Maximum likelihood estimation of mixed C-vines with application to exchange rates

CLAUDIA CZADO, ULF SCHEPSMEIER and ALEKSEY MIN
Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
85747 Garching bei München, Germany
corresponding author: U. Schepsmeier
email: schepsmeier@ma.tum.de

ABSTRACT

Multivariate copulas are commonly used in economics, finance and risk management. They allow for very flexible dependency structures, even though they are applied to transformed financial data after marginal time dependencies are removed. This is necessary to facilitate statistical parameter estimation. In this paper we consider a very flexible class of mixed C-vines, which allows the variables to be ordered according to their influence. Vines are build from bivariate copulas only and the term "mixed" refers to allowing the pair-copula family to be chosen individually for each term. In addition there are many C-vine structure specifications possible and therefore we propose a novel data driven sequential selection procedure, which selects both the C-vine structure and its attached pair-copula families with parameters. After the model selection ML estimation of the parameters is facilitated using the above found sequential estimates as starting values. An extensive simulation study shows a satisfactory performance of ML estimates in small samples. Finally an application involving US-exchange rates demonstrates the need for mixed C-vine models.

Keywords: copulas, C-vine, maximum likelihood estimation, pair-copula construction, US-exchange rates

1 Introduction

Pairwise construction principals have been very useful for building multivariate distributions. The first such pairwise construction was given by Joe (1996) based on the famous Sklar theorem (Sklar (1959)) using cumulative distribution functions (cdf). Bedford and Cooke (2001, 2002) realized that there were many such constructions possible, thus they organized them in graphical way by sequentially designing trees which identify the bivariate copula densities needed to make up a d-dimensional density. It involves only products of bivariate copulas, which we call pair-copulas. Since the trees are intrinsically related they called these distributions regular vines. Their primary interest was to use vines in the modeling of large networks so they restricted themselves to the case of Gaussian pair-copulas.
Aas et al. (2009) were the first to recognize that this construction principle can be extended by using arbitrary pair-copulas, since the construction principle has no restriction on the choice of pair-copulas. They developed standard maximum likelihood (ML) estimation for special vine copulas, where the challenge was to provide a good starting point for the required high dimensional optimization. Vine copulas are vine distributions with uniform margins. Regular vines include two simple tree structures, such as line trees and star trees, the first one corresponds to D-vines, while the second one corresponds to C-vines. Czado (2010) showed that C- and D-vines can be constructed by simple recursive conditioning frequently used in time series. Similar recursive conditioning arguments are used in time series.
Aas et al. (2009) used a sequential estimation procedure to provide starting values for the ML estimation in C- and D-vines. They utilized D-vines as a building block for a two step risk model for financial assets. The margins were estimated by standard ARMA-GARCH models and standardized residuals were formed. In a second step copula data is formed using a parametric or a non-parametric innovation specification. This allows to have an approximately i.i.d. sample for copula estimation. The parametric transformation was suggested by Joe (2005), while a rank based transformation was used by Genest et al. (1995). Both approaches follow a two step parameter estimation approach, first estimating marginal parameters and then copula parameters. To improve efficiency in the semi-parametric approach of Genest et al. (1995), Chen and Fan (2006) and Chan et al. (2009) based their inference for the copula parameters on the Kullback-Leibler information criteria. Later Min and Czado (2010c) extended their approach to D-vines and simple R-vine specifications using only t-copulas as pair-copulas. While such an approach allows for different symmetric tail behavior for pairs of variables, it does not allow for non-symmetric tail behavior such as provided by a Clay-

ton or Gumbel copula. This is the starting point for this paper. We want to allow for different pair-copula families and concentrate on C-vines, which have not been investigated in applications so far.
Such mixed C-vines have several selection problems attached to them. First we need to select the appropriate C-vine structure, since there exist as for example Aas et al. (2009) showed $d!/2$ different C-vines and additionally to choose from a catalogue of pair-copula families for each required pair-copula. At the moment only selection procedures within specified D-vine structures exist. Min and Czado (2010b) use reversible jump MCMC to simplify a D-vine with specific single pair-copula family by discovering conditional independences, while Smith et al. (2010) use indicator variables for identifying conditional independence in a Bayesian setup.
The goal of this paper is to provide a comprehensive solution to the selection of C-vines by identifying an appropriate C-vine structure and selecting a fitting pair-copula family. To accomplish this a sequential approach is developed based on the cardinality of the conditioning variables in conjunction with individual choices for each pair-copula as a best fitting pair-copula family from a large catalogue of families. The catalogue of pair-copula families includes elliptical copulas such as Gaussian and t-copulas, single parameter Archimedean copulas such as Gumbel and Clayton, as well as two parameter families such as BB1 and BB7 of Joe (1997). Finally the Joe and Frank copula are included as well. For the copula family selection a goodness-of-fit procedure studied in Genest et al. (2006), Genest et al. (2009) and Berg (2009) is used, as well as model comparison tests based on Vuong (1989) and Clarke (2007) suitable for non-nested model comparison. In addition, scatter and contour plots, as well as plots of the λ-function introduced by Genest and Rivest (1993) are evaluated.
The selection of a C-vine structure and its pair-copulas is determined by developing an appropriate sequential estimation procedure which is used as starting value for the ML estimation. A large simulation study shows very good small sample performance of the ML estimation in mixed C-vines. Finally the usefulness of these models is demonstrated in an application involving US-exchange rates by using model comparison criteria as AIC, BIC, Vuong and Clarke tests suitable to compare joint copula models.
In summary the main contributions of this paper are

1. Development and implementation of sequential and ML estimation procedures for copula parameters in a previously specified mixed C-vine.

