Bayesian model selection for D-vine pair-copula constructions

Aleksey MIN and Claudia CZADO

Abstract

In recent years analyses of dependence structures using copulas have become more popular than the standard correlation analysis. Starting from Aas, Czado, Frigessi, and Bakken (2009) regular vine pair-copula constructions (PCCs) are considered the most flexible class of multivariate copulas. PCCs are involved objects but (conditional) independence present in data can simplify and reduce them significantly. In this paper the authors detect (conditional) independence in a particular vine PCC model based on bivariate $t-$copulas by deriving and implementing a reversible jump Markov chain Monte Carlo algorithm. However the methodology is general and can be extended to any regular vine PCC and to all known bivariate copula families. The proposed approach considers model selection and estimation problems for PCCs simultaneously. The effectiveness of the developed algorithm is shown in simulations and its usefulness is illustrated in two real data applications.

Keywords: copula, D-vine, Metropolis-Hastings algorithm, pair-copula construction, reversible jump Markov chain Monte Carlo.

1 Introduction

Over the past decade there has been a large interest in copulas as a tool for capturing the dependence structure between random variables. Since Frees and Valdez (1998), Li (2000) and Embrechts, McNeil, and Straumann (2002), copulas have been widely used in economics, finance and risk management and subsequently applied to other fields. For a comprehensive review on this topic we refer readers to Genest and Favre (2007), Genest, Gendron, and Bourdeau-Brien (2009) and Patton (2009).

Most copula applications deal with bivariate data while examples involving multivariate copulas of dimension $d \geq 3$ are often restricted to Archimedean copulas, elliptical (usually Gaussian or t) copulas or their extensions (see e.g. Song, 2000; Frahm, Junker,

and Szimayer, 2003; Demarta and McNeil, 2005; Fischer, Köck, Schlüter, and Weigert, 2009; McNeil and Neslehova, 2009; Berg and Aas, 2009). In this paper we focus on multivariate copulas known as pair-copula constructions (PCCs), which have gained popularity recently. Here a multivariate copula density is factorized as a product of bivariate copula densities called pair-copulas or building blocks. This construction was first discovered by Joe (1996) mainly in terms of copula distribution functions. Bedford and Cooke (2001, 2002) systemized the approach of Joe (1996) in terms of copula densities by introducing the notion of regular vines. An excellent introduction to regular vines as well as a statistical inference for Gaussian regular vines is given by Kurowicka and Cooke (2006). Aas, Czado, Frigessi, and Bakken (2009) recognized the flexibility and generality of regular vine constructions and moved beyond the Gaussian case by employing bivariate t, Clayton and Gumbel copulas as building blocks for PCCs. In addition they considered statistical inference for PCCs based on maximum likelihood (ML). According to recent empirical investigations of Berg and Aas (2009) and Fischer, Köck, Schlüter, and Weigert (2009), the vine constructions based on bivariate $t-$copulas dominate other multivariate copulas in fitting multivariate financial data.

There is a variety of estimation procedures for copulas. For independent identically distributed (i.i.d) multivariate data, copulas are usually estimated using a semiparametric (SP) approach of Genest, Ghoudi, and Rivest (1995) or a parametric inference for margins (IFM) approach of Joe (2005). These methods are two-stage estimation procedures, where at the first step marginal cumulative distribution functions (CDFs) are fitted and at the second step a parametric copula is fitted by ML. In the SP approach, marginal CDFs are fitted non-parametrically and this results in copula estimation based on empirical ranks. In contrast, the IFM approach assumes parametric marginal CDFs and therefore it is a sequential two-step ML estimation. While these approaches can be used to overcome computational difficulties and numerical instabilities that may be encountered in a joint ML estimation, there is a trade-off as the SP and IFM estimates are generally less efficient than joint ML estimates. In financial applications univariate marginal data is hardly ever i.i.d. Chen and Fan (2006) and Chan, Chen, Chen, Fan, and Peng (2009) combine univariate time series models and copulas to obtain flexible copula-based models for multivariate time series. Further they propose a semiparametric estimation procedure

for copula parameters and show that it leads to consistent and asymptotic normal estimates. Using regular vine PCCs, Min and Czado (2010b) consider copula-based models for multivariate time series with PCCs and use the minimal Kullback-Leibler divergence to estimate copula parameters. The asymptotic variance of estimates for copula parameters does not have a closed analytical form and therefore its computation is based on bootstrap and/or numerically evaluated partial derivatives of the log-likelihood function. For high dimensional data this is very computationally expensive.

In the past decade Markov chain Monte Carlo (MCMC) methods have been successfully used for estimation and inference problems of highly parameterized models. They have first been introduced by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) and Hastings (1970) and nowadays they become standard statistical tools for data analysis. Since a reversible jump MCMC (RJ MCMC) was introduced by Green (1995), MCMC applications become even more widespread as non-nested models of variable dimension can be compared. Nevertheless the Bayesian literature on copulas has, until recently, been sparse. Most Bayesian treatments on copulas have focused either on bivariate families (see Huard, Évin, and Favre, 2006; Silva and Lopes, 2008; Arakelian and Dellaportas, 2009) or multivariate Gaussian and t copulas (see Pitt, Chan, and Kohn, 2006; Dalla Valle, 2009). Moreover Bayesian model selection has only been considered for the bivariate case.

In this paper we approach model selection for D-vine PCCs by identifying (conditional) independence present in data, which can reduce them significantly. This corresponds to identifying whether individual pair-copulas in the D-vine PCC are identical to the independence copula or not. In Min and Czado (2010a) we focused on developing MCMC estimation of D-vine PCCs based on bivariate t copulas and made a first attempt at model selection using the easy to implement approach of Congdon (2006). Since Congdon's approach is biased (see Robert and Marin, 2008), we now make use of more advanced Bayesian model selection methods. Even when a specific D-vine PCC for a d dimensional data and parametric family for bivariate copula building blocks are chosen, there remain $2^{d(d-1)/2}$ PCC models allowing for (conditional) independencies to compare. These models are generally non-nested and as d becomes large comparison of these models is only tractable using computationally intensive methods such as RJ MCMC. Here we derive a

RJ MCMC algorithm for selecting the best model for a chosen D-vine PCC (see Section 3 for definition) when building pair-copulas are bivariate $t-$copulas. However the methodology is generic and applicable to any regular vine as well as to other families of bivariate copulas.

The remainder of the paper is organized as follows. In Section 2 we define copulas and D-vine PCCs. Section 3 presents our RJ MCMC algorithm. It describes the key steps of our algorithm and gives the acceptance probability of birth and death moves. Section 4 contains two simulation studies investigating the small sample performance of the proposed algorithm. In Section 5.1 we revisit the Euro swap data from Min and Czado (2010a). In Section 5.2 we apply our methodology to data from Flury and Riedwyl (1988) on the counterfeit old Swiss 1000-franc bank notes. The paper closes with a conclusion and discussion section.

2 Multivariate copulas and D-vine PCC

Copulas are d-dimensional multivariate distributions with uniformly distributed marginal distributions on $[0,1]$. According to Sklar's theorem (see Sklar (1959)) any continuous multivariate cumulative distribution function (CDF) $F(x_1,\ldots,x_d)$ is determined by its unique copula $C(u_1,\ldots,u_d)$ and marginal CDF $F_i(x_i)$, $i=1,\ldots,d$ through the relationship

$$F(x_1,\ldots,x_d) = C(F_1(x_1), F_2(x_2), ..., F_d(x_d)). \tag{2.1}$$

Excellent introductions to copulas are given in the books by Joe (1997) and Nelsen (1999). From now on we consider only absolutely continuous distributions $F(x_1,\ldots,x_d)$ with a joint density function $f(x_1,\ldots,x_d)$ and marginal density functions $f_i(x_i)$ for $i=1,\ldots,d$. Then relationship (2.1) implies that

$$f(x_1,\ldots,x_d) = c(F_1(x_1),\ldots,F_d(x_d)) \cdot f_1(x_1) \cdot \ldots \cdot f_d(x_d), \tag{2.2}$$

where $c(u_1,\ldots,u_d)$ is the density function of $C(u_1,\ldots,u_d)$. One of the main attraction of copulas is that they allow for the construction of multivariate distributions with given marginal distributions. In addition the dependence structure is captured by the copula

independently of margins since the copula is invariant with respect to increasing transformations of marginal variables.

Any multivariate copula can be represented as a regular vine PCC and in particular as a D-vine PCC. To illustrate this point for D-vines, let us first introduce some notation. Consider a random vector $\mathbf{U} = (U_1, \ldots, U_d)'$ with uniformly $U(0,1)$ distributed margins U_i, $i = 1, \ldots, d$. Let $C(u_1, \ldots, u_d)$ be a copula of $\mathbf{U}$ and $c(u_1, \ldots, u_d)$ be the copula density. For a pair of integers r and s ($1 \leq r \leq s \leq d$) the set $r:s$ denotes all integers between r and s inclusively, i.e. $r:s := \{r, \ldots, s\}$. If $r > s$ then $r:s = \varnothing$. Let $\mathbf{U}_{r:s}$ denote the set of variables $\{U_r, \ldots, U_s\}$. Further $u_{i|r:s}$ denotes the conditional CDF $F_{i|r:s}(u_i|\mathbf{u}_{r:s})$ and $c_{i|r:s}(u_i|\mathbf{u}_{r:s})$ is the corresponding conditional density of U_i given $\mathbf{U}_{r:s}$. In our notation a set of subindices after the vertical line always corresponds to a set of conditioning variables. Using a well known recursive decomposition for any $d-$dimensional density f

$$f(x_1, \ldots, x_d) = f(x_1) \cdot \prod_{k=2}^{d} f(x_k|x_1, \ldots, x_{k-1}),$$

for the copula density c we obtain

$$c(u_1, \ldots, u_d) = 1 \cdot \prod_{k=2}^{d} c_{k|1:(k-1)}(u_k|\mathbf{u}_{1:(k-1)}), \tag{2.3}$$

where 1 stands for the marginal density of U_1. Now consider each factor on the right hand side of (2.3). For $k = 2$ and using $U_1 \sim U(0,1)$ the conditional density $c(u_2|u_1)$ is given by

$$c_{2|1}(u_2|u_1) = c(u_1, u_2). \tag{2.4}$$

Starting for $k = 3$ we use Sklar's theorem for bivariate conditional densities (see Patton, 2004) and obtain

$$\begin{aligned} c_{3|1:2}(u_3|u_1, u_2) &= \frac{c(u_1, u_2, u_3)}{c(u_1, u_2)} = \frac{c(u_1, u_3|u_2)}{c_{1|2}(u_1|u_2)} \\ &= \frac{c_{13|2}\big(F(u_1|u_2), F(u_3|u_2)\big) \cdot c_{1|2}(u_1|u_2) \cdot c_{3|2}(u_3|u_2)}{c_{1|2}(u_1|u_2)} \\ &= c_{13|2}\big(u_{1|2}, u_{3|2}\big) \cdot c(u_2, u_3), \end{aligned} \tag{2.5}$$

where $c_{13|2}$ is a conditional copula density of U_1 and U_3 given $U_2 = u_2$.

