate D-vine having the first level in common with the DAG PCC derived from the DAG D in Figure 2.
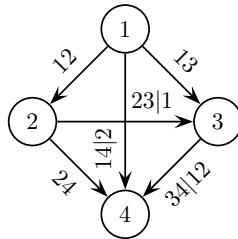

Figure 4: A complete DAG specifying the same PCC as the D-vine in Figure 3 given that none of the associated pair copulas is the product copula. We ordered the parent sets according to $\mathrm{pa}(1; \emptyset) = \emptyset$, $\mathrm{pa}(2; 1) = \emptyset$, $\mathrm{pa}(3; 1) = \emptyset$, $\mathrm{pa}(3; 2) = \{1\}$, $\mathrm{pa}(4; 2) = \emptyset$, $\mathrm{pa}(4; 1) = \{2\}$, and $\mathrm{pa}(4; 3) = \{1, 2\}$.

Since the Gaussian DAG-copula model can be viewed as a misspecified model for the data given in scenarios 1 through 12, it is not surprising that error estimates in this model are generally higher than those in the non-Gaussian DAG-copula models. In fact, Figure 5 shows that the estimates of Kendall's τ obtained in the Gaussian DAG-copula model are considerably worse than those obtained in the non-Gaussian model. Differences between these models with regard

Scenario		1	2	3	4	5	6	7	8	9	10	11	12
C_{12}	Clayton	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
	Gumbel, Gauss, Student, Product: all 0												
C_{13}	Clayton	0	0	0	0	0	0	**100**	**100**	**100**	**100**	**100**	**100**
	Gumbel	**100**	**100**	**99**	**100**	**99**	**98**	0	0	0	0	0	0
	Gauss	0	0	1	0	1	0	0	0	0	0	0	0
	Student	0	0	0	0	0	2	0	0	0	0	0	0
	Product	0	0	0	0	0	0	0	0	0	0	0	0
C_{24}	Clayton	0	0	0	0	0	0	**100**	**100**	**100**	**100**	**100**	**100**
	Gumbel	0	0	2	0	0	1	0	0	0	0	0	0
	Student	**100**	**100**	**98**	**100**	**100**	**99**	0	0	0	0	0	0
	Gauss, Product: all 0												
$C_{14\|2}$	Clayton	3	0	4	1	1	0	95	29	39	100	98	98
	Gumbel	36	86	36	4	36	98	0	0	8	0	0	0
	Gauss	58	0	43	89	54	1	2	6	18	0	2	1
	Student	3	14	4	6	8	1	3	65	4	0	0	1
	Product	0	0	13	0	1	0	0	0	31	0	0	0
$C_{23\|1}$	Clayton	3	3	1	6	8	2	4	2	5	7	6	3
	Gumbel	3	3	3	4	5	4	3	1	5	4	4	2
	Gauss	13	9	7	8	10	10	16	7	10	7	11	14
	Student	5	1	0	2	5	4	1	1	4	3	1	1
	Product	**76**	**84**	**89**	**80**	**72**	**80**	**76**	**89**	**76**	**79**	**78**	**80**
$C_{34\|12}$	Clayton	0	0	0	13	0	0	100	81	100	31	99	100
	Gumbel	0	0	0	3	0	0	0	0	0	2	0	0
	Gauss	93	22	95	78	93	95	0	0	0	62	0	0
	Student	7	78	5	6	7	5	0	19	0	5	1	0
	Product	0	0	0	0	0	0	0	0	0	0	0	0

Table 3: For each scenario (column) and run we identified the copula families with optimal AIC. The resulting frequencies are given above. Frequencies of the true copula families for C_{12}, C_{13}, C_{24}, and $C_{23\|1}$ appear in bold. As an example, the top left entry of the table states that in all of the 100 runs of scenario 1 the Clayton family as a choice for C_{12} yields a higher AIC than all other families considered (Gumbel, Gaussian, Student's t, product). Also, the Clayton family is the true copula family for C_{12} in this scenario.