2. Development of a data driven sequential selection procedure for jointly choosing the C-vine structure and pair-copula families.

The paper is organized as follows: In Section 2 we discuss bivariate copula families, while Section 3 introduces the mixed C-vine copula model. Here also sequential and ML estimation of copula parameters are studied. In Section 4 the data driven sequential procedure for jointly selecting the C-vine structure with pair-copula families is developed. In Section 5 a simulation study investigates the small sample performance of the ML procedure for mixed C-vines. A second study in Section 6 demonstrates the satisfactory performance of the C-vine selection method. A mixed C-vine is chosen and investigated in Section 7 to model the dependencies among the standardized residuals in US-exchange rates. The paper closes with a summary and discussion section.

2 Bivariate copula families

A d-dimensional copula is a multivariate distribution function $C(u_1, \ldots, u_d)$ defined on the unit cube $[0,1]^d$, with uniformly distributed marginals. It can be used to characterize the dependency between d random variables, while allowing for arbitrary marginal distributions. In the next section we develop multivariate copulas using only bivariate copulas as building blocks, therefore we concentrate here on $d = 2$. In particular the famous theorem of Sklar (1959) gives the connection between marginals and copula to the joint distribution. For this let $F(\cdot, \cdot)$ denote a bivariate cdf with marginal cdf's F_1 and F_2, respectively, then there exists a two dimensional copula cdf $C(\cdot, \cdot)$, such that for all $(x_1, x_2) \in \overline{\mathbb{R}}^2$

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

holds. For continuous F_1 and F_2, $C(\cdot, \cdot)$ is unique and is defined through

$$C(x_1, x_2) = F(F_1^{-1}(x_1), F_2^{-1}(x_2)).$$

If F is in addition absolutely continuous then the copula density is well defined and given by

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}.$$

For details see Joe (1997) and Nelsen (2006).
The most important and most commonly used copulas in finance are the Gaussian and the t-copula. Both belong to the class of elliptical copulas and for a precise definition see for example Frahm et al. (2003), who discuss their applicability and limitations. Another class often discussed and utilized are Archimedean copulas; see for example Embrechts et al. (2003) or Nelsen

(2006). Non-linear dependence is often measured using Kendall's τ (see e.g. Kruskal (1958)), while dependence in the tail is measured by upper and lower tail dependence coefficients (see e.g. Joe (1997)) as

$$\lambda_U = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u} = \lim_{u \nearrow 1} P(U_1 > u | U_2 > u)$$

and

$$\lambda_L = \lim_{u \searrow 0} \frac{C(u,u)}{u} = \lim_{u \searrow 0} P(U_1 \leq u | U_2 \leq u).$$

While the Gaussian copula has $\lambda_U = \lambda_L = 0$, the tail dependence coefficients for the t-copula are symmetric, i.e.

$$\lambda = \lambda_U = \lambda_L.$$

The Archimedean Clayton and Gumbel copula are reflection-asymmetric. They allow either for lower by no upper tail dependence and vise versa. In particular for the Clayton copula we have $\lambda_L > 0$, but $\lambda_U = 0$ and for the Gumbel copula $\lambda_L = 0$ but $\lambda_U > 0$. Therefore Joe (1997, Section 5.2) introduced two bivariate copula families called BB1 and BB7, respectively, which allow for different $\lambda_L > 0$ and $\lambda_U > 0$ simultaneously. We will utilize these bivariate copula families as possible building blocks in our mixed C-vine. Since they are not so commonly discussed, we summarize their definition and properties in Table 1. In addition top panel of Figure 1 visualizes a scatter plot of a bivariate random sample of size $n = 1000$ from a BB1 (left top panel) and a BB7 (right top panel) copula distribution. In the lower panels the corresponding contour plots of the transformed vector $\mathbf{Z} = (\Phi^{-1}(U_1), \Phi^{-1}(U_2))$ are given, where (U_1, U_2) follows a BB1 and a BB7 copula distribution, respectively. Here $\Phi^{-1}(\cdot)$ denotes the quantile function of a standard normal random variable. Note that $\mathbf{Z}$ has standard normal margins and the copula dependence structure is not changed under these monotone transformations.

3 Pair-copula construction (PCC) of mixed C-vines

Using two-dimensional copulas it is possible to construct general multivariate distributions by specifying the dependence and conditional dependence of selected pairs of random variables and all marginal distribution functions. We will define such a construction in this section. Our presentation follows Aas et al. (2009) and Czado (2010), but the idea was first developed by Joe (1996) for cdf's and organized graphically using densities and a sequence of