By induction, the kth factor ($k = 4, \ldots, d$) in (2.3) can be now factorized as

$$\begin{aligned} c_{k|1:(k-1)}(u_k|\mathbf{u}_{1:(k-1)}) &= \frac{c(u_1, u_k|\mathbf{u}_{2:(k-1)})}{c_{1|2:(k-1)}(u_1|\mathbf{u}_{2:(k-1)})} \\ &= c_{1k|2:(k-1)}(u_{1|2:(k-1)}, u_{k|2:(k-1)}) \cdot c_{k|2:(t-1)}(u_k|\mathbf{u}_{2:(k-1)}) \\ &= \prod_{t=1}^{k-2} c_{tk|(t+1):(k-1)}(u_{t|(t+1):(k-1)}, u_{k|(t+1):(k-1)}) \\ &\times \; c_{(k-1)k}(u_{k-1}, u_k). \end{aligned} \tag{2.6}$$

Substituting the right hand sides of (2.4)–(2.6) into (2.3) and applying convention $i(i+1)|\varnothing := i(i+1)$, we obtain

$$c(u_1, \ldots, u_d) = \prod_{k=2}^{d} \prod_{t=1}^{k-1} c_{tk|(t+1):(k-1)}(u_{t|(t+1):(k-1)}, u_{k|(t+1):(k-1)}) \tag{2.7}$$

$$= \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i(i+j)|(i+1):(i+j-1)}\left(u_{i|(i+1):(i+j-1)}, u_{i+j|(i+1):(i+j-1)}\right). \tag{2.8}$$

Thus, the copula density $c(u_1, \ldots, u_d)$ is factorized as the product of $d(d-1)/2$ unconditional and conditional bivariate copula densities called pair-copulas. There are $(d-1)$ unconditional copulas with subindices $i(i+1)$, $i = 1, \ldots, d-1$. Remaining $(d-1)(d-2)/2$ pair-copulas are conditional and they are evaluated at univariate conditional distribution functions $F_{i|(i+1):(i+j-1)}(u_i|\mathbf{u}_{(i+1):(i+j-1)})$ and $F_{i+j|(i+1):(i+j-1)}(u_{i+j}|\mathbf{u}_{(i+1):(i+j-1)})$.

Equation (2.8) provides a D-vine PCC representation for an arbitrary multivariate copula density and is clearly invariant with respect to any permutation of the variable labels. Since the index j indicates the number of conditioning variables, it is convenient to work with (2.8) while (2.7) is easy to derive. Further pair-copulas on the right hand size of (2.8) can be easily determined with a help of a D-vine, whose edge labels represent subindices of the pair-copulas. For a graphical representation of a $d-$dimensional D-vine we refer the reader to Kurowicka and Cooke (2006), Aas, Czado, Frigessi, and Bakken (2009) or Min and Czado (2010a). To illustrate, the five dimensional D-vine copula density $c(u_1, \ldots, u_5)$ is given by

$$c(u_1, \ldots, u_5) = c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} \cdot c_{13|2} \cdot c_{24|3} \cdot c_{35|4} \cdot c_{14|23} \cdot c_{25|34} \cdot c_{15|234}, \tag{2.9}$$

where we omit arguments of the pair-copulas for clearness.

Aas, Czado, Frigessi, and Bakken (2009) have first noticed the power of PCCs for designing flexible multivariate copulas when a restriction on conditional pair-copulas is

imposed. In particular, the pair-copulas $c_{i(i+j)|(i+1):(i+j-1)}$'s in (2.8) generally depend on the conditioning values $\mathbf{u}_{(i+1):(i+j-1)}$, which makes (2.8) not feasible in its all generality for statistical applications. However the right hand side of (2.8) defines still a valid copula density if we assume that $c_{c_{i(i+j)|(i+1):(i+j-1)}}(\cdot,\cdot)$ are independent of $\mathbf{u}_{(i+1):(i+j-1)}$. Under this restriction Aas, Czado, Frigessi, and Bakken (2009) show how arguments of conditional pair-copulas can be computed (see e.g. Section 3.1). Further Hobæk Haff, Aas, and Frigessi (2010) discuss the above simplification of Aas, Czado, Frigessi, and Bakken (2009) in examples and illustrate that it is not severe. In particular they show that elliptical (e.g. Gaussian or t) copulas meet this assumption of pair-copula independence on conditioning values and therefore the corresponding regular vine PCC representations are equivalent and invariant with respect to variable labeling. Note that under the above restriction the arguments of the conditional pair-copulas still depend on conditioning values and a different variable labeling gives no longer the same multivariate copula density.

3 RJ MCMC

3.1 D-vine PCCs based on $t-$copulas

In this paper we specify the building pair-copulas of the D-vine PCC model (2.8) as bivariate $t-$copulas. However the methodology is generic and is applicable much more widely. The bivariate $t-$copula (see Embrechts, Lindskog, and McNeil, 2003) has 2 parameters: the association parameter $\rho \in (-1,1)$ and the degrees of freedom (df) parameter $\nu \in (0,\infty)$ and its density is given by

$$c(u_1,u_2|\rho,\nu) = \frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{1-\rho^2}\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^2} \cdot \frac{\left(\left[1+\frac{\left(t_\nu^{-1}(u_1)\right)^2}{\nu}\right]\left[1+\frac{\left(t_\nu^{-1}(u_2)\right)^2}{\nu}\right]\right)^{\frac{\nu+1}{2}}}{\left(1+\frac{\left(t_\nu^{-1}(u_1)\right)^2+\left(t_\nu^{-1}(u_2)\right)^2-2\rho t_\nu^{-1}(u_1)t_\nu^{-1}(u_2)}{\nu(1-\rho^2)}\right)^{\frac{\nu+2}{2}}}, \quad (3.1)$$

where $t_\nu^{-1}(\cdot)$ is a quantile function of a $t-$distribution with ν degrees of freedom. Here $\Gamma(a)$ denotes the gamma function given by $\Gamma(a) := \int_0^\infty x^{a-1}e^{-x}dx$. If the df parameter ν of a bivariate $t-$copula $c(\cdot,\cdot|\rho,\nu)$ converges to infinity then a Gaussian copula (see Song, 2000) with parameter ρ is obtained in the limit. Figure 1 compares contour plots

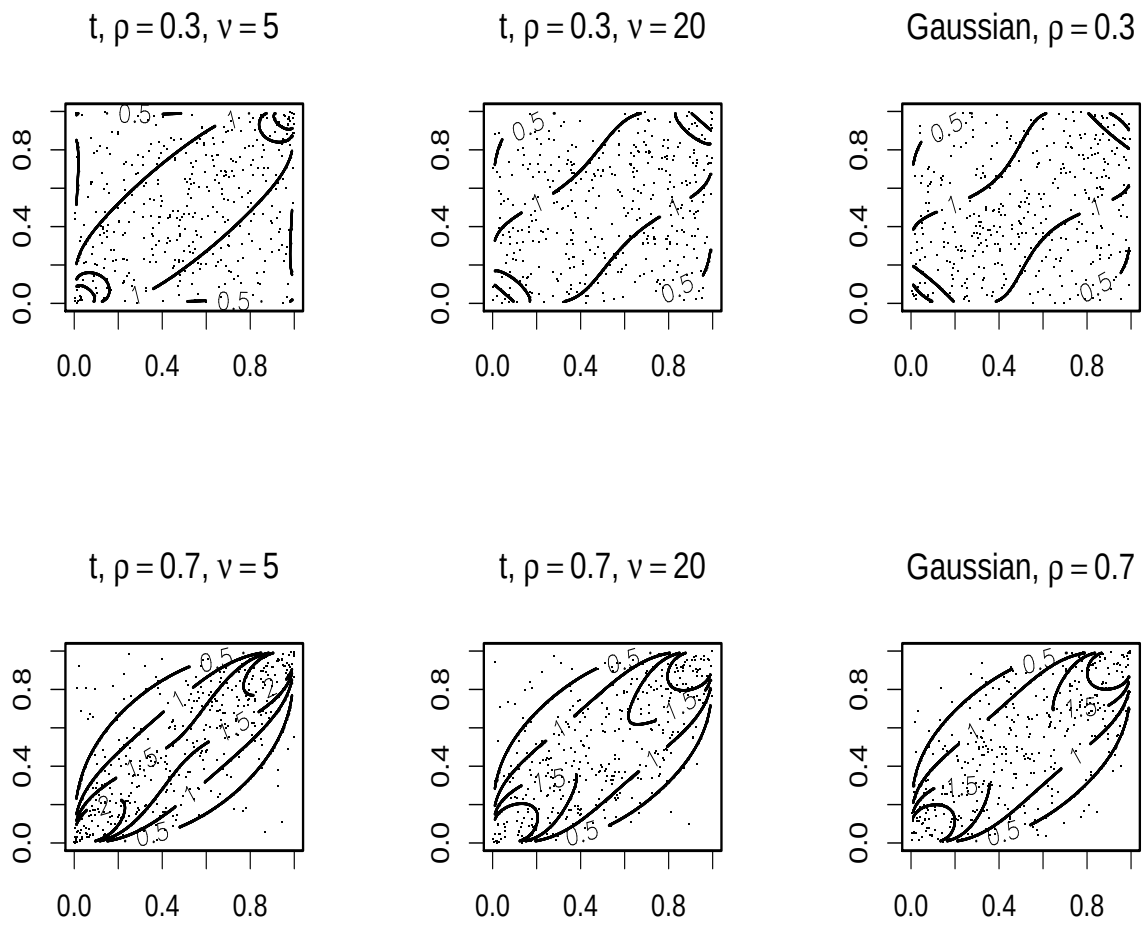

Figure 1: Contour plots of bivariate t and Gaussian copula densities for different parameters. Each contour plot also displays a scatter plot of a random sample of size 400 from the corresponding copula density.

of bivariate t and Gaussian copula densities and indicates that starting from $\nu = 20$ the both copula families are very close to be distinguish numerically.