to bias and MSE for Kendall's τ are more pronounced in scenarios featuring high correlations and asymmetric tail dependence. The smallest difference in performance is thus found in the low-correlation scenarios 1 and 7, whereas in the high-correlation scenarios 2 and 8 the Gaussian DAG-copula model fails by a huge amount. For instance, in scenario 8 the ratios of biases for Kendall's τ in the non-Gaussian and Gaussian DAG-copula model, respectively, are of order 10^3 or higher. The corresponding ratios of MSEs are of order 10^2 or higher. We shall emphasise that Figure 5, with its focus on estimation of Kendall's τ, presents only one aspect in the comparison between non-Gaussian and Gaussian DAG-copula models. For instance, estimation of the degrees of freedom of a Student's t copula is neglected in this figure. It is clear that Gaussian DAG-copula models are useless when interest lies in the estimation of TDCs, say.
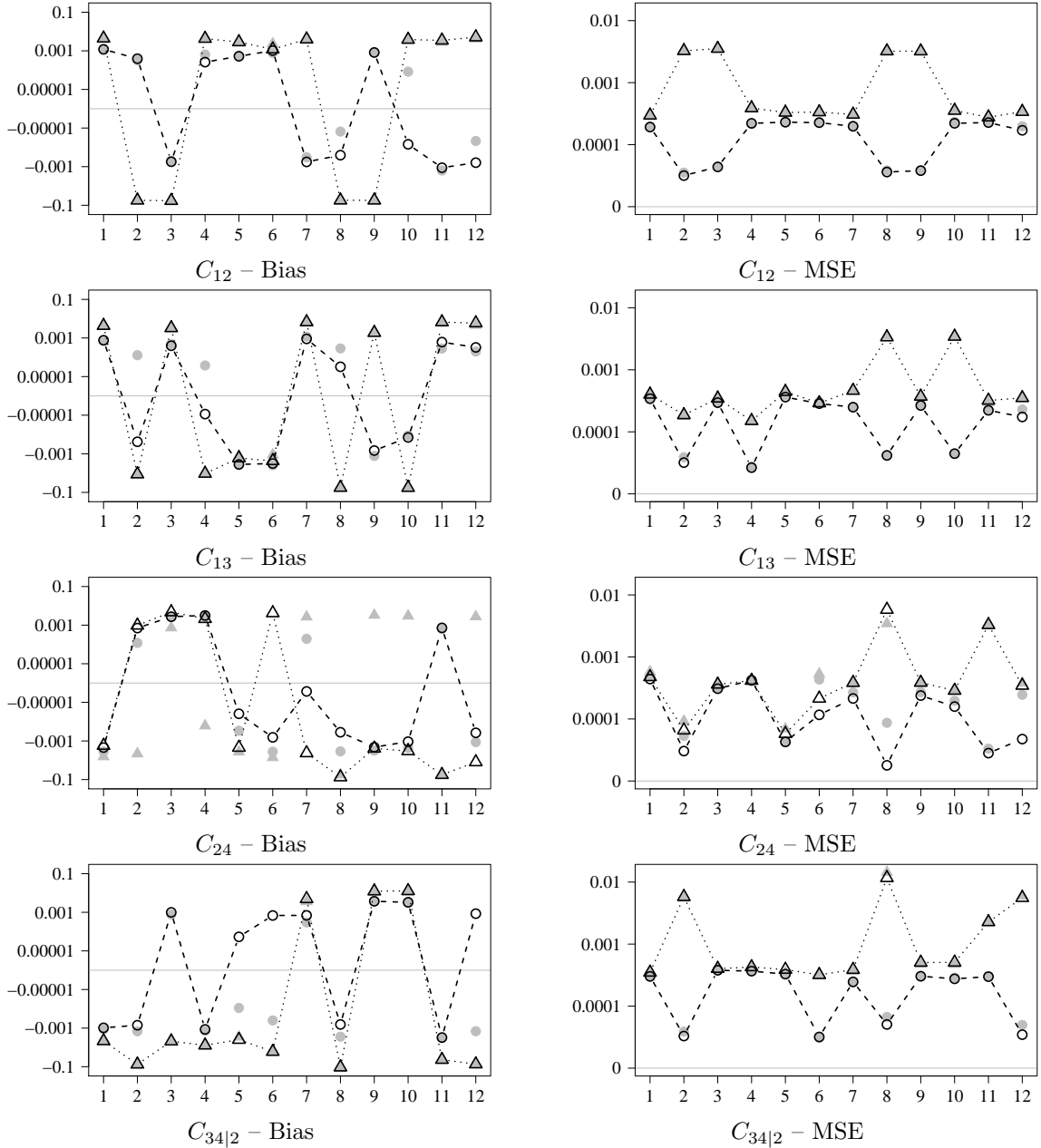
Figure 5: Bias (left) and MSE (right) of estimates of Kendall's τ associated with copulas C_{12}, C_{13}, C_{24}, $C_{34|2}$. The vertical axis uses a transformed log-scale for better visibility. In each of the 12 scenarios (horizontal axis) described in Table 1 parameter estimates were obtained in the true (circle) and in the Gaussian (triangle) DAG-copula model using sequential (solid grey) and joint (outline black) ML estimation.

Since DAG and D-vine-copula models have different parameter sets, direct comparisons of parameter estimates are infeasible. Hence we compared maximised log-likelihoods for the Gaussian DAG, the non-Gaussian DAG, and the D-vine-copula model. Figure 6 shows that maximised log-likelihoods in the non-Gaussian DAG-copula models are roughly 50% higher than those

obtained in their Gaussian counterparts. Even though the Gaussian DAG-copula model in scenarios 1 through 6 has one parameter less than its non-Gaussian competitor this difference in log-likelihood clearly shows the latter model's superiority. Figure 6 also shows the performance of the approximating D-vine-copula model. As this model and the non-Gaussian DAG-copula model share the same first level of pair copulas, it is not surprising that the maximised log-likelihoods in these models differ mainly in those scenarios with high values of $\boldsymbol{\theta}_{34|2}$, namely, scenarios 2, 6, 8, and 12. In other words, the effects of structural misspecification are primarily noticeable in those four scenarios. Note, however, that the D-vine-copula model involves a higher number of parameters than the non-Gaussian DAG-copula model and that its maximised log-likelihood is always slightly inferior to that of the latter model.
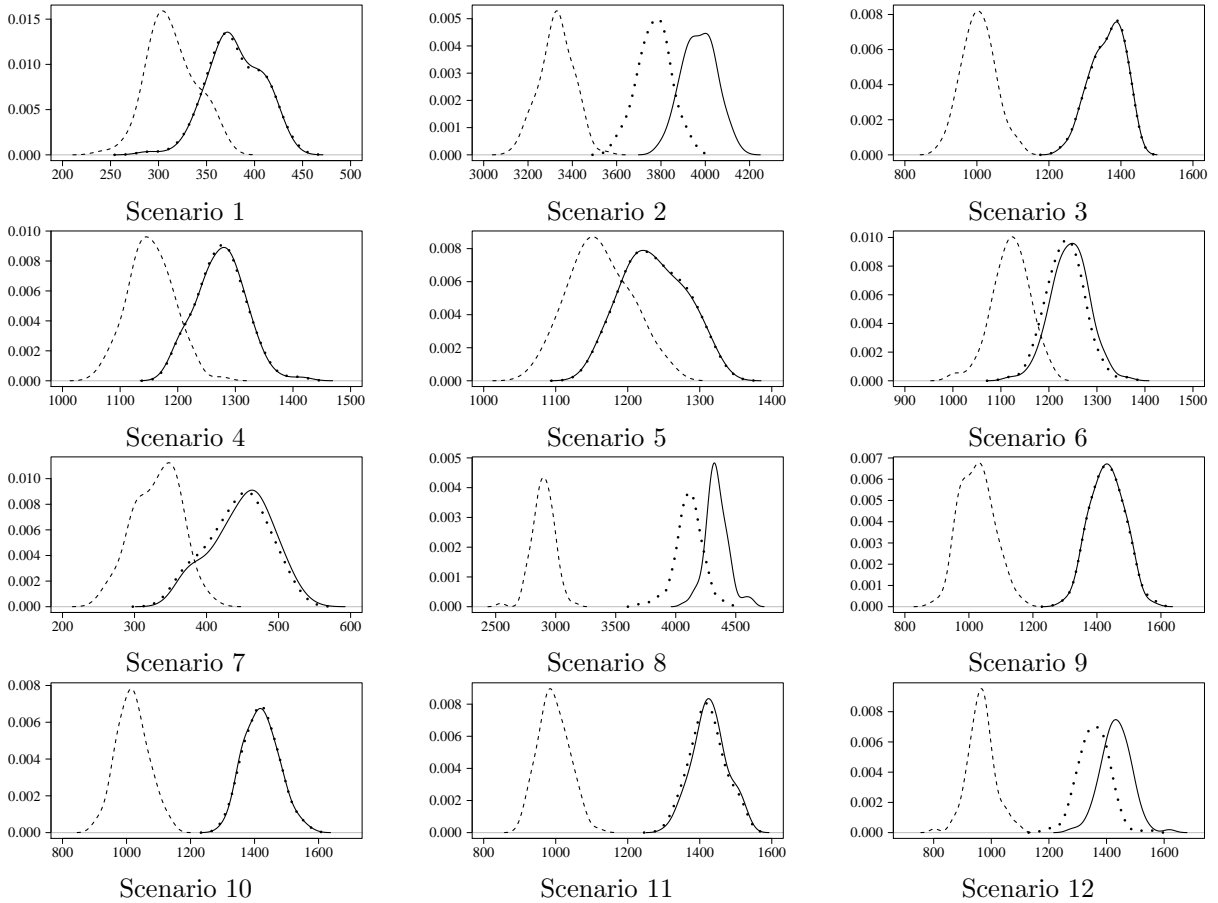