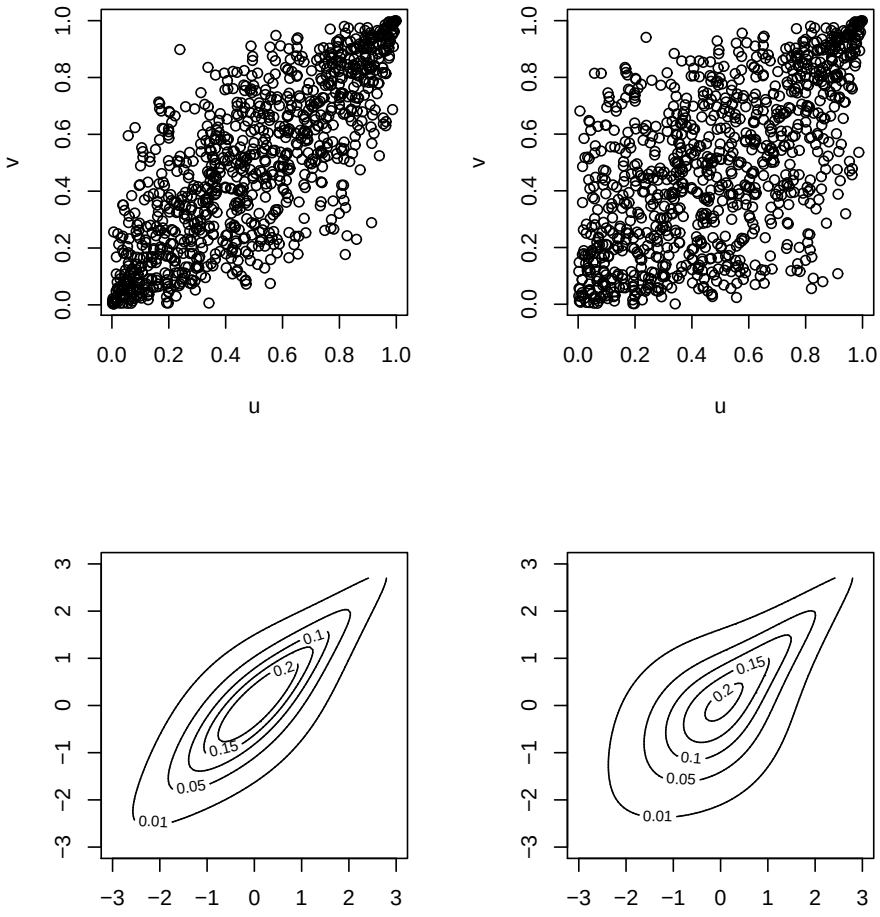


Figure 1: Top row: scatter plot BB1 (left) with $\theta = 0.3$ and $\delta = 2.1(\tau = 0.59, \lambda_L = 0.33, \lambda_U = 0.61)$ and BB7 (right) with $\theta = 2.1$ and $\delta = 0.3(\tau = 0.43, \lambda_L = 0.72, \lambda_U = 0.61)$. Bottom row: Contour plots of Z transformed from BB1 with $\theta = 0.3$ and $\delta = 2.1$ (left) and BB7 with $\theta = 2.1$ and $\delta = 0.3$ (right).

BB1	
cdf	$C(u,v;\theta,\delta) = \left\{1 + [(u^{-\theta}-1)^{\delta} + (v^{-\theta}-1)^{\delta}]^{\frac{1}{\delta}}\right\}^{-\frac{1}{\theta}}$
	$\theta > 0, \delta \geq 1$
density	$c(u,v;\theta,\delta) = \left\{1 + [(u^{-\theta}-1)^{\delta} + (v^{-\theta}-1)^{\delta}]^{\frac{1}{\delta}}\right\}^{-\frac{1}{\theta}-2}$
	$\times[(u^{-\theta}-1)^{\delta} + (v^{-\theta}-1)^{\delta}]^{\frac{2}{\delta}-2}$
	$\times\{\theta\delta + 1 + \theta(\delta-1)[(u^{-\theta}-1)^{\delta} + (v^{-\theta}-1)^{\delta}]^{-\frac{1}{\delta}}\}$
	$\times(u^{-\theta}-1)^{\delta-1}u^{-\theta-1}(v^{-\theta}-1)^{\delta-1}v^{-\theta-1}$
Kendall's τ	$\tau = 1 - \frac{2}{\delta(\theta+2)}$
upper/lower tail dependence	$\lambda_U = 2 - 2^{1/\delta}$
	$\lambda_L = 2^{-1/(\delta\theta)}$
BB7	
cdf	$C(u,v;\theta,\delta) =$
	$1 - \left(1 - [(1-(1-u)^{\theta})^{-\delta} + (1-(1-v)^{\theta})^{-\delta} - 1]^{-\frac{1}{\delta}}\right)^{\frac{1}{\theta}}$
	$\theta \geq 1, \delta > 0$
density	$c(u,v;\theta,\delta) = (-\frac{1}{\theta})(\frac{1}{\delta}-1)\cdot h^{\frac{1}{\theta}-2}dvh\cdot duh - \frac{1}{\theta}\cdot h^{\frac{1}{\theta}-1}duvh$
	$h = 1 - ((1-(1-u)^{\theta})^{-\delta} + (1-(1-v)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}}$
	$duh = -\theta((1-(1-u)^{\theta})^{-\delta} - (1-(1-v)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}-1}$
	$\times(1-(1-u)^{\theta})^{-\delta-1}(1-u)^{\theta-1}$
	$dvh = -\theta((1-(1-u)^{\theta})^{-\delta} - (1-(1-v)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}-1}$
	$\times(1-(1-v)^{\theta})^{-\delta-1}(1-v)^{\theta-1}$
	$duvh = \frac{1}{\delta}(-\frac{1}{\delta}-1)((1-(1-u)^{\theta})^{-\delta}$
	$-(1-(1-v)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}-2}\cdot duS\cdot dvS$
	$duS = -\theta\delta(1-(1-u)^{\theta})^{-\delta-1}(1-u)^{\theta-1}$
	$dvS = -\theta\delta(1-(1-v)^{\theta})^{-\delta-1}(1-v)^{\theta-1}$
Kendall's τ	$\tau = 1 - \frac{2}{\delta(2-\theta)} + \frac{4}{\theta^2\delta}B(\frac{2-2\theta}{\theta}+1, \delta+2)$
	with Beta-function $B(x,y) = \int_0^1 t^{x+1}(1-t)^{y-1}dt$
upper/lower tail dependence	$\lambda_U = 2 - 2^{1/\theta}$
	$\lambda_L = 2^{-1/\delta}$