Specifying the pair-copulas as bivariate t-copulas given in (3.1), the conditional CDF of U_1 given $U_2 = u_2$ is given by

$$h(u_1|u_2, \rho, \nu) := t_{\nu+1}\left(\frac{t_\nu^{-1}(u_1) - \rho\, t_\nu^{-1}(u_2)}{\sqrt{\frac{\left(\nu + \left(t_\nu^{-1}(u_2)\right)^2\right)(1-\rho^2)}{\nu+1}}}\right), \tag{3.2}$$

where $t_{\nu+1}$ denotes the CDF of a $t-$distribution with $\nu + 1$ df. Following Aas, Czado, Frigessi, and Bakken (2009) we call $h(u_1|u_2, \rho, \nu)$ as the $h-$function for the $t-$copula with parameters ρ and ν. The $h-$function (3.2) is essential for any D-vine PCC based on t

pair-copulas since it is needed in the computation of the arguments for the conditional pair-copula. Namely, it turns out that the conditional CDF in (2.8) with k conditioning variables ($2 \leq k < d-1$) is the kth fold nested superposition of the $h-$function (3.2) (see e.g. Aas, Czado, Frigessi, and Bakken, 2009; Min and Czado, 2010a). For example, the first argument $u_{1|23}$ of the pair-copula $c_{14|23}$ in the D-vine PCC (2.9) based on bivariate $t-$copulas is given by

$$h\left[h(u_1|u_2,\rho_{12},\nu_{12})\Big|h(u_3|u_2,\rho_{23},\nu_{23}),\rho_{13|2},\nu_{13|2}\right],$$

where $(\rho_{12},\nu_{12})'$, $(\rho_{23},\nu_{23})'$ and $(\rho_{13|2},\nu_{13|2})'$ are the parameter vectors of pair-copulas c_{12}, c_{23} and $c_{13|2}$, respectively.

We define $\boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)} := (\rho_{i(i+j)|(i+1):(i+j-1)}, \nu_{i(i+j)|(i+1):(i+j-1)})'$ as the parameter vector of $c_{i(i+j)|(i+1):(i+j-1)}(\cdot,\cdot)$, where $j = 1,\ldots,d-1$ and $i = 1,\ldots,d-j$ ($d \geq 3$). To construct the joint parameter vector $\boldsymbol{\theta}$ of the D-vine PCC, we group pair-copula parameters according to the number of conditioning variables of pair-copulas and then order them within each group according to their first subindices, i.e. $\boldsymbol{\theta} := (\boldsymbol{\theta}'_{12},\boldsymbol{\theta}'_{23},\ldots,\boldsymbol{\theta}'_{1d|2:(d-1)})'$ (cf. (2.9)). Now a joint D-vine PCC density of i.i.d. $d-$dimensional observations $\mathbf{u} := \{\mathbf{u}_1,\ldots,\mathbf{u}_N\}$ for $\boldsymbol{\theta}$ is given by

$$c(\mathbf{u}|\boldsymbol{\theta}) := \prod_{n=1}^{N} c(\mathbf{u}_n|\boldsymbol{\theta}) = \prod_{n=1}^{N} c(u_{1,n},\ldots,u_{d,n}|\boldsymbol{\theta}) \tag{3.3}$$

with

$$\begin{aligned} c(\mathbf{u}_n|\boldsymbol{\theta}) &:= \prod_{i=1}^{d-1} c\left(u_{i,n},u_{i+1,n}|\boldsymbol{\theta}_{i(i+1)}\right) \\ &\times \prod_{j=2}^{d-1}\prod_{i=1}^{d-j} c\left(v_{j-1,2i-1,n},v_{j-1,2i,n}|\boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)}\right). \end{aligned} \tag{3.4}$$

Arguments $v_{j-1,2i-1,n}$'s and $v_{j-1,2i,n}$'s of the conditional copulas $c\left(\cdot,\cdot|\boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)}\right)$ for $n = 1,\ldots,N$ have a complex structure. However simple but tedious calculations show (see e.g. Aas, Czado, Frigessi, and Bakken, 2009; Min and Czado, 2010a) that using the $h-$function (3.2), they can be recursively determined as:

$$\begin{aligned} v_{1,1,n} &:= h(u_{1,n}|u_{2,n},\boldsymbol{\theta}_{12}) \\ v_{1,2i,n} &:= h(u_{i+2,n}|u_{i+1,n},\boldsymbol{\theta}_{(i+1)(i+2)}) \quad \text{for} \quad i = 1,\ldots,d-3, \end{aligned}$$

$$
\begin{aligned}
v_{1,2i+1,n} &:= h(u_{i+1,n}|u_{i+2,n}, \boldsymbol{\theta}_{(i+1)(i+2)}) \quad \text{for} \quad i = 1, \ldots, d-3, \\
v_{1,2d-4,n} &:= h(u_d|u_{d-1}, \boldsymbol{\theta}_{d(d-1)}), \\
v_{j,1,n} &:= h(v_{j-1,1,n}|v_{j-1,2,n}, \boldsymbol{\theta}_{1(1+j)|2:j}) \quad \text{for} \quad j = 2, \ldots, d-2, \\
v_{j,2i,n} &:= h(v_{j-1,2i+2,n}|v_{j-1,2i+1,n}, \boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)}) \quad \text{for} \quad d > 4, \\
&\quad j = 2, \ldots, d-3 \quad \text{and} \quad i = 1, \ldots, d-j-2 \\
v_{j,2i+1,n} &:= h(v_{j-1,2i+1,n}|v_{j-1,2i+2,n}, \boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)}) \quad \text{for} \quad d > 4, \\
&\quad j = 2, \ldots, d-3 \quad \text{and} \quad i = 1, \ldots, d-j-2 \\
v_{j,2d-2j-2,n} &:= h(v_{j-1,2d-2j,n}|v_{j-1,2d-2j-1,n}, \boldsymbol{\theta}_{(d-j)d|(d-j+1):(d-1)}) \\
&\qquad\qquad \text{for} \quad j = 2, \ldots, d-2.
\end{aligned}
$$

D-vine PCCs based on bivariate t copulas extend the class of multivariate $t-$ copulas. In particular for any labeling of variables, a multivariate $t-$copula density with association matrix Σ and df parameter $\nu > 2$ (see Embrechts, Lindskog, and McNeil, 2003) can be represented as a D-vine PCC with t pair-copulas, whose ρ parameters are partial correlations (see Yule and Kendall, 1965) computed from Σ and df parameters are equal to $\nu+j$, where j is the number of conditioning variables in a pair-copula. Thus a multivariate $t-$copula density has $d!/2$ different D-vine PCC representations. Therefore, in the sequel we fix one chosen variable labeling and consider the resulting D-vine PCC to avoid the model identifiability problem taking place for multivariate t copulas. In applied work different strategies for variable labeling in D-vine PCCs can be considered and we discuss this point in Section 5.

3.2 Model indicator

If independence or conditional independence in the data is present then the decomposition on the right hand side of (3.4) may be simplified to a sub-decomposition since some factors on the right hand side of (3.4) can be replaced with the density of the independence copula, i.e. with 1. The novelty of this paper consists in detecting this independence, conditional or not, in a chosen D-vine PCC model fully within the Bayesian framework. In other words we determine pair-copula factors in (3.4) which are not the independence copula. We call these copula pairs present. At each MCMC iteration we consider a model vector

$\mathbf{m}$. The model indicator $\mathbf{m}$ will contain information on present pair-copulas and may vary from iteration to iteration.

To define the model indicator $\mathbf{m}$, we have to fix an ordering of pair-copulas in (3.4) as above for the joint parameter vector $\boldsymbol{\theta}$ in (3.3). Further let n_c denote the number of pair-copula factors in (3.4), i.e. $n_c = d(d-1)/2$. First we define the model vector $\mathbf{m}_f$ for the full decomposition (3.4) as a vector of dimension n_c containing subindices of the ordered pair-copulas in (3.4), i.e.

$$\mathbf{m}_f := \big(\underbrace{12, 23, \ldots, (d-1)d}_{(d-1) \text{ components}}, \underbrace{13|2, \ldots, (d-2)d|d-1}_{(d-2) \text{ components}}, \underbrace{\ldots}_{\ldots}, \underbrace{\ldots}_{\ldots}, \underbrace{1d|2, \ldots, d-1}_{1 \text{ component}}\big).$$

This allows to avoid identifiability problems since the decomposition (3.4) is invariant with respect to permutation of pair-copula factors. A model vector $\mathbf{m}$ of a sub-decomposition is obtained from $\mathbf{m}_f$ by replacing components of $\mathbf{m}_f$, which correspond to pair-copula factors $c_{i,i+j|(i+1):(i+j-1)}(u, v) \equiv 1$, $\forall\ u, v \in (0, 1)$, with 0. This can be interpreted as conditional independence of U_i and U_{i+j} given $U_{i+1}, \ldots, U_{i+j-1}$. Further the $s-$th component of a model vector $\mathbf{m}$ is denoted by m_s, i.e. $\mathbf{m} = (m_1, \ldots, m_{n_c})$.