Figure 6: Kernel density estimates of the maximised log-likelihoods in 100 runs for each scenario (see Table 1), based on the Gaussian DAG (dashed line), the non-Gaussian DAG (solid) and the D-vine-copula model (dotted).

The average computation time for joint ML estimation in the non-Gaussian DAG-copula model was twelve seconds in scenarios 1 through 6 and five seconds in scenarios 7 through 12 on a 2 GHz dual-core computer with 2 GB of RAM. Fitting the D-vine-copula model instead reduced

computation time by about 70% in scenarios 1, 3, 4, 5, and 6, and by about 90% in scenarios 7, 9, 10, 11, and 12. This difference in computation time is due to the numerical integration involved in computing joint ML estimates for the non-Gaussian DAG-copula model. In the high-correlation scenarios 2 and 8, however, ML estimation in both models was performed equally fast. Overall there is an increase in computation time when using the non-Gaussian DAG-copula model instead of the D-vine-copula model but it is small compared to the associated gain in statistical precision.

Besides maximised log-likelihoods we also investigated how well the three models capture rank correlations of the bivariate margins $\boldsymbol{U}_{14}$, $\boldsymbol{U}_{23}$, and $\boldsymbol{U}_{34}$ which were not directly included in our DAG and D-vine PCCs. To this end we performed the following procedure in each run. First, we computed estimates $\widehat{\tau}_{14}$, $\widehat{\tau}_{23}$, and $\widehat{\tau}_{34}$ of Kendall's τ for the three margins. Then we generated a sample of $n = 1000$ i.i.d. observations from the Gaussian DAG, the non-Gaussian DAG, and the D-vine-copula model, respectively, using the joint ML parameter estimates obtained before. For each of these samples we again computed estimates of Kendall's τ and compared the results to $\widehat{\tau}_{14}$, $\widehat{\tau}_{23}$, and $\widehat{\tau}_{34}$, respectively. Figure 7 gives an overview of our findings in terms of empirical bias and MSE for each scenario. The patterns described by these error estimates resemble those of the maximised log-likelihoods summarised above and hence yield similar conclusions. The range of bias and MSE values is roughly the same as in Figure 5.

In order to check whether our findings are valid for smaller sample sizes we conducted an additional simulation study in which each simulation run contained $n = 500$ (as opposed to $n = 1000$) observations. As far as estimation of Kendall's τ is concerned we found biases similar to those in Figures 5 and 7 while MSEs tended to be twice as high as those given in Figures 5 and 7. The comparison between Gaussian and non-Gaussian DAG-copula models may be summarised along the same lines as above. By comparison with its competitors the performance of the vine-copula model decreases for smaller sample size and is subject to higher variability.