Table 1: Cdf, density, Kendall's τ, upper and lower tail dependence coefficients of the BB1 and BB7 copula, respectively.

nested trees by Bedford and Cooke (2002). The class of multivariate copulas constructed in such a way are called regular vines. The so called C- and D-vines are subclasses of regular vines.

Our general assumption is that all joint, marginal and conditional distributions are absolutely continuous with corresponding densities. Under this regularity condition, Czado (2010) showed that a multivariate density can be constructed as a product of pair-copulas, acting on several different conditional probability distributions. Given the recursive decomposition of the conditional distribution of (X_{t-1}, X_t) given $X_1, \ldots, X_{t-2}$ as

$$f(x_t|x_1, \ldots, x_{t-1}) = c_{t-1,t|1,\ldots,t-2} \cdot f(x_t|x_1, \ldots, x_{t-2}), \tag{3.1}$$

for $t = 2, \ldots, d$, Czado (2010) showed that the joint distribution is

$$f(x_1, \ldots, x_d) = \prod_{k=1}^{d} f(x_k) \times \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+i|1,\ldots,j-1}, \tag{3.2}$$

where she used the following abbreviation for a bivariate conditional copula density of X_i and X_j given $x_{i_1}, \ldots, x_{i_k}$

$$c_{i,j|i_1,\ldots,i_k} := c_{i,j|i_1,\ldots,i_k}(F(x_i|x_{i_1}, \ldots, x_{i_k}), F(x_j|x_{i_1}, \ldots, x_{i_k}))$$

for arbitrary distinct indices $i, j, i_1, \ldots, i_k$ with $i < j$ and $i_1 < \ldots < i_k$. Here $f(\cdot|\cdot)$ and $F(\cdot|\cdot)$ denote conditional densities and distribution functions, respectively. According to Bedford and Cooke (2002) this PCC (3.2) is called a canonical vine distribution or short C-vine.
For $d = 4$ the C-vine density (3.2) can, for example, be written as

$$f(x_1, x_2, x_3, x_4) = \prod_{k=1}^{d} f(x_k) \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{23|1} \cdot c_{24|1} \cdot c_{34|12}. \tag{3.3}$$

Both in (3.2) and (3.3) the choice for the bivariate pair-copulas $c_{ij|i_1,\ldots,i_k}$ is completely arbitrary and in this paper we allow for an individual choice for each of these pair-copulas from a catalogue of copula families. We call such a C-vine a mixed C-vine. It is also clear that the construction is iterative by nature, and that given a specific factorization in (3.2), there are many different orderings of the variables yielding different C-vines.
In (3.1) we implicitly assume that the bivariate conditional copula does not depend on the specific values of the conditioning variables, other than through its arguments given by conditional distribution functions. This restriction is however not so severe, see Haff et al. (2010).
It should be mentioned that even with this restriction the class of possible C-vines covers a huge range of dependencies among variables, since we are allowed to use arbitrary pair-copula families in (3.2), thus extending the range

of available multivariate distributions enormously. In this paper we therefore want to find the best approximating C-vine model out of this class, while maintaining tractability of estimation and model selection. This is currently no longer the case if one wants to allow for additional dependencies on the value of the conditioning variables.
We note that under the assumption of no further dependency on $\boldsymbol{x}_{1:(j-1)} = (x_1, \ldots, x_{j-1})$ of the pair-copula term
$c_{j,j+i|1,\ldots,j-1}(F(x_j|\boldsymbol{x}_{1:(j-1)}), F(x_{j+i}|\boldsymbol{x}_{1:(j-1)}))$ than through the two arguments, the corresponding multivariate density in (3.2) does not need any further restrictions on the pair-copula specifications. If a Gaussian bivariate copula with correlation parameter $\rho_{j,j+i|1,\ldots,j-1}$ for each $c_{j,j+i|1,\ldots,j-1}$ is chosen, the parameter $\rho_{j,j+i|1,\ldots,j-1}$ can be interpreted as a partial correlation of the variables X_j and X_{j+i} given the variables $X_1, \ldots, X_{j-1}$. Partial correlations of Gaussian variables are unrestricted in $[-1, 1]$ and always induce a positive definite correlation matrix (Kurowicka and Cooke (2006, pp 101) and Kurowicka and Cooke (2001, Prop 3.19)). This means that a C-vine with all Gaussian pair-copulas is just a multivariate Gaussian distribution. The above property shows the advantage of using a C-vine representation of multivariate Gaussian variables. Similarly, a multivariate t-distribution with a common degree of freedom can be represented as a C-vine with all t-pair-copulas with some restrictions on the degrees of freedom for each t-pair-copula (Min and Czado (2010c)).
Here we concentrate on C-vines, however a similar argument can be used to derive the class of D-vines (see Czado (2010)). C-vines are especially useful, when there exists a variable order with sequentially decreasing driving force. More precisely the order starts with a variable X_{i_1} which has the highest dependency with all other variables. Now conditioning all remaining variables on X_{i_1}, the next variable X_{i_2} has highest dependency with all other variables. Similarly, variable X_{i_3} has highest dependency with all variables conditional on variables X_{i_1} and X_{i_2}, etc. In (3.2) one observes that the variable X_{i_1} is mostly involved in all pair-copulas, then the variable X_{i_2} is second mostly involved in all pair-copulas and so on. Thus the above described behavior is captured by (3.2) with the order $X_{i_1}, X_{i_2}, \ldots, X_{i_{d-2}}$.
For the pair-copula construction marginal conditional distributions of the form $F(x|\mathbf{v})$ are needed. For every v_j in the vector $\mathbf{v}$ $F(x|\mathbf{v})$ can be written as