The model vector $\mathbf{m}$ completely describes the corresponding decomposition. It helps us to determine which pair-copula factors could be reduced or which factors could be added. We exclude from our consideration the trivial case when all margins are independent. Therefore the model vector $\mathbf{m}$ may move among all possible $(2^{n_c}-2)$ sub-decompositions of (3.4) and decomposition (3.4) itself. In our RJ MCMC algorithm we put a noninformative proper uniform prior on $\mathbf{m}$, i.e. $\pi(\mathbf{m}) = 1/(2^{n_c}-1)$ for $\forall\ \mathbf{m}$. Thus, RJ MCMC allows us to travel within this large model space without having to fit all possible model specifications. Figure 2 illustrates how model vector $\mathbf{m}$ may move along all possible sub-decompositions when $d = 3$. Further the parameter vector corresponding to $\mathbf{m}_f$ is $\boldsymbol{\theta}_{\mathbf{m}_f} := \boldsymbol{\theta}$ and it contains $2n_c$ components. Similarly to $\mathbf{m}$, a parameter vector $\boldsymbol{\theta}_{\mathbf{m}}$ of $\mathbf{m}$ is obtained from $\boldsymbol{\theta}_{\mathbf{m}_f}$ by replacing two dimensional $\boldsymbol{\theta}_{i(i+j)|(i+1):(i+j-1)}$'s in $\boldsymbol{\theta}_{\mathbf{m}_f}$, which correspond to unit pair-copulas, with $(0, 0)$.

3.3 Prior distributions

In the Bayesian framework unknown parameters are assumed to be random. Available information on the parameters is incorporated through the prior distributions. We use

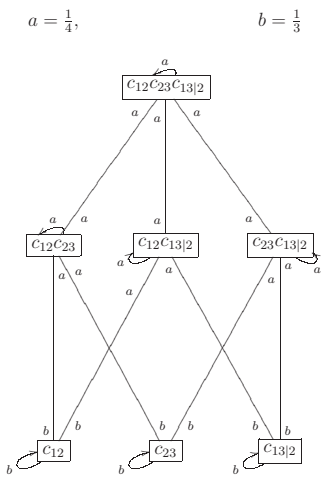

Figure 2: Graph of all possible PCCs for $d = 3$ with the corresponding model switching probabilities

non-informative proper priors for ρ's, i.e. the uniform distribution on $(-1, 1)$. Further we restrict ν's to be between 1 and U. The lower limit is imposed to avoid numerical problems. The upper limit is chosen in such a way that the dependence of bivariate $t-$ copulas on ν for $\nu > U$ is negligible. For ν in $(1, U)$ we also use a noninformative proper prior, i.e. the uniform distribution on $(1, U)$. The upper bound U will be specified later. Finally we assume that all ν's and ρ's are jointly independent a priori which results in the product of prior distributions. Thus the prior distribution of $\boldsymbol{\theta}_{\mathbf{m}}$ in model $\mathbf{m}$ is given by

$$\pi(\boldsymbol{\theta}_{\mathbf{m}}) \;:=\; \pi(\boldsymbol{\theta}|\mathbf{m}) = \prod_{m_s \in \mathbf{m}: m_s \neq 0} \left(\frac{1}{2} \mathbb{1}_{(-1,1)}(\rho_{m_s}) \cdot \frac{1}{U-1} \mathbb{1}_{(1,U)}(\nu_{m_s}) \right). \tag{3.5}$$

Note that in the prior specification we have followed Min and Czado (2010a).

3.4 Move types.

For our RJ MCMC algorithm we consider three types of moves of current model $\mathbf{m}$:

S – "stay" move in the current model $\mathbf{m}$. The parameters of $\mathbf{m}$ are updated.

B – "birth" of factor c_{m_s} in model $\mathbf{m}$, where $m_s \notin \mathbf{m}$. This corresponds to "birth " of $\boldsymbol{\theta}_{m_s}$.

D – "death" of factor c_{m_s} in model $\mathbf{m}$, where $m_s \in \mathbf{m}$. This corresponds to "death " of $\boldsymbol{\theta}_{m_s}$.

An update of parameters takes place in the stay move S. In the full decomposition (3.4) the death move D, the stay move S and no birth move B are possible while in sub-decompositions consisting of only one pair-copula the birth move B, the stay move S and no death move D are possible. Thus we want to exclude the trivial case where all marginal variables are independent. In all other decompositions all three type of moves are possible.

Note that the dimension of a parameter vector without zero components changes in the sense that the number of factors in decomposition may vary. Thus in our case the maximal dimension of the parameter vector with non-zero components is known while in many applications of the RJ MCMC this may not be the case. Further all three moves require a Metropolis-Hastings (MH) step. The MH-step for the stay move is performed similarly to Min and Czado (2010a) and is therefore omitted here. We update parameters individually using a random walk normal proposal truncated to the supports of the parameters and tune proposal variances to achieve an acceptance rate for parameters between 20% and 80%. In the next two sub-sections we derive the acceptance probabilities for the birth and death moves, respectively.

3.5 Model switching probabilities

At each MCMC iteration a decision on the type of moves needs to be taken. In general, we do not have any preference of models, therefore we suppose that each model can be left for another one or can be kept with equal probability. For the full model $\mathbf{m}_f$ there are n_c possibilities to leave the model and there is one possibility to stay and update the parameters. This implies that when the number of non-zero components $p_{\mathbf{m}}$ of $\mathbf{m}$ is equal to n_c (i.e. $\mathbf{m} = \mathbf{m}_f$) then the probability of leaving for a some sub-model or staying at the full model is equal to $1/(n_c + 1)$. In particular the probability $g(\mathbf{m}_f \to \mathbf{m}^{(\text{new})})$ to move from $\mathbf{m}_f$ to any $\mathbf{m}^{(\text{new})}$ with $n_c - 1$ non-zero components is equal to $1/(n_c + 1)$. For $p_{\mathbf{m}} = 1$ the model vector $\mathbf{m}$ contains only one non-zero component. This means that $\mathbf{m}$ cannot be reduced any more. Further there are $n_c - 1$ possibilities to enlarge $\mathbf{m}$ and there is one possibility to stay at $\mathbf{m}$. Therefore the probability $g(\mathbf{m} \to \mathbf{m}^{(\text{new})})$ of enlarging the model

$\mathbf{m}$ to $\mathbf{m}^{(\text{new})}$ is equal to $1/n_c$ if $p_{\mathbf{m}} = 1$. It is not difficult to see that in all other cases the probability $g(\mathbf{m} \to \mathbf{m}^{(\text{new})})$ of enlarging or reducing $\mathbf{m}$ to $\mathbf{m}^{(\text{new})}$ is equal to $1/(n_c + 1)$. Thus

$$g(\mathbf{m} \to \mathbf{m}^{(\text{new})}) = \begin{cases} \frac{1}{n_c+1}, & \text{if} \quad 1 < p_{\mathbf{m}} \leq n_c; \\ \frac{1}{n_c}, & \text{if} \quad p_{\mathbf{m}} = 1. \end{cases} \tag{3.6}$$

Figure 2 displays model switching probabilities for $d = 3$.

3.6 Acceptance probability for birth and death move

First we derive the Metropolis-Hastings step for the birth move B. Let $\mathbf{m}$ be an actual model vector with $q_{\mathbf{m}}$ zeroes, i.e. the corresponding decomposition has $p_{\mathbf{m}} = n_c - q_{\mathbf{m}}$ pair-copula factors. Therefore there are $q_{\mathbf{m}}$ states to which model vector may move and we fix one of them, further denoted by $\mathbf{m}^{(\text{new})}$. The model vectors $\mathbf{m} = (m_1, \ldots, m_{n_c})'$ and $\mathbf{m}^{(\text{new})} = (m_1^{(\text{new})}, \ldots, m_{n_c}^{(\text{new})})'$ differ only in one component, let say in the $s-$th component m_s. Thus, in contrast to $\mathbf{m}$, the $s-$th component of $\mathbf{m}^{(\text{new})}$ is not 0 and $m_i^{(\text{new})} = m_i$ for $i \neq s$.