The main conclusion from our simulation study is that non-Gaussian DAG-PCC models are capable of capturing features in data that neither Gaussian DAG-PCC nor vine-PCC models can reflect. Gaussian DAG-PCC models exhibit particularly poor performance in presence of non-normal tail behaviour. Vine-PCC models are not sufficiently flexible to observe certain Markov properties. Both non-normality and Markov properties are easily incorporated into our non-Gaussian DAG-PCC models.
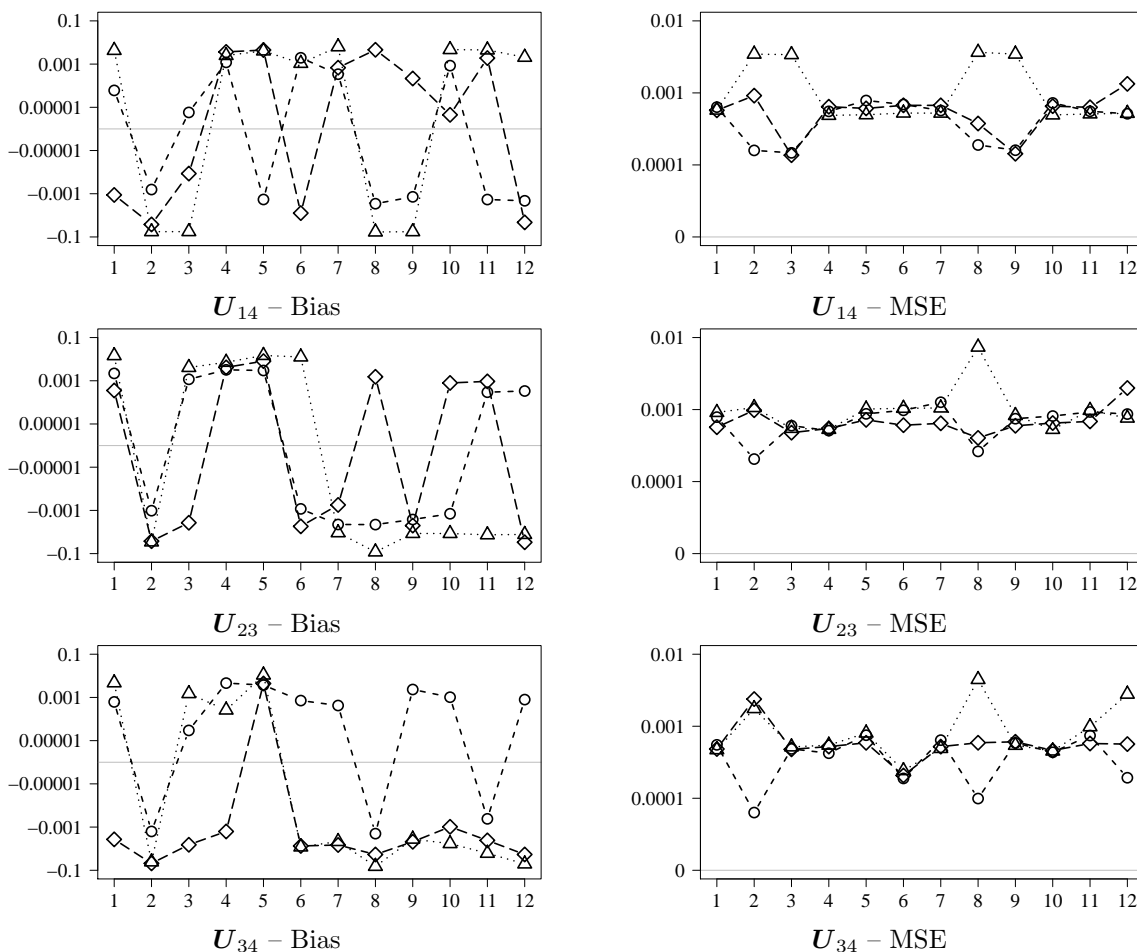
Figure 7: Bias (left) and MSE (right) of estimates of Kendall's τ associated with margins $\boldsymbol{U}_{14}$, $\boldsymbol{U}_{23}$, and $\boldsymbol{U}_{34}$. The vertical axis uses a transformed log-scale for better visibility. In each of the 12 scenarios (horizontal axis) described in Table 1 estimates of Kendall's τ were compared to estimates of Kendall's τ obtained from samples generated from the Gaussian DAG (triangle), the non-Gaussian DAG (circle), and the D-vine-copula model (diamond).

5 Application: Financial returns

We applied the DAG-PCC model to a four-variate financial data set comprising US and German stock and bond indices. More precisely, we modeled the dependence structure of daily log-returns of the Dow Jones Industrial Average (DJI), the Dow Jones Corporate Bond Index (DJCB), the German stock index (DAX), and the corresponding German corporate bond index (RDAX) from 3 April 2007 to 30 September 2010 ($n = 854$ observations). US indices are given in US Dollars and German indices in Euros, that is, we did not correct the data for exchange rate fluctuations. Figure 8 shows the four time series of daily log-returns.

By Sklar's theorem we can model univariate marginal distributions without regard to the
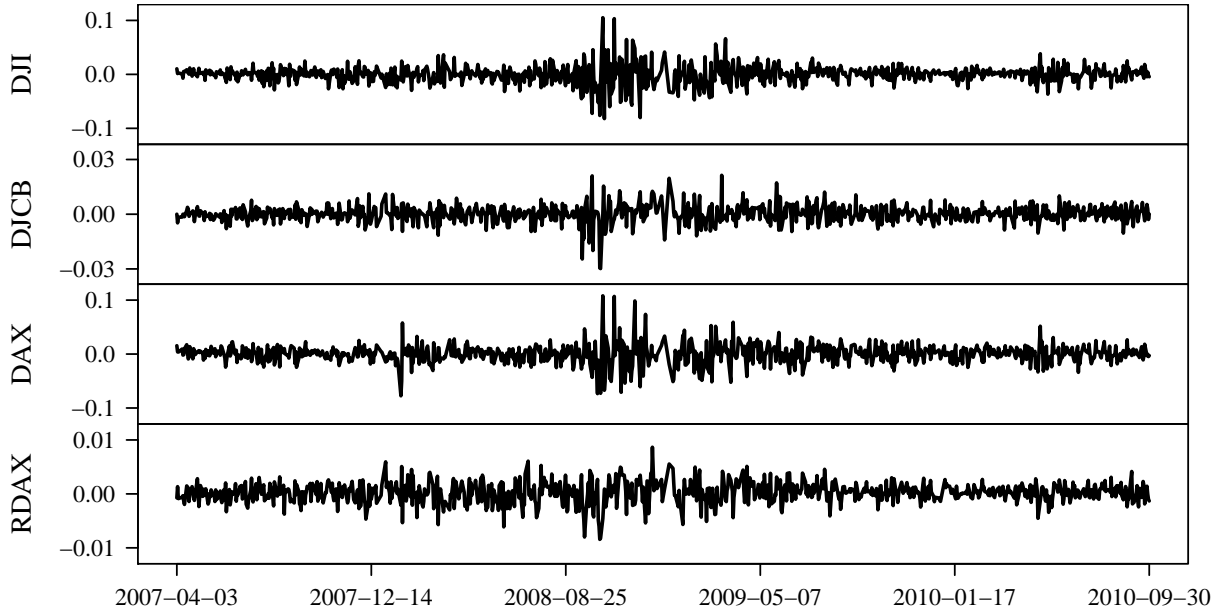
Figure 8: Daily log-returns of the Dow Jones Industrial Average (DJI), the Dow Jones Corporate Bond Index (DJCB), the German stock index (DAX), and the German corporate bond index (RDAX).