$$F(x|\mathbf{v}) = \frac{\partial C_{x,v_j|\mathbf{v}_{-j}}\{F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})\}}{\partial F(v_j|\mathbf{v}_{-j})} \tag{3.4}$$

with $C_{x,v_j|\mathbf{v}_{-j}}$ an arbitrary bivariate copula cdf (see Aas et al. (2009)).
Since we will apply representation (3.2) to copula data $\mathbf{u} = (u_1, \ldots, u_d)^t$ on

Copula	***h*-function**
Gaussian	$h(u\|v;\rho) = \Phi\left(\frac{\Phi^{-1}(u)-\rho\Phi^{-1}(v)}{\sqrt{1-\rho^2}}\right)$
t	$h(u\|v;\rho,\nu) = t_{\nu+1}\left\{\frac{t_\nu^{-1}(u)-\rho t_\nu^{-1}(v)}{\sqrt{\frac{(\nu+(t_\nu^{-1}(v))^2)(1-\rho^2)}{\nu+1}}}\right\}$
BB1	$h = \left(1+((u^{-\theta}-1)^\delta+(v^{-\theta}-1)^\delta)^{\frac{1}{\delta}}\right)^{-\frac{1}{\theta}-1}$ $\cdot((u^{-\theta}-1)^\delta+(v^{-\theta}-1)^\delta)^{\frac{1}{\delta}-1}(v^{-\theta}-1)^{\delta-1}v^{-\theta-1}$
BB7	$h = \left(1-[(1-(1-u)^\theta)^{-\delta}+(1-(1-v)^\theta)^{-\delta}-1]^{-\frac{1}{\delta}}\right)^{\frac{1}{\theta}-1}$ $\cdot[(1-(1-u)^\theta)^{-\delta}+(1-(1-v)^\theta)^{-\delta}-1]^{-\frac{1}{\delta}-1}$ $\cdot(1-(1-v)^\theta)^{-\delta-1}(1-v)^{\theta-1}$

Table 2: *h*-functions of the Gaussian, the t-, the BB1 and BB7 copula

$[0,1]^d$ we denote (3.2) in this case as C-vine copula density. Note that in this case $f(u_i) = 1$ and $F(u_i) = u_i \ \forall i = 1,\dots,d$. In the following we assume a parametric specification for $C_{i,j|i_1,\dots,i_k}$ given by an appropriate parameter (vector) $\boldsymbol{\theta}$. For parametric pair-copula densities and univariate conditioning set (3.4) simplifies to

$$h(u|v;\boldsymbol{\theta}) := F(u|v;\boldsymbol{\theta}) = \frac{\partial C_{u,v}(u,v;\boldsymbol{\theta})}{\partial v}, \tag{3.5}$$

where $\boldsymbol{\theta}$ is the parameter vector for $C_{u,v}$.
Table 2 gives the *h*-functions of the Gaussian, the t-, the BB1 and the BB7 copula, respectively. Here ρ is the parameter of the Gaussian copula and $\Phi^{-1}(\cdot)$ is the inverse of the standard normal distribution function. Further $t_\nu^{-1}(\cdot)$ in the t-copula case is the quantile-function of the univariate standard t distribution with ν degrees of freedom, expected value 0 and variance $\frac{\nu}{\nu-2}$ for $\nu > 2$.

To illustrate the usefulness of (3.4) and (3.5) we derive the conditional cdf $F(u_3|u_1,u_2;\boldsymbol{\theta}_{23|1},\boldsymbol{\theta}_{12},\boldsymbol{\theta}_{13})$ needed as argument for $c_{34|12}$ in a 4 dimensional C-vine copula density (compare to (3.3)). First (3.4) implies that

$$F(u_3|u_1,u_2;\boldsymbol{\theta}_{23|1},\boldsymbol{\theta}_{12},\boldsymbol{\theta}_{13}) = \frac{\partial C_{u_3,u_2|u_1}(F(u_3|u_1;\boldsymbol{\theta}_{13}),F(u_2|u_1;\boldsymbol{\theta}_{12});\boldsymbol{\theta}_{23|1})}{\partial F(u_2|u_1;\boldsymbol{\theta}_{12})}$$

holds. Now using (3.5)

$$F(u_3|u_1;\boldsymbol{\theta}_{13}) = h(u_3|u_1;\boldsymbol{\theta}_{13}) \text{ and } F(u_2|u_1;\boldsymbol{\theta}_{12}) = h(u_2|u_1;\boldsymbol{\theta}_{12}),$$

it follows that

$$F(u_3|u_1,u_2;\boldsymbol{\theta}_{23|1},\boldsymbol{\theta}_{12},\boldsymbol{\theta}_{13}) = h(h(u_3|u_1;\boldsymbol{\theta}_{13})|h(u_2|u_1;\boldsymbol{\theta}_{12});\boldsymbol{\theta}_{23|1}).$$

Here $\boldsymbol{\theta}_{13},\boldsymbol{\theta}_{12}$ and $\boldsymbol{\theta}_{23|1}$ denote the parameters of c_{13},c_{12} and $c_{23|1}$, respectively. Therefore higher order conditioning requires recursive application of appropriate h-functions.