According to the reversible jump MCMC algorithm (see Green, 1995), we have to propose a new value for $\boldsymbol{\theta}^{(\text{new})}_{m_s^{(\text{new})}} = (\rho_s^{(\text{new})}, \ \nu_s^{(\text{new})})'$. We do this by generating a random vector $\boldsymbol{\eta}_{m_s^{(\text{new})}}$ from a bivariate normal distribution $N_2(\hat{\boldsymbol{\theta}}^{\text{MLE}}_{m_s^{(\text{new})}}, \Sigma_{m_s^{(\text{new})}})$ truncated to $(-1, 1) \times (1, U)$ independently of $\boldsymbol{\theta}_{\mathbf{m}}$. Here $\hat{\boldsymbol{\theta}}^{\text{MLE}}_{m_s^{(\text{new})}} = (\rho_s^{\text{MLE}}, \nu_s^{\text{MLE}})'$ denotes the corresponding two dimensional sub-vector of the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\theta}}^{\text{MLE}}_{\mathbf{m}_f}$ in the full model $\mathbf{m}_f$. Aas, Czado, Frigessi, and Bakken (2009) derive an algorithm to determine the likelihood in a PCC necessary for the MLE. Note that there are n_c covariance matrices $\Sigma_{m_s^{(\text{new})}}$'s, which govern the reversible jump mechanism. They are taken of the form $\Sigma_{m_s} = diag(\sigma^2_{m_s,\rho}, \sigma^2_{m_s,\nu})$, where $diag(a, b)$ denotes a two dimensional diagonal matrix with a and b on the main diagonal. The choice of the variances $\sigma^2_{m_s,\rho}$'s and $\sigma^2_{m_s,\nu}$'s for $\rho_s^{(\text{new})}$'s and $\nu_s^{(\text{new})}$'s will be discussed in the next section. The next step for the RJ MCMC algorithm is to determine a bijection between $(\boldsymbol{\theta}_{\mathbf{m}}, \boldsymbol{\eta}_{m_s^{(\text{new})}})$ and $\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}}$. This can be done in an obvious way by setting for $i = 1, \ldots, n_c$

$$\boldsymbol{\theta}^{(\text{new})}_{m_i^{(\text{new})}} := \begin{cases} \boldsymbol{\theta}_{m_i}, & \text{when} \quad m_i^{(\text{new})} = m_i; \\ \boldsymbol{\eta}_{m_s^{(\text{new})}}, & \text{when} \quad m_s^{(\text{new})} \neq m_s. \end{cases}$$

It is not difficult to see that the Jacobian of the above bijection is equal to 1. According to Green (1995), the acceptance probability for the birth move is now given by

$$\alpha_B := \min\left\{1, \quad \frac{c(\mathbf{u}|\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}})}{c(\mathbf{u}|\boldsymbol{\theta}_{\mathbf{m}})} \cdot \frac{\pi(\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}})}{\pi(\boldsymbol{\theta}_{\mathbf{m}})} \cdot \frac{g(\mathbf{m}^{(\text{new})} \to \mathbf{m})}{g(\mathbf{m} \to \mathbf{m}^{(\text{new})})\varphi_2(\boldsymbol{\eta}_{m_s^{(\text{new})}})} \cdot 1\right\}, \tag{3.7}$$

where $\varphi_2(\boldsymbol{\eta}_{m_s^{(\text{new})}})$ is the bivariate density of the normal distribution $N_2(\hat{\boldsymbol{\theta}}^{\text{MLE}}_{m_s^{(\text{new})}}, \Sigma_{m_s^{(\text{new})}})$ truncated to $(-1,1)\times(1,U)$ and 1 corresponds to the Jacobian. Since (3.5) holds the prior ratio $\pi(\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}})/\pi(\boldsymbol{\theta}_{\mathbf{m}})$ results in $2(U-1)$. Note that prior model probabilities cancel out.

The death move D, where some $\boldsymbol{\theta}_{m_s}$ should be replaced with $(0,0)'$, is the reverse move to the birth move. Now the model vector $\mathbf{m}^{(\text{new})}$ moves to $\mathbf{m}$. Therefore the acceptance probability is obtained here by reverting the fractions on the right hand side of (3.7), i.e.

$$\alpha_D := \min\left\{1, \quad \frac{c(\mathbf{u}|\boldsymbol{\theta}_{\mathbf{m}})}{c(\mathbf{u}|\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}})} \cdot \frac{\pi(\boldsymbol{\theta}_{\mathbf{m}})}{\pi(\boldsymbol{\theta}^{(\text{new})}_{\mathbf{m}^{(\text{new})}})} \cdot \frac{g(\mathbf{m} \to \mathbf{m}^{(\text{new})})\varphi_2(\boldsymbol{\theta}^{(\text{new})}_{m_s})}{g(\mathbf{m}^{(\text{new})} \to \mathbf{m})} \cdot 1\right\}.$$

4 Simulation study

In this simulation study we want to investigate the performance of the RJ MCMC algorithm with regard to its ability to identify the true underlying model. Further we want to study the influence of U and Σ on the performance. We present results for two simulated five-dimensional data sets of sample size $n = 1000$. The copula data are simulated from two D-vine specifications. The first D-vine PCC model (Data 1) consist of 4 bivariate $t-$copulas and 6 bivariate Gaussian copulas with zero correlation and is given by

$$c(u_1,u_2,u_3,u_4,u_5) \;=\; c^t_{12}\cdot c^t_{23}\cdot c^t_{34}\cdot c^t_{45}\cdot c^G_{13|2}\cdot c^G_{24|3}\cdot c^G_{35|4}\cdot c^G_{14|23}\cdot c^G_{25|34}\cdot c^G_{15|234}. \tag{4.1}$$

Here the upper indices t and G indicate on $t-$ and Gaussian copula, respectively. In the first specification we consider the following parameters $\rho_{12} = \rho_{23} = \rho_{34} = \rho_{45} = 0.7$ and $\nu_{12} = \nu_{23} = \nu_{34} = \nu_{45} = 5$ for bivariate $t-$copulas. The second D-vine PCC model (Data 2) consist of 5 bivariate $t-$copulas and 5 bivariate Gaussian copulas with zero correlation and is given by

$$c(u_1,u_2,u_3,u_4,u_5) \;=\; c^t_{12}\cdot c^t_{23}\cdot c^G_{34}\cdot c^t_{45}\cdot c^G_{13|2}\cdot c^G_{24|3}\cdot c^t_{35|4}\cdot c^G_{14|23}\cdot c^t_{25|34}\cdot c^G_{15|234}. \tag{4.2}$$

The following parameters of bivariate $t-$copulas $\rho_{12} = -0.25$, $\rho_{23} = 0.47$, $\rho_{45} = 0.3$, $\rho_{35|4} = -0.2$, $\rho_{25|34} = 0.7$, $\nu_{12} = 4$, $\nu_{23} = 15$, $\nu_{45} = 7$, $\nu_{35|4} = 5$ and $\nu_{25|34} = 10$ are chosen

in the second specification. Since a bivariate Gaussian copula with association parameter $\rho = 0$ is the independence copula with density 1, in both cases the true decomposition is reduced and our RJ MCMC algorithm should detect this.

According to any RJ MCMC procedure the most visited model (decomposition) is considered as the best model. We present results only on model (decomposition) selection since Bayesian estimates within the best model exhibit high quality. Model visits are quantified by posterior model probabilities. A posterior probability estimate of a model is given by the relative frequency of the number of RJ MCMC iterations, at which the model has been visited, with respect to the total number of RJ MCMC iterations. Thus the best model has the highest estimated posterior model probability.

We consider two choices of U, namely $U = 20$ and $U = 300$, for each data set. The first choice $U = 20$ is justified by the fact that the bivariate $t-$ copula with parameters ρ and $\nu > 20$ does not differ much from the Gaussian copula with parameter ρ (see e.g. Figure 1). If $\nu < 20$ then a difference between bivariate Gaussian and t copulas is numerically more pronounced. This allows the independence copula (or the Gaussian copula with $\rho = 0$) and $t-$copula to be well separated. For the second choice $U = 300$ a density of the bivariate t-copula with $\rho = 0$ can be made much closer to 1, the density of independence copula, to be distinguished numerically.

First we run the MH algorithm of Min and Czado (2010a) to tune proposal variances for the full PCC in (4.1) to achieve acceptance rates of the MH algorithm between 20% and 80% as suggested by Besag, Green, Higdon, and Mengersen (1995). Further they are used in the stay move to update the corresponding parameters. It is worth to note that we have observed only a negligible influence of the proposal variances, chosen according to Besag, Green, Higdon, and Mengersen (1995), for the stay move on the posterior probabilities. For the birth move we propose new values for $\boldsymbol{\theta}_{m_s}$ $(s = 1, \ldots, 10)$ according to a $N_2(\hat{\boldsymbol{\theta}}^{\text{MLE}}_{m_s}, \Sigma)$ distribution truncated to $(-1, 1) \times (1, U)$, where $\hat{\boldsymbol{\theta}}^{\text{MLE}}_{\mathbf{m}_f}$ is the MLE for $\boldsymbol{\theta}$ restricted to be in $[(-1, 1) \times (1, 20)]^{10}$. Thus we use the same covariance matrix Σ for all 10 pair-copulas. For each choice of U we consider two choices for Σ. In the first case we consider $\Sigma_1 = diag(100^2, 15000^2)$. This choice of the covariance matrix actually corresponds to a nearly uniform distribution on $(-1, 1) \times (1, U)$. In contrast to the first case we also consider $\Sigma_2 = diag(0.2^2, 5^2)$. This gives more weight to the proposal means.

Further we use the ML estimate $\hat{\boldsymbol{\theta}}^{\text{MLE}}_{\mathbf{m}_f}$ and the model vector $\mathbf{m}_f$ of the full decomposition (4.1) as initial values for $\boldsymbol{\theta}$ and $\mathbf{m}$, respectively.

There are 10 copula terms in (4.1) and in (4.2), which can be present or not in a model. This implies that there are $1023 = 2^{10} - 1$ models to be explored by the RJ MCMC algorithm. In order to enumerate them we treat their corresponding model vectors $\mathbf{m}$ as a binary representation of a decimal number. If a pair-copula is present (not present) in the decomposition than the corresponding entry of $\mathbf{m}$ is 1 (0). Thus the full decomposition (4.1) has the model vector $\mathbf{m} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ and it corresponds to 1023. The decomposition with four pair-copulas c_{12}, c_{23}, c_{34} and c_{45} has the model vector $\mathbf{m} = (1, 1, 1, 1, 0, 0, 0, 0, 0, 0)$ corresponding to 960. There is a one-to-one correspondence between all possible models and the sequence of $1, \ldots, 1023$. Now we can enumerate all possible submodels as $M_1, \ldots, M_{1023}$ by using a decimal representation of $\mathbf{m}$ as the sub-index of M. The true model vector of the second D-vine PCC specification (4.2) is $\mathbf{m} = (1, 1, 0, 1, 0, 0, 1, 0, 1, 0)$ and corresponds to M_{842}.