dependence structure between variables. Hence we first removed serial correlation in the four time series of log-returns by applying an AR(1) filter with conditional heteroskedasticity being captured by a GARCH(1,1) process, see Bollerslev (1986). The log-return of series $i \in \{DJI, DJCB, DAX, RDAX\}$ at time t is hereby given as

$$x_{t,i} = \mu_i + a_i\, x_{t-1,i} + \varepsilon_{t,i},$$

$$\varepsilon_{t,i} = \sigma_{t,i}\, z_{t,i},$$

$$\sigma_{t,i}^2 = \omega_i + \alpha_i\, \varepsilon_{t-1,i}^2 + \beta_i\, \sigma_{t-1,i}^2$$

with parameters $\omega_i > 0$, $\alpha_i, \beta_i \geq 0$, $\alpha_i + \beta_i < 1$, $|a_i| < 1$, and $\mu_i \in \mathbb{R}$, where $\mathrm{E}\left[z_{t,i}\right] = 0$ and $\mathrm{Var}\left[z_{t,i}\right] = 1$. The standardised residuals $z_{t,i}$ are assumed to follow a univariate Student's t distribution with ν degrees of freedom, that is, $\sqrt{\frac{\nu}{\nu-2}}\, z_{t,i} \sim \mathrm{t}_\nu$. ML estimates and their standard errors derived from numerical evaluation of the Hessian of the AR(1)-GARCH(1,1) parameters for the four time series of log-returns are given in Table 4. Using these standard errors and a 5% significance level we cannot reject the null hypothesis of the Ljung-Box test that there is no autocorrelation left in the residuals and squared residuals, see Ljung and Box (1978). The same holds true for the null hypothesis of the Lagrange-multiplier ARCH test that the residuals exhibit no conditional heteroskedasticity, see Engle (1982). We converted the standardised residuals to uniformly distributed observations $u_{t,i} = \mathrm{t}_\nu\!\left(\sqrt{\frac{\nu}{\nu-2}}\, z_{t,i}\right)$ before modeling the joint dependence structure of the four time series of log-returns by a DAG copula.

	μ [$\times 10^4$]	a	ω [$\times 10^6$]	α	β	ν
DJI	10.15 (3.66)	-0.10 (0.03)	1.99 (37.68)	0.12 (0.02)	0.88 (0.02)	6.33 (1.73)
DJCB	1.62 (1.38)	-0.11 (0.04)	0.56 (28.76)	0.08 (0.02)	0.89 (0.02)	6.75 (1.52)
DAX	5.81 (4.37)	0.01 (0.03)	3.92 (37.95)	0.09 (0.02)	0.89 (0.02)	8.24 (2.30)
RDAX	2.54 (0.73)	0.10 (0.04)	0.02 (23.80)	0.05 (0.01)	0.95 (0.01)	17.30 (9.03)

Table 4: ML estimates and standard errors (in parentheses) of AR(1)-GARCH(1,1) parameters for the DJI, DJCB, DAX, and RDAX daily log-returns.

Based on the economic consideration that the German stock index is driven by its US counterpart and that within the US and Germany corporate bond indices are driven by the respective national stock indices, we propose a conditional independence model for the transformed residuals $u_{t,i}$. The DAG D from Figure 2 with vertices $1, 2, 3, 4$ representing the variables DJI, DJCB, DAX, RDAX, respectively, reflects the above-mentioned dependences and specifies the conditional independence assumptions

$$\text{DJCB} \perp\!\!\!\perp \text{DAX} \mid \text{DJI} \qquad \text{and} \qquad \text{DJI} \perp\!\!\!\perp \text{RDAX} \mid \{\text{DJCB}, \text{DAX}\}.$$

Besides, we also obtain the DAG D when applying the PC algorithm (Spirtes et al., 2000, Section 5.4.2) to the data $\Phi^{-1}(u_{t,i})$, where Φ denotes the standard normal cdf. This transformation is needed since the tests for conditional independence performed by the PC algorithm (at the 5% significance level) are based on the assumption of normality. Since we aim to specify a model with non-Gaussian dependence structure the PC algorithm may, however, only be considered a screening method for potential Markov properties. Given the results presented below we retrospectively measured the reliability of the PC algorithm for the analysed financial data set as follows. We first generated $N = 100$ i.i.d. samples of size $n = 854$ from the non-Gaussian DAG-copula model specified by the joint ML parameter estimates in Table 5 and then applied the PC algorithm to recover the conditional independence properties of each of these samples. The true DAG structure was recovered in 90% of the cases, which supports our model assumptions. By contrast we obtained an overall recovery rate of only 30% when applying the PC algorithm to the 1200 data sets of our simulation study described in Section 4. Highly non-normal data hence require suitable tests for conditional independence to obtain more satisfactory recovery rates. An implementation of the PC algorithm is readily available in the R package `pcalg` (Kalisch et al., 2011).

Given the DAG D Theorem 3.1 prescribes which pair copulas need specification in the definition of our model. Note that vertex 4 (RDAX) has two parents (DJCB and DAX), which necessitates the selection of an ordering of the parents. We decided to use vertex 2 (DJCB) as the first parent based on the heuristic rule of modeling strong bivariate dependences prior to weak dependences. Our decision was based on estimates $\hat{\tau}$ of Kendall's τ between vertices $2, 4$ ($\hat{\tau} = 0.39$) and $3, 4$

($\hat{\tau} = -0.25$), respectively.