We now turn to parameter estimation in C-vines. Let $\boldsymbol{\theta}_i$ be the vector of all pair-copulas in (3.2), which involve a conditioning set of cardinality i for $i = 0,\ldots,d-2$. For the C-vine in (3.3), $\boldsymbol{\theta}_0$ contains the parameters of the pair-copulas c_{12},c_{13} and c_{14}, $\boldsymbol{\theta}_1$ the parameters of $c_{23|1}$ and $c_{24|1}$ and $\boldsymbol{\theta}_2$ the parameters of $c_{34|12}$. More generally $\boldsymbol{\theta}_i = (\boldsymbol{\theta}_{1i}^t,\ldots,\boldsymbol{\theta}_{d-(i+1),i}^t)^t$, where $\boldsymbol{\theta}_{j,i}$ is the parameter vector of $c_{i+1,j+i+1|1,\ldots,i}$. Finally we denote by $\boldsymbol{\theta} = (\boldsymbol{\theta}_0^t,\ldots,\boldsymbol{\theta}_{d-2}^t)^t$ the set of all parameters to be estimated.

We present two estimation methods, one is a sequential estimator (SE) and the other one is the maximum likelihood estimator (MLE). Suppose i.i.d. data $\mathbf{u}_t := (u_{1,t},\ldots,u_{d,t})^t$ for $t = 1,\ldots,T$ is available. For SE the parameters of unconditional copulas are first estimated then they are used to estimate parameters of pair-copulas with single conditioning variable. These estimates will then be used for estimation of pair-copula parameters with two conditioning variables. We proceed sequentially until all parameters are estimated.

More precisely for $\boldsymbol{\theta}_{j,0}$ of $c_{1,j+1}$ estimate $\boldsymbol{\theta}_{j,0}$ based on data $(u_{1,t},u_{j+1,t}), t = 1,\ldots,T$ for $j = 1,\ldots,d-1$. For one-parameter families with a known relationship to Kendall's τ, one can invert the empirical Kendall's τ based on $(u_{1,t},u_{j+1,t}), t = 1,\ldots,T$. Alternatively one can maximize the corresponding bivariate likelihood to get $\hat{\boldsymbol{\theta}}_{j,0}^S$ for $j = 1,\ldots,d-1$.

In the next step we want to estimate $\boldsymbol{\theta}_{j,1}$ corresponding to $c_{2,j+2|1}$ for $j = 1,\ldots,d-2$. Define

$$\hat{v}_{2|1,t} := F(u_{2,t}|u_{1,t};\hat{\boldsymbol{\theta}}_{1,0}^S) = h(u_{2,t}|u_{1,t};\hat{\boldsymbol{\theta}}_{1,0}^S)$$
$$\hat{v}_{j+2|1,t} := F(u_{j+2,t}|u_{1,t};\hat{\boldsymbol{\theta}}_{j+1,0}^S) = h(u_{j+2,t}|u_{1,t};\hat{\boldsymbol{\theta}}_{j+1,0}^S)$$

for $j = 1,\ldots,d-2$. Use data $(\hat{v}_{2|1,t},\hat{v}_{j+2|1,t})$, $t = 1,\ldots,T$ to estimate $\boldsymbol{\theta}_{j,1}$ for $j = 1,\ldots,d-2$ and denote these estimates by $\hat{\boldsymbol{\theta}}_{j,1}^S$. For $\boldsymbol{\theta}_{j,2}$ corresponding to $c_{3,j+3|1,2}$ for $j = 1,\ldots,d-3$, define

$$\hat{v}_{3|1,2,t} := h(\hat{v}_{3|1,t}|\hat{v}_{2|1,t};\hat{\boldsymbol{\theta}}_{1,1}^S)$$
$$\hat{v}_{j+3|1,2,t} := h(\hat{v}_{j+3|1,t}|\hat{v}_{2|1,t};\hat{\boldsymbol{\theta}}_{j+1,1}^S)$$

and estimate $\boldsymbol{\theta}_{j,2}$ based on $(\hat{v}_{3|1,2,t}, \hat{v}_{j+3|1,2,t})$, $t = 1, \ldots, T$ for $j = 1, \ldots, d-3$. For general $\boldsymbol{\theta}_{j,i}, i = 1, \ldots, d-2$ base estimation on

$$\hat{v}_{i+1|1,\ldots,i,t} := h(\hat{v}_{i+1|1,\ldots,i-1,t}|\hat{v}_{i|1,\ldots,i-1,t}; \hat{\boldsymbol{\theta}}^S_{1,i-1})$$
$$\hat{v}_{j+i+1|1,\ldots,i,t} := h(\hat{v}_{j+i+1|1,\ldots,i-1,t}|\hat{v}_{i+1|1,\ldots,i-1,t}; \hat{\boldsymbol{\theta}}^S_{j+1,i-1})$$

for $t = 1, \ldots, T$ and $j = 1, \ldots, d-(i+1)$. These sequential estimates $\boldsymbol{\theta}^S_{j,i}, i = 0, \ldots, d-2$, $j = 1, \ldots, d-2$ can be used on their own, they are asymptotically normal under regularity conditions as recently shown by Haff (2010). However their asymptotic covariance expression is intractable. To improve efficiency one can use MLE's, which require high dimensional optimization of the log-likelihood. Here the sequential estimates can be used as starting values for the optimization.