For each data set we performed 4 RJ MCMC runs for each combination of Σ and U. Each run consists out of 100,000 iterations. Using trace plots (not shown) for parameters of visited models we see that 10,000 iterations are sufficient for burn-in. The resulting posterior probability estimates $\hat{P}_k := \hat{P}(M_k|\text{data})$ are presented in Table 1. We identify only models with $\sum_k \hat{P}_k > 0.9$ at least. From these we see that in all choices of U and Σ, the true models M_{960} for Data 1 and M_{842} for Data 2 have the highest estimated posterior model probability. However these estimates are considerable higher for $U = 20$ compared to $U = 300$ for both data sets and both choices of Σ. This is to be expected since for $U = 300$ we are much closer to an identifiability problem than for $U = 20$. For all cases considered the true model is detected with at least 0.8855 probability for $U = 20$, while for $U = 300$ this value is only 0.5460. Therefore $U = 20$ provides a reasonable choice for upper prior limit U. In the following applications we therefore choose $U = 20$. The influence of the proposal covariances for the birth/death moves on $\hat{P}_k$s is minimal indicating robustness of the RJ MCMC algorithm with regard to this choice.

Table 1 also displays proportions in % of accepted birth-death moves for each of the simulated data. They vary between 1.38% and 5.52% and are of order or lower than ones (4%, 7% and 18%) reported by Richardson and Green (1997). This is to be expected since

Table 1: Estimated posterior model probabilities $\hat{P}_k = \widehat{P}(M_k|\text{data})$ of all 1023 models using different birth proposal covariances and upper prior limits for df's for simulated Data 1 and Data 2.

Data 1					
Model	Model indicator	$\hat{P}_k$			
		$\Sigma_1 = diag(10^2, 1500^2)$		$\Sigma_2 = diag(0.2^2, 5^2)$	
		$U = 20$	$U = 300$	$U = 20$	$U = 300$
M_{960}	$(1,1,1,1,0,0,0,0,0,0)$	0.8855	0.5460	0.9000	0.5743
M_{992}	$(1,1,1,1,1,0,0,0,0,0)$	0.0292	0.1748	0.0191	0.2069
M_{968}	$(1,1,1,1,0,0,1,0,0,0)$	0.0533	0.0878	0.0535	0.0545
M_{964}	$(1,1,1,1,0,0,0,1,0,0)$	0.0169	0.0363	0.0138	0.0397
M_{1000}	$(1,1,1,1,1,0,1,0,0,0)$	0.0013	0.0288	0.0015	0.0254
M_{976}	$(1,1,1,1,0,1,0,0,0,0)$	0.0026	0.0227	0.0034	0.0190
M_{962}	$(1,1,1,1,0,0,0,0,1,0)$	0.0060	0.0207	0.0055	0.0224
		13 M_k's with $\sum_k \hat{P}_k < 0.01$	44 M_k's with $\sum_k \hat{P}_k < 0.1$	13 M_k's with $\sum_k \hat{P}_k < 0.01$	28 M_k's with $\sum_k \hat{P}_k < 0.1$
% of accepted birth-death moves		1.57%	4.90%	1.89%	5.52%
Data 2					
M_{842}	$(1,1,0,1,0,0,1,0,1,0)$	0.9582	0.8288	0.9758	0.8379
M_{970}	$(1,1,1,1,0,0,1,0,1,0)$	0.0088	0.0239	0.0112	0.0454
M_{874}	$(1,1,0,1,1,0,1,0,1,0)$	0.0057	0.0807	0.0064	0.0768
		16 M_k's with $\sum_k \hat{P}_k < 0.01$	23 M_k's with $\sum_k \hat{P}_k < 0.1$	14 M_k's with $\sum_k \hat{P}_k < 0.01$	15 M_k's with $\sum_k \hat{P}_k < 0.1$
% of accepted birth-death moves		1.38%	1.97%	1.93%	2.28 %

we have a multivariate problem, where not only the number of pair-copulas but also which pair-copulas are chosen is very important. In general, proportions of accepted birth-death moves are higher for D-vine PCC models capturing more (conditional) independencies present in data (see also Section 5). Further it seems that using Σ_2 as the proposal covariance leads to a RJ MCMC scheme that traverses the model space slightly more efficiently, as highlighted by a higher percentage of accepted birth/death moves.

5 Applications

5.1 Swap data

As a first application we consider daily Euro swap rates with maturity of 2, 3, 5, 7 and 10 years, respectively, from Min and Czado (2010a). Swap rates are based on annually compounded zero coupon swaps. The data covers the time slot from December 7, 1988 until May 21, 2001 containing 3182 working days. Following Min and Czado (2010a) each margin of the original data has been filtered using an ARMA(1,1)-GARCH(1,1) to obtain standardized independent residuals. Further the marginally independent residuals are transformed with their empirical distribution function in order to achieve approximately uniform i.i.d. margins as advocated by Genest, Ghoudi, and Rivest (1995). Our algorithm is applied to these transformed data which are denoted by S_2, S_3, S_5, S_7 and S_{10}, respectively. Note that for convenience the subindices indicate the maturity years of swaps.

Deciding to fit a D-vine PCC model in a practical application, the variable labeling for a D-vine PCC should be fixed. Aas, Czado, Frigessi, and Bakken (2009) originally advocated to choose a bivariate dependence measure such as a tail dependence coefficient (TDC) and reflect the most bivariate dependence in unconditional pair-copulas. However one can also label variables according to some hypotheses one wants to check or even do this somewhat arbitrarily. Here we follow Aas, Czado, Frigessi, and Bakken (2009) to fix the variable labeling for the swap data. In Min and Czado (2010a) we looked at all bivariate TDC of bivariate t copulas and found out that the ordering $(S_2, S_3, S_5, S_7, S_{10})$ reflects the most bivariate tail dependence in unconditional pair-copulas of a D-vine PCC. Therefore the investigated full D-vine PCC with t pair-copulas contains 10 pair-copulas and is given by

$$\begin{aligned} c(u_{S_2}, u_{S_3}, u_{S_5}, u_{S_7}, u_{S_{10}}) &= c_{S_2S_3} c_{S_3S_5} c_{S_5S_7} c_{S_7S_{10}} c_{S_2S_5|S_3} c_{S_3S_7|S_5} c_{S_5S_{10}|S_7} \\ &\times c_{S_2S_7|S_3S_5} c_{S_3S_{10}|S_5S_7} c_{S_2S_{10}|S_3S_5S_7}, \end{aligned} \tag{5.1}$$

where for brevity we drop the arguments and the parameters of the pair-copulas. Min and Czado (2010a) estimated parameters of Model (5.1) in a Bayesian framework and came to conclusion that (5.1) could be reduced by the pair $c_{S_2S_{10}|S_3S_5S_7}$. In particular the

Table 2: Estimated posterior model probabilities $\hat{P}_k = \widehat{P}(M_k|\text{data})$ of all 1023 models for the swap data. The upper prior limit U for ν is equal to 20.

Model	PCC	Formula	$\hat{P}_k$	
			$\Sigma_1 = diag(10^2, 150^2)$	$\Sigma_2 = diag(0.1^2, 5^2)$
M_{1023}:	with all pairs	$c_{S_2S_3}c_{S_3S_5}c_{S_5S_7}c_{S_7S_{10}}$ $\times c_{S_2S_5\|S_3}c_{S_3S_7\|S_5}c_{S_5S_{10}\|S_7}$ $\times c_{S_2S_7\|S_3S_5}c_{S_3S_{10}\|S_5S_7}c_{S_2S_{10}\|S_3S_5S_7}$	0.2866	0.3017
M_{1022}:	without $c_{S_2S_{10}\|S_3S_5S_7}$	$c_{S_2S_3}c_{S_3S_5}c_{S_5S_7}c_{S_7S_{10}}$ $\times c_{S_2S_5\|S_3}c_{S_3S_7\|S_5}c_{S_5S_{10}\|S_7}$ $\times c_{S_2S_7\|S_3S_5}c_{S_3S_{10}\|S_5S_7}$	0.7134	0.6983
M_i		for $i \neq 1022, 1023$	0	0
% of accepted birth-death moves			0.96%	1.53%

95% credibility interval for $\rho_{S_2S_{10}|S_3S_5S_7}$ included zero and the corresponding estimated df is large.

We run 100,000 RJ MCMC iterates for upper prior limit $U = 20$ for the df ν and consider the first 10,000 iterations as burn-in. In (5.1) we tune proposal variances for the stay move. Proposal covariance matrices $\Sigma_1 = diag(10^2, 150^2)$ and $\Sigma_2 = diag(0.1^2, 5^2)$ are used for the birth of new components. Further the MLE's restricted to $(-1, 1) \times (1, 20)$ are chosen as the starting values and $\mathbf{m}_f = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)'$ is chosen as the starting model indicator. Note that there are a total of $2^{10} - 1 = 1023$ models to be compared. Table 2 gives posterior model probability estimates for all models. Only two models have been visited and this a consequence of high dependence present in data. As result a low proportion of accepted birth and death moves is observed (see Table 2). Nevertheless RJ MCMC iterates showed acceptable mixing. Note that the models are enumerated as in Section 4. Model M_{1023} corresponds to the full decomposition (5.1) and Model M_{1022} is associated with (5.1) reduced by the pair $c_{S_2S_{10}|S_3S_5S_7}$. Since Model M_{1022} has been visited in 71.34% and 69.83% cases for the covariance matrix Σ_1 and Σ_2, respectively, we conclude that the influence of the proposal distribution for birth components of the RJ mechanism is negligible. Thus, M_{1022} is the best model and we restrict ourself now to its RJ MCMC iterates.

Having chosen the best model, its parameters can be estimated from those iterations which belong to this best model. The estimated autocorrelation for each parameter in RJ MCMC iterations for the best model is very high while the cross-correlation is of order 10^{-2}. This high autocorrelation is to be expected since at each iteration the model can move to $n_c - 1$ or n_c models, or stay in the current model resulting in a update of the parameters. It will take some time, especially for large values of n_c, before a possibility to move in a "right" direction is given. Then the RJ mechanism will decide whether a change of model should be performed.