Having fixed our model's Markov structure our next step was the selection of parametric copula families for the four copulas C_{12}, C_{13}, C_{24}, and $C_{34|2}$ obtained from Theorem 3.1. We considered Gaussian, Student's t, Frank, Clayton, and Gumbel copula families as well as reflected versions of the Clayton and Gumbel copula families in order to account for negative correlations. Among these candidates we selected copula families based on comparisons of AIC values, which resulted in modeling all four copulas C_{12}, C_{13}, C_{24}, and $C_{34|2}$ by Student's t copulas. Figure 9 displays these choices along with kernel density estimates of the respective true copula pdfs. Visual comparison of these plots strongly affirms our choices of the Student's t copula family. We shall note that the estimates of the correlation parameters ρ and the degrees of freedom ν used to compute AIC values are nothing else than the sequential ML estimates to be obtained from the selected DAG-copula model. Our choice of Student's t copula is consistent with popular modeling approaches in the literature on statistical finance, see Ignatieva and Platen (2010).

Copula pdf c_{12} (DJI, DJCB)   Copula pdf c_{13} (DJI, DAX)

Copula pdf c_{24} (DJCB, RDAX)   Copula pdf $c_{34|2}$ ({DAX, RDAX} | DJCB)
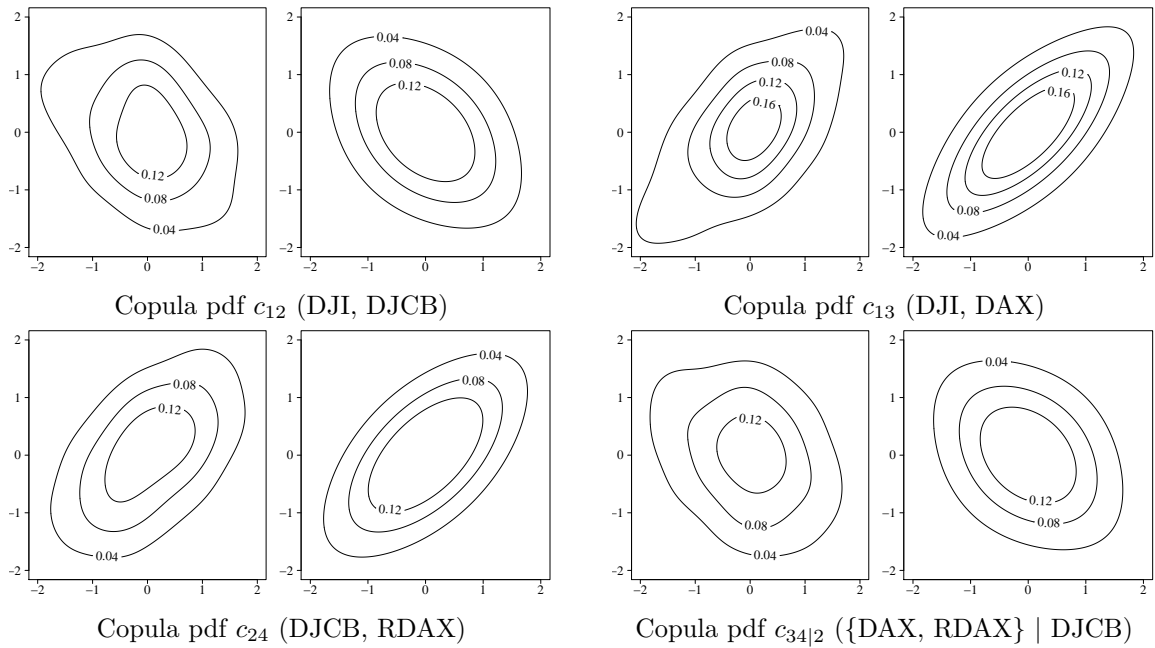
Figure 9: Kernel density estimates of pair-copula pdfs (left) and our choices of Student's t copula pdfs (right) for modeling the DJI, DJCB, DAX, and RDAX data. All copulas are displayed with standard normal margins.

We computed joint ML estimates of the parameters of the selected non-Gaussian DAG-copula model using the routines described in Section 4. In view of reducing model complexity we also applied a semiparametric ML estimator inspired by Hobæk Haff (2010). To this end we replaced the integral in the log-likelihood function in equation (4.3) by a non-parametric conditional-cdf

estimator of $F_{3|2}$ given by

$$\widehat{F}_{3|2}(u_{t,3} \mid u_{t,2}) = \frac{\sum_{s=1}^{n} \Phi\left(\frac{\Phi^{-1}(u_{t,3}) - \Phi^{-1}(u_{s,3})}{h_3}\right) \varphi\left(\frac{\Phi^{-1}(u_{s,2}) - \Phi^{-1}(u_{t,2})}{h_2}\right)}{\sum_{s=1}^{n} \varphi\left(\frac{\Phi^{-1}(u_{s,2}) - \Phi^{-1}(u_{t,2})}{h_2}\right)}, \quad t \leq n,$$

where φ denotes the standard normal pdf and h_2, h_3 are normal-reference rule-of-thumb bandwidths, see Li and Racine (2007, Section 6.2). We used transformed observations $\Phi^{-1}(u_{t,i})$ for the kernel smoothing to avoid boundary effects. The log-likelihood function of this model, which we will refer to as the *semiparametric non-Gaussian DAG-copula model*, takes the form

$$\widehat{L}(\boldsymbol{\theta}; \boldsymbol{u}) = \sum_{t=1}^{n} \log c_{12}(u_{t,1}, u_{t,2}; \boldsymbol{\theta}_{12}) + \log c_{13}(u_{t,1}, u_{t,3}; \boldsymbol{\theta}_{13}) + \log c_{24}(u_{t,2}, u_{t,4}; \boldsymbol{\theta}_{24})$$

$$+ \log c_{34|2}\left(\widehat{F}_{3|2}(u_{t,3} \mid u_{t,2}), h_{24}(u_{t,2}, u_{t,4}; \boldsymbol{\theta}_{24}); \boldsymbol{\theta}_{34|2}\right).$$

ML estimation in this model is performed as in the fully parametric case. Last, we also applied the Gaussian DAG-copula model from Section 4 to the data and compared its performance to the non-Gaussian and the semiparametric non-Gaussian DAG-copula model. Parameter estimates and their standard errors as well as AIC values for all three models are given in Table 5.