We now turn to the determination of MLE's of the parameters of a C-vine copula distribution. For this the corresponding representation (3.2) can be used to construct the log-likelihood for an i.i.d. d-variate copula sample $\mathbf{u}_t = (u_{1,t}, \ldots, u_{d,t})^t$ for $t = 1, \ldots, T$. Let $\mathbf{u} = (\mathbf{u}_1^t, \ldots, \mathbf{u}_T^t)^t$ and $\boldsymbol{\theta}$ the parameter vector to be estimated, then the log-likelihood can be written as

$$\begin{aligned} l(\boldsymbol{\theta}, \mathbf{u}) = \sum_{t=1}^{T} \Bigg[& \sum_{j=1}^{d-1} \log\{c(u_{1,t}, u_{j+1,t}; \boldsymbol{\theta}_{j,0})\} \\ & + \sum_{i=1}^{d-2} \sum_{j=1}^{d-(i+1)} \log\{c(v_{i+1|1,\ldots,i,t}, v_{j+i+1|1,\ldots,i,t}; \boldsymbol{\theta}_{j,i})\} \Bigg] \end{aligned} \tag{3.6}$$

where

$$v_{j+2|1,t} := h(u_{j+2,t}|u_{1,t}; \boldsymbol{\theta}_{j+1,0}) \quad j = 0, \ldots, d-2 \tag{3.7}$$

$$\begin{aligned} v_{j+i+1|1,\ldots,i,t} := h(v_{i+j+1|1,\ldots,i-1,t}|v_{i+1|1,\ldots,i-1}; \boldsymbol{\theta}_{j+1,i-1}) \\ i = 1, \ldots, d-2, j = 1, \ldots, d-(i+1). \end{aligned} \tag{3.8}$$

The log-likelihood (3.6) together with definition (3.7) and (3.8) can now be numerically optimized using the sequential estimates $\hat{\boldsymbol{\theta}}^S_{j,i}$ as starting values. We denote the MLE's by $\hat{\boldsymbol{\theta}}^{MLE}_{j,i}$ to distinguish them from $\hat{\boldsymbol{\theta}}^S_{j,i}$.

4 Selection of vine copula models

As already noted there are many different orderings of the variables in C-vine models possible. Aas et al. (2009) have shown that there exactly $d!/2$ different C-vines and thus C-vine copulas available. In a mixed C-vine copula

model we need in addition to choose a bivariate copula family for each of the $d(d-1)/2$ pair-copulas. We will now consider these selection problems.
As noted in Aas et al. (2009) it is preferable to choose models with high dependence in the bivariate conditional distribution characterized by $c_{i,j|i_1,\ldots,i_k}$, where the number of conditioning variables k is small. This suggests a data driven sequential approach starting with determining the $d-1$ unconditional pair-copulas needed in a C-vine copula. For this estimate all pairwise Kendall's $\tau_{i,j}$ values by $\hat{\tau}_{i,j}$ and find the variable i^* which maximizes

$$\hat{S}_i := \sum_{j=1}^{d} |\hat{\tau}_{i,j}| \tag{4.1}$$

over $i = 1, \ldots, d$. Here we set $\hat{\tau}_{i,i} = 1$ for $i = 1, \ldots, d$. To ease notation we reorder the variables in such a way that the first variable is now i^*. For this reordering $c_{1,j+1}, j = 1, \ldots, d-1$ are selected as unconditional pair-copulas. We call variable 1 also the root of all unconditional pair-copulas. Before determining the pair-copulas with the single conditioning variable 1, a choice of the pair-copula family and its parameter value for $c_{1,j+1}$ for $j = 1, \ldots, d-1$ has to be made. We will discuss this choice later and assume at this point that we are able to choose a pair-copula family with parameter estimate $\hat{\boldsymbol{\theta}}^S_{j,0}$ for $c_{1,j+1}$ for $j = 1, \ldots, d-1$. As in the sequential estimation procedure $d-1$ transformed variables

$$\hat{v}_{j+2|1,t} := h(u_{j+2,t}|u_{1,t}; \hat{\boldsymbol{\theta}}^S_{j+1,0}) \quad j = 0, \ldots, d-2, t = 1, \ldots, T \tag{4.2}$$

are defined. Again $(d-1)$ data samples of size T are used and all pairwise Kendall's τ values are estimated and we can find the corresponding maximum as in (4.1) based now $d-1$ variables. Assume that this maximum is obtained at i^{**} and again reorder the variables $i = 2, \ldots, d$ in such a way that i^{**} is now variable 2. We select now $c_{2,j+2|1}$ for $j = 1, \ldots, d-2$ as pair-copulas with single conditioning variable 1. Here $2 = i^{**}$ can be considered as root of $c_{2,j+2|1}$ for $j = 1, \ldots, d-2$. We continue now with this procedure with transformed variables as defined in (4.2) until we have all pair-copulas and their sequential estimates $\hat{\boldsymbol{\theta}}^S_{j,s}$ determined. Note that this sequential procedure both determines the C-vine copula structure and corresponding parameter estimates.
We now consider the problem of choosing the copula family. This has been a well studied problem and many procedures have been suggested. Note that for the sequential selection procedure we only require a copula selection in two dimensions. Copula goodness-of-fit tests have been studied by Genest et al. (2009), Genest et al. (2006) and Berg (2009). One recommended test