For the best model M_{1022} we subsample corresponding MCMC iterations for each margin. Each 200-th iteration has been recorded and this ensures low autocorrelation in all marginal MCMC iterations. Figure 3 shows posterior kernel density estimates for the parameters of M_{1022} and for both choices of the covariance matrices Σ_1 (solid lines) and Σ_2 (dashed lines) based on every 200-th iterations. We see that both choices of Σ result in very close posterior density estimates. Further the estimated posterior density of the ρ's are close to normal densities. Posterior mode estimates for ν are very low. Taking into account high posterior estimates of ρs, this demonstrates heavy tail dependence of the swap data. Finally note that Figure 3 also displays posterior density estimates (dotted line) for the Bayesian analysis of the full decomposition (5.1) without model selection based on each 20-th iteration of overall 10,000 iterations with burn-in 1000. It should be also mentioned that numerical estimates of standard errors for MLE's of PCC (5.1) are only partially available since the numerical evaluation of the Hessian matrix leads to a nonpositive variance estimate. In particular numerical variance estimates of $\hat{\rho}^{\mathrm{MLE}}_{S_2S_3}, \hat{\rho}^{\mathrm{MLE}}_{S_3S_5}, \hat{\rho}^{\mathrm{MLE}}_{S_5S_7}$ and $\hat{\rho}^{\mathrm{MLE}}_{S_7S_{10}}$ are found to be negative.

5.2 Swiss counterfeit bank notes data

Following the Associate Editor's suggestion to consider i.i.d. multivariate data, we apply our RJ MCMC algorithm to part of the data set on genuine and counterfeit old Swiss 1000-franc bank notes from Flury and Riedwyl (1988). The analyzed data represent geometric measurements (in mm) on 100 Swiss 1000-franc counterfeit bank notes and a short variable description is given in Table 3. Ignoring marginal models, we utilize empirical probability transformation to create the copula data for our analysis. In contrast to the previous

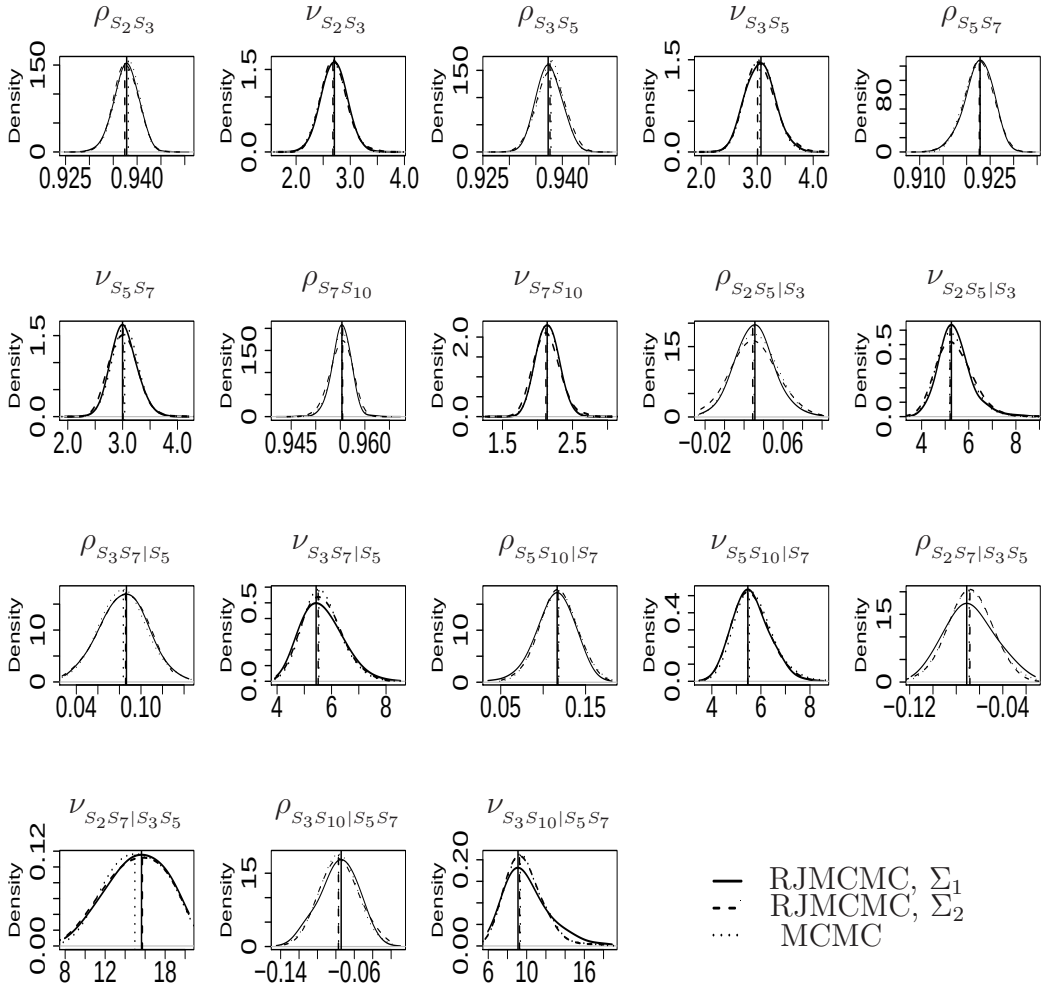

Figure 3: Estimated posterior densities of the parameters in Model M_{1022} for the swap rate data based on each 200-th iterations and obtained with the covariance matrices Σ_1 (solid lines) and Σ_2 (dashed lines) for birth move. The corresponding estimated posterior densities and posterior mode estimates of MCMC for the full decomposition are shown as dotted lines. Vertical lines (solid, dashed or dotted) indicate posterior mode estimates.

application, we do not label variables to reflect highest tail dependencies. Instead we consider just the original labeling of variables and the corresponding D-vine PCC model for the six dimensional copula data with 15 pair-copulas is given by

$$\begin{aligned} c(u_1,\ldots,u_6) &= c_{12}\cdot c_{23}\cdot c_{34}\cdot c_{45}\cdot c_{56}\cdot c_{13|2}\cdot c_{24|3}\cdot c_{35|4}\cdot c_{46|5} \\ &\times c_{14|23}\cdot c_{25|34}\cdot c_{36|45}\cdot c_{15|234}\cdot c_{26|345}\cdot c_{16|2345}. \end{aligned} \tag{5.2}$$

Table 3: Variables of the data on Swiss counterfeit bank notes.

Notation	Short description
"1"	– Length in mm of the bank note
"2"	– Height in mm of the bank note, measured on the left
"3"	– Height in mm of the bank note, measured on the right
"4"	– Distance in mm of inner frame to the lower border
"5"	– Distance in mm of inner frame to the upper border
"6"	– Length in mm of the diagonal

As in the preceding application we run 100,000 RJ MCMC iterates for upper prior limit $U = 20$ for the df ν and consider the first 10,000 iterations as burn-in. For the stay move we tune proposal variances in the full model (5.2). The proposal covariance matrices $\Sigma_1 = diag(10^2, 150^2)$ and $\Sigma_2 = diag(0.1^2, 5^2)$ are used for the birth of new components. Further the MLE's restricted to $(-1, 1) \times (1, 20)$ are chosen as the starting values and $\mathbf{m}_f = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)'$ is chosen as the starting model indicator. Note that there are a total of $2^{15} - 1 = 32767$ models to be compared. Table 4 summarizes posterior model probability estimates for all models. In contrast to the previous example, this data set contains more conditional independencies and therefore significantly more models have been visited and relatively high acceptance rate of birth-death moves is observed (see Table 4). Models M_{26728} and M_{26729} possess the highest estimated posterior model probabilities and are of equal magnitude. Estimated posterior model probabilities of all other visited models are smaller or even significantly smaller than 0.03.

From Table 4 we see that depending on the RJ mechanism (i.e. on the birth proposal covariance matrix) the superiority of the models alternates slightly. Thus we cannot distinguish between Models M_{26728} and M_{26729}. However if we apply our algorithm in the same setup as above but with the upper prior limit $U = 10$ then we obtain the estimated posterior model probabilities of order 0.36 for M_{26728} and 0.24 for M_{26729} independently of both choices of the covariance matrices. As our simulation study, this application clearly illustrates the influence of the upper prior limit U for the df parameter ν on model selection results. In general a lower value of U is preferred since this allows to separate models

Table 4: Estimated posterior model probabilities $\hat{P}_k = \widehat{P}(M_k|\text{data})$ of all 32767 models for the data on Swiss counterfeit bank notes. The upper prior limit U for ν is equal to 20.

Model	Model indicator	$\hat{P}_k$	
		$\Sigma_1 = diag(10^2, 150^2)$	$\Sigma_2 = diag(0.1^2, 5^2)$
M_{26728}:	(1,1,0,1,0,0,0,0,1,1,0,1,0,0,0)	0.2731	0.2449
M_{26729}:	(1,1,0,1,0,0,0,0,1,1,0,1,0,0,1)	0.2427	0.2660
M_k		$\hat{P}_k < 0.0300$ for 372 Models	$\hat{P}_k < 0.0300$ for 427 Models
% of accepted birth-death moves		6.54%	9.08%

more.

Note that models M_{26728} and M_{26729} are actually in some sense very close. They only disagree on the conditional independence between "1" (banknote length) and "6" (diagonal length) given "2", "3", "4" and "5" and assume the same eight (conditional) independencies. In particular it seems that the height of the banknote ("3") on the right and the distance of inner frame to the lower border ("4") are independent. This could indicate on a nonstable (nonprofessional) process of frame stamping. To check this hypothesis, we have analyzed the data on genuine old Swiss 1000-franc bank notes and found that there is no unconditional independencies (i.e all unconditional copulas are present) even among the four most visited D-vine PCC models for any choice of the covariance matrices.