		Student C_{12}	Student C_{13}	Student C_{24}	Student $C_{34\|2}$	AIC
nG	S	$-0.35, 10.1 \, (0.03, 3.8)$	$0.66, 9.2 \, (0.02, 4.3)$	$0.57, 17.4 \, (0.02, 4.0)$	$-0.29, 8.9 \, (0.03, 3.4)$	-1002.0
	J	$-0.35, 10.4 \, (0.03, 4.3)$	$0.66, 9.3 \, (0.02, 4.4)$	$0.56, 14.0 \, (0.02, 4.4)$	$-0.29, 8.7 \, (0.03, 3.6)$	-1002.4
sG	S	$-0.35, 10.1 \, (0.03, 3.9)$	$0.66, 9.2 \, (0.02, 4.3)$	$0.57, 17.4 \, (0.02, 4.0)$	$-0.30, 9.8 \, (0.03, 4.3)$	-999.6
	J	$-0.35, 10.1 \, (0.03, 4.2)$	$0.66, 9.2 \, (0.02, 4.4)$	$0.56, 15.6 \, (0.02, 4.2)$	$-0.30, 9.5 \, (0.03, 4.3)$	-999.8
		Gauss C_{12}	Gauss C_{13}	Gauss C_{24}	Gauss $C_{34\|2}$	AIC
G	S	$-0.34 \, (0.03)$	$0.66 \, (0.02)$	$0.57 \, (0.02)$	$-0.28 \, (0.04)$	-971.5
	J	$-0.34 \, (0.03)$	$0.66 \, (0.02)$	$0.57 \, (0.02)$	$-0.28 \, (0.04)$	-971.5

Table 5: Sequential (S) and joint (J) ML estimates, standard errors (in parentheses), and AIC values for the Gaussian (G), the non-Gaussian (nG), and the semiparametric non-Gaussian (sG) DAG-copula model applied to the log-return data.

Due to non-normal tail behaviour observed in the data it is not surprising that the Gaussian DAG-copula model again turns out inferior to its non-Gaussian competitors. Applying the Vuong test with AIC correction (Vuong, 1989) for model selection yields the same conclusion. The results for the non-Gaussian and the semiparametric non-Gaussian DAG-copula model, however, are rather close. In fact, the null hypothesis of the Vuong test that both models are equally close to the true model cannot be rejected at the 5% significance level. Choosing the semiparametric over the fully parametric non-Gaussian DAG-copula model cut down computation time by the factor 50.

6 Conclusion

We have defined non-Gaussian DAG-copula models based on density factorisations that are driven by two ideas: First, the *D*-recursive factorisation known from the literature on DAG models and second, pair-copula constructions (PCCs) for the conditional densities resulting from that factorisation. Such PCCs have been widely used in regular-vine models. Our approach in Section 3, however, is tailored to constructing models with pre-specified Markov properties (as given by a DAG), a feature that cannot be precisely implemented in vine models. The theoretical justification for that approach is given in Theorem 3.1, which also states that univariate margins and pair copulas in these models can be freely chosen and are not limited to the Gaussian case. As is demonstrated in Section 4 our non-Gaussian DAG-copula models are suitable for likelihood inference and outperform both Gaussian DAG and D-vine-copula models in certain settings. In Section 5 we have presented an application to US and German stock and bond indices in which our non-Gaussian DAG-copula models proved superior to their Gaussian competitors.

We shall mention that our approach, which extends an idea from Hanea et al. (2006), involves considerable computational effort. First, the pair-copula constructions used in our expansion of the likelihood function give rise to conditional cdfs whose evaluation requires numerical integration. Second, the routine for finding joint ML estimates as described in Section 4 faces the usual difficulties of non-linear high-dimensional optimisation problems. The first problem may be tackled by approximating certain conditional cdfs by non-parametric estimates as suggested by Hobæk Haff (2010). Including this idea in the optimisation routines used in Section 5 cut down computation time by the factor 50. One way of dealing with the second problem above may be to consider sequential instead of joint ML estimates. Although sequential ML estimates do not maximise the full log-likelihood function they often, as far as our experience shows, turn out to be good approximations to joint ML estimates. However, there is yet no systematic study of this relation.

Selecting a suitable model from the class of non-Gaussian DAG-PCC models is still an open problem and involves both the question for the assumed Markov structure and the question for suitable pair copulas. In Section 5 we used economic considerations to address the problem of selecting a non-Gaussian DAG-PCC model for the given financial data. We are currently researching whether existing structure learning algorithms such as the PC algorithm by Spirtes et al. (2000) can be adapted to our continuous non-Gaussian framework. As mentioned in Section 5, an application of the PC algorithm with conditional independence tests based on the assumption of joint normality to the simulated data sets of Section 4 recovered the true DAG structure in only 30% of the cases. Suitable tests for conditional independence are hence required to obtain more satisfactory recovery rates. One may also investigate whether the Bayesian methodology of model selection suggested for vine-PCC models by Min and Czado

(2011) and Smith et al. (2010) may be applied to non-Gaussian DAG-PCC models. In a similar vein, there is yet no clear answer to the question of whether regular-vine models or non-Gaussian DAG-PCC models are preferable for modeling a given data set.