6 Conclusions and discussions

This paper considers a RJ MCMC algorithm for D-vine PCCs based on bivariate t copulas to simplify them by detecting (conditional) independence in data. The proposed RJ MCMC algorithm consists of only birth and death moves since so-called split and combine moves (see e.g. Richardson and Green, 1997) do not have sense in our setup. RJ MCMC iterates visit only those models which are important for data. It should be noted

that naive model selection for D-vine PCCs based on credibility interval from MCMC is always possible but our approach is more advanced and the first one, which considers RJ MCMC in the context of copulas. Further the proposed RJ MCMC algorithm estimates a D-vine PCC model and its parameters simultaneously.

The simulation study shows that the prior upper limit U for df ν has a significant impact on estimates of posterior model probabilities. The reason is that for large U a near identifiability problem between a $t-$copula and the independence copula can occur. Therefore we suggest to use $U = 20$, for which RJ MCMC identified true models of artificial data with high posterior model probability. Further we have observed that the results of RJ MCMC are robust with regard to the choice of proposal distribution for the birth move. We have used other bivariate truncated normal distributions with different magnitude of marginal variances and obtained very similar results. We have even employed a uniform distribution on $(-1, 1) \times (1, 20)$ and the corresponding results hardly deviate from the results presented here. As a rule of thumb we propose to use the covariance matrix $\Sigma = diag(1, 10^2)$ for $U = 20$ since the support intervals of ρ $(-1, 1)$ and of ν $(1, 20)$ are always covered by the corresponding 95% probability intervals of the $N(\hat{\rho}_s^{\mathrm{MLE}}, 1^2)$ and $N(\hat{\nu}_s^{\mathrm{MLE}}, 10^2)$ distributions whatever the MLE's $\hat{\rho}_s^{\mathrm{MLE}}$ and $\hat{\nu}_s^{\mathrm{MLE}}$ are. Min and Czado (2010a) observed that the full Bayesian estimation of D-vine PCCs is robust with respect to prior distributions. Therefore we believe that any reasonable choice of the proposal distribution and prior distributions for the proposed RJ MCMC algorithm will give similar results if there is one dominating model behind data. Finally note that computational time of our RJ MCMC algorithm for the presented analyses vary between 5 and 24 hours on a single core machine, which makes them practical and usable.

A richer family of multivariate copulas can be obtained by using pair-copulas from a catalogue of bivariate copulas including the independence copula. For model selection one can address now the choice of a copula family through the choice of each bivariate copula in a PCC. We envision here that appropriate RJ MCMC algorithms can be developed and implemented to address this model selection problem. Further recent econometric research work on copula models (e.g. Patton (2004)) show that copulas with parameters having time varying structure can capture the dependence structure better than copulas with constant parameters. Naturally a generalization for PCCs emerges here. These are

topics of future research.

Joint estimation of marginal and copula parameters has recently been found to be important. Thus Kim, Silvapulle, and Silvapulle (2007) have shown that a separate estimation of the marginal parameters may have an essential influence on the parameter estimation of multivariate copulas. Therefore inference based on joint estimates might be lead to quite different results compared to the inference ignoring estimation errors in the marginal parameters. For financial applications, one usually starts with multivariate time series and in a first step, one estimates for each marginal time series its structure as for example an ARMA or GARCH structure. In a second step, one determines standardized residuals, which are assumed to form an i.i.d marginal sample. Depending on whether the distribution of the residuals are known or unknown, one uses a parametric or empirical probability transform to transform to data with approximate uniform margins. This separates the marginal distribution from the dependence structure. In a final step, this dependency is modelled using a multivariate copula and copula parameter are estimated. The statistical properties of such two step estimation procedures are investigated by Joe (2005) for a known standardized residual distribution and by Chen and Fan (2006) for unknown standardized residual distribution, respectively. Czado, Gärtner, and Min (2010) have solved the above joint estimation problem in a Bayesian framework when marginal models are described by AR(1) models. Joint Bayesian estimation of more general marginal models such as ARMA-GARCH or stochastic volatility and PCC parameters is another topic of future research.

Acknowledgements

Claudia Czado acknowledges the support of the Deutsche Forschungsgemeinschaft (Cz 86/1-3). The authors are grateful to K. Aas and D. Berg for providing their R-package for likelihood calculations and to D. Pauler Ankerst for helpful discussions. They also thank G. Kuhn and N. Wermuth for providing the Euro swap rates and Swiss bank notes data, respectively. The authors would like to thank the Editor P. Gustafson, an Associate Editor and an anonymous referee for constructive comments and helpful remarks which led to a significant improved presentation of the manuscript.

References

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

Arakelian, V. and Dellaportas, P. (2009). Contagion tests via copula threshold models. Working Paper, Department of Statistics, Athens University of Economics and Business, page forthcomming in *Quantitative Finance*.

Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann. Math. Artif. Intell.*, **32**(1-4), 245–268.

Bedford, T. and Cooke, R. M. (2002). Vines–a new graphical model for dependent random variables. *Ann. Statist.*, **30**(4), 1031–1068.

Berg, D. and Aas, K. (2009). Models for construction of multivariate dependence. *European Journal of Finance*, **15**, 639–659.

Besag, J., Green, P., Higdon, D., and Mengersen, K. (1995). Bayesian computation and stochastic systems. *Statist. Sci.*, **10**(1), 3–66. With comments and a reply by the authors.

Chan, N.-H., Chen, J., Chen, X., Fan, Y., and Peng, L. (2009). Statistical inference for multivariate residual copula of GARCH models. *Statist. Sinica*, **19**, 53–70.

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

Congdon, P. (2006). Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. *Computational Statistics & Data Analysis*, **50**, 346 – 357.

Czado, C., Gärtner, F., and Min, A. (2010). Analysis of Australian electricity loads using joint Bayesian inference of D-vines with autoregressive margins. In D. Kurowicka and H. Joe, editors, *Dependence modelling: Handbook on Vine Copulae*. World Scientific, *forthcomming*.

Dalla Valle, L. (2009). Bayesian copulae distributions with application to operational risk management. *Methodol. Comput. Appl. Probab.*, **11**(1), 95–115.

Demarta, S. and McNeil, A. J. (2005). The t-copula and related copulas. *Internat. Statist. Rev.*, **73**(1), 111–129.

Embrechts, P., McNeil, A. J., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk management: value at risk and beyond (Cambridge, 1998)*, pages 176–223. Cambridge Univ. Press, Cambridge.

Embrechts, P., Lindskog, F., and McNeil, A. J. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of heavy tailed distributions in finance*. Elsevier/North-Holland, Amsterdam.

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

Flury, B. and Riedwyl, H. (1988). *Multivariate statistics : a practical approach*. Chapman & Hall, New York.

Frahm, G., Junker, M., and Szimayer, A. (2003). Elliptical copulas: applicability and limitations. *Statist. Probab. Lett.*, **63**(3), 275–286.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *N. Am. Actuar. J.*, **2**(1), 1–25.

Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, **12**, 347–368.

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

Genest, C., Gendron, M., and Bourdeau-Brien, M. (2009). The advent of copulas in finance. *European Journal of Finance*, **15**, (in press).

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

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

Huard, D., Évin, G., and Favre, A.-C. (2006). Bayesian copula selection. *Comput. Statist. Data Anal.*, **51**(2), 809–822.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf and B. Schweizer and M. D. Taylor, editor, *Distributions with Fixed Marginals and Related Topics*.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Joe, H. (2005). Asymptotic effiency of the two stage estimation method for copula-based models. *Journal of Multivariate Analysis*, **94**, 401–419.

Kim, G., Silvapulle, M. J., and Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Comput. Statist. Data Anal.*, **51**(6), 2836–2850.

Kurowicka, D. and Cooke, R. (2006). *Uncertainty analysis with high dimensional dependence modelling*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.

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

McNeil, A. J. and Neslehova, J. (2009). Multivariate Archimedean copulas, $d-$monotone functions and l_1 norm symmetric distributions. *Ann. Statist.*, **37**(5b), 3059–3097.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Phisics*, **21**, 1087–1092.

Min, A. and Czado, C. (2010a). Bayesian inference for multivariate copulas using pair-copula constructions. Working Paper, Zentrum Mathematik,Technische Universtät München*, forthcoming in The Journal of Financial Econometrics*.

Min, A. and Czado, C. (2010b). SCOMDY models based on pair-copula constructions with application to exchange rates. Working Paper, Zentrum Mathematik,Technische Universtät München*, submitted to Quantitative Finance*.

Nelsen, R. B. (1999). *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer-Verlag, New York.

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

Patton, A. J. (2009). Copula-based models for financial time series. In T.G. Andersen and R.A. Davis and J.-P. Kreiss and T. Mikosch, editor, *Handbook of Financial Time Series*.

Pitt, M., Chan, D., and Kohn, R. (2006). Efficient Bayesian inference for Gaussian copula regression models. *Biometrika*, **93**(3), 537–554.

Richardson, S. and Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *J. Roy. Statist. Soc. Ser. B*, **59**(4), 731–792.

Robert, C. and Marin, J. (2008). Some difficulties with some posterior probability approximations. *Bayesian Analysis*, **3**, 427–442.

Silva, R. d. S. and Lopes, H. F. (2008). Copula, marginal distributions and model selection: A Bayesian note. *Statistics and Computing*, **18**, 313–320.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.

Song, P. X.-K. (2000). Multivariate dispersion models generated from Gaussian copula. *Scand. J. Stat.*, **27**, 305–320.

Yule, G. U. and Kendall, M. G. (1965). *An Introduction to the Theory od Statistics*. Charles Griffin & Company, 14th edition.

Aleksey MIN
(**corresponding author)**
Zentrum Mathematik
Technische Universität München
Parkring 11
D-85748 Garching, Germany
E-mail: aleksmin@ma.tum.de
Tel.: +49 89 28917404
Fax: +49 89 28917407

Claudia CZADO
Zentrum Mathematik
Technische Universität München
Parkring 13
D-85748 Garching, Germany
E-mail: cczado@ma.tum.de
Tel.: +49 89 28917428
Fax: +49 89 28917435