Also, we are currently working on a general algorithm for simulating DAG copulas with arbitrarily prescribed DAGs as well as for likelihood inference in models based on these DAG copulas. The computational effort mentioned above necessitates a C++ implementation (as opposed to the R implementation that we started out with). The current development stage of this implementation shows promising results in terms of numerical efficiency.

Acknowledgements

References

K. Aas. Modelling the dependence structure of financial assets: A survey of four copulas. Research report SAMBA/22/04, Norwegian Computing Centre, 2004.

K. Aas and D. Berg. Models for construction of multivariate dependence—a comparison study. *The European Journal of Finance*, 15(7–8):639–659, 2009.

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

H. Akaike. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

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

T. Bedford and R. M. Cooke. Vines—A New Graphical Model for Dependent Random Variables. *The Annals of Statistics*, 30(4):1031–1068, 2002.

T. Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

References

E. Brechmann. Truncated and simplified regular vines and their applications. Diploma thesis, Technische Universität München, 2010.

U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance.* John Wiley & Sons, Chichester, 2004.

L. Chollete, A. Heinen, and A. Valdesogo. Modeling International Financial Returns with a Multivariate Regime-switching Copula. *Journal of Financial Econometrics*, 7(4):437–480, 2009.

R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems.* Springer, New York, second edition, 2003.

D. R. Cox and N. Wermuth. *Multivariate Dependencies.* Chapman & Hall, London, 1996.

C. Czado, U. Schepsmeier, and A. Min. Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Statistical Modelling*, 2011. To appear.

G. Elidan. Copula Bayesian Networks. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (Eds.), *Advances in Neural Information Processing Systems*, volume 23, pages 559–567. NIPS Foundation, La Jolla, California, 2010a.

G. Elidan. Inference-less Density Estimation using Copula Networks. In P. Grünwald and P. Spirtes (Eds.), *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*, pages 151–159. AUAI Press, Corvallis, Oregon, 2010b.

R. F. Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007, 1982.

M. Fischer, C. Köck, S. Schlüter, and F. Weigert. An empirical analysis of multivariate copula models. *Quantitative Finance*, 9(7):839–854, 2009.

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

A. M. Hanea, D. Kurowicka, and R. M. Cooke. Hybrid Method for Quantifying and Analyzing Bayesian Belief Nets. *Quality and Reliability Engineering International*, 22:709–729, 2006.

I. Hobæk Haff. Parameter estimation for pair-copula constructions. Research report SAMBA/-36/10, Norwegian Computing Centre, 2010.

I. Hobæk Haff, K. Aas, and A. Frigessi. On the simplified pair-copula construction—Simply useful or too simplistic? *Journal of Multivariate Analysis*, 101(5):1296–1310, 2010.

References

M. Hofmann and C. Czado. Assessing the VaR of a portfolio using D-vine copula based multivariate GARCH models. Submitted for publication, 2010.

K. Ignatieva and E. Platen. Modelling Co-movements and Tail Dependency in the International Stock Market via Copulae. *Asia-Pacific Financial Markets*, 17(3):261–302, 2010.

H. Joe. Families of m-Variate Distributions With Given Margins and $m(m-1)/2$ Bivariate Dependence Parameters. In L. Rüschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics*, volume 28 of *Lecture Notes—Monograph Series*, pages 120–141. Institute of Mathematical Statistics, Hayward, California, 1996.

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

H. Joe and J. J. Xu. The Estimation Method of Inference Functions for Margins for Multivariate Models. Technical report 166, Department of Statistics, University of British Columbia, 1996.

M. Kalisch, M. Mächler, and D. Colombo. *Estimation of CPDAG/PAG and causal inference using the IDA algorithm*, 2011. Manual for the R package `pcalg`.

G. Kim, M. J. Silvapulle, and P. Silvapulle. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, 51(6):2836–2850, 2007.

D. Koller and N. Friedman. *Probabilistic Graphical Models*. MIT Press, Cambridge, Massachusetts, 2009.

D. Kurowicka and R. Cooke. *Uncertainty Analysis with High Dimensional Dependence Modelling*. John Wiley & Sons, Chichester, 2006.

D. Kurowicka and H. Joe (Eds.). *Dependence Modeling: Vine Copula Handbook*. World Scientific, Singapore, 2011.

S. L. Lauritzen. *Graphical Models*. Oxford University Press, Oxford, 1996.

Q. Li and J. S. Racine. *Nonparamatric Econometrics: Theory and Practice*. Princeton University Press, Princeton, New Jersey, 2007.

G. M. Ljung and G. E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton University Press, Princeton, New Jersey, 2005.

A. Min and C. Czado. Bayesian Inference for Multivariate Copulas Using Pair-Copula Constructions. *Journal of Financial Econometrics*, 8(4):511–546, 2010.

References

A. Min and C. Czado. Bayesian model selection for D-vine pair-copula constructions. *The Canadian Journal of Statistics*, 39(2):239–258, 2011.

R. E. Neapolitan. *Learning Bayesian Networks*. Prentice Hall, Upper Saddle River, New Jersey, 2003.

R. B. Nelsen. *An Introduction to Copulas*. Springer, New York, second edition, 2006.

A. J. Patton. Modelling Asymmetric Exchange Rate Dependence. *International Economic Review*, 47(2):527–556, 2006.

O. Pourret, P. Naïm, and B. Marcot (Eds.). *Bayesian Networks: A Practical Guide to Applications*. John Wiley & Sons, Chichester, 2008.

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

M. Smith, A. Min, C. Almeida, and C. Czado. Modeling Longitudinal Data Using a Pair-Copula Decomposition of Serial Dependence. *Journal of the American Statistical Association*, 105(492):1467–1479, 2010.

P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT Press, Cambridge, Massachusetts, second edition, 2000.

Q. H. Vuong. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2):307–333, 1989.

J. Whittaker. *Graphical Models in Applied Multivariate Statistics*. John Wiley & Sons, Chichester, 1990.