is based on the Cramér-von-Mises statistics, which we will utilize later. It is implemented in the R-package "copula". Genest et al. (2009) introduced and studied a λ-function based on the Kendall's process. This λ-function is especially easy for Archimedean copulas.
Joe (1997, Section 10.3) proposed a model selection based on the Akaike information criterion (AIC) of Akaike (1974).

$$AIC := -2\sum_{i=1}^{n} \log f(x_i; \hat{\boldsymbol{\theta}}) + 2k,$$

where $\hat{\boldsymbol{\theta}}$ denotes the estimate of $\boldsymbol{\theta}$ and k is the number of parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$ in the model. Specifying the AIC to a specific copula with density c we get

$$AIC := -2\sum_{i=1}^{n} \log c(u_{i1}, u_{i2}; \hat{\boldsymbol{\theta}}) + 2k,$$

which can be used as a copula selection criterion. The advantage of this selection method is that it can be automatized in a copula selection program. Additionally we also look at likelihood ratio based tests suggested by Vuong (1989) and Clarke (2007) suitable for non-nested model comparison. Finally bivariate scatter plots of the copula data and empirical contour plots of the transformed copula data with normal margins are examined.
For our implementation for the copula family choice we consider the Gaussian, t-, Clayton, Gumbel, Frank, BB1 and BB7 copula family, which cover a wide range of dependence behavior.
Since the Vuong and Clarke tests require two specified copula models to be compared to, we first specify a copula model, denoted by A, we conduct then tests comparing A to any other copula family considered. A score for A is determined as follows; each time model A is preferred to another model, the score is increased by 1. If model A cannot be distinguished from the other model, the score is left unchanged. If the other model is preferred to A, then 1 is subtracted from the score. This scoring is done for considered copulas and the model with highest score is chosen. We note that the copula goodness-of-fit tests have a more general alternative, while the Vuong and Clarke tests consider a single parametric copula in the alternative.

Scenario	τ_{12}	τ_{13}	τ_{14}	$\tau_{23\|1}$	$\tau_{24\|1}$	$\tau_{34\|12}$
1	H	H	H	H	H	H
2	L	L	L	L	L	L
3	H	H	H	L	L	L
4	L	L	L	H	H	H

Table 3: Choice of conditional Kendall's τ for 4 scenarios in a four dimensional C-vine copula model (H $\hat{=}\{\tau = 0.8\}, L\hat{=}\{\tau = 0.2\}$)

5 Small sample performance of the ML estimates in a fully specified C-vine with different pair-copula types

We investigated the accuracy, stability and robustness of the maximum likelihood estimation described in the previous section by performing a small simulation study in R for different sample sizes, different C-vine dimensions and different copula family combinations. For this we used the R-routines written by Schepsmeier (2010), which contains the MLE algorithm as well as a routine for the simulation of C-vines (see Aas et al. (2009)).
The advantage of this R-package is that mixed C-vine copula models can be used. Unlike Aas et al. (2009), Min and Czado (2010a) or Nikoloulopoulos et al. (2011) one is not limited to simple models with only one copula family for all pair-copulas in a C-vine PCC. In this paper we give a summary of our simulation study of the ML estimates in a fully specified C-vine. Note that in this simulation study the C-vine and copula selection methods are not used, i.e. the C-vine structure as well as the pair-copula families are fixed. Detailed results of this small sample performance study can be found in Appendix C of Schepsmeier (2010). First we restrict to a four dimensional C-vine copula set. In this case six pair-copulas have to be chosen. As a common measure of dependence across pair-copula families we used Kendall's τ. Two choices, $\tau = 0.2(L)$ and $\tau = 0.8(H)$, are investigated. To be more precise, the four scenarios of Table 3 are determined.
Possible pair-copula families were Gaussian (N), t-copula with ν degrees of freedom (t_ν), Clayton (C), Gumbel (G), Frank (F), Joe (J), BB1 and BB7. As sample size $n = 500, 1000$ and 2000 were investigated. In a first setup we used the same pair-copula family for each of the six pair-copulas, while in a second setup we allowed for mixed C-vine copulas and investigated a total of 8 combinations given in Table 4.
For the one parameter pair-copulas the choice of Kendall's τ determines the corresponding parameter. For the two parameter families BB1 and BB7 an

Model	c_{12}	c_{13}	c_{14}	$c_{23\|1}$	$c_{24\|1}$	$c_{34\|12}$
1	N	C	G	F	J	t_3
2	t_3	G	J	G	N	J
3	C	t_5	F	G	N	J
4	G	t_5	J	C	t_{10}	F
5	N	C	BB1	BB7	t_{10}	F
6	BB1	t_5	J	BB7	t_{10}	F
7	BB7	t_5	F	N	t_{10}	BB1
8	BB1	N	BB7	BB1	t_5	BB7

Table 4: Mixed copula models investigated

additional constraint is needed to fix the two parameter values. Here we used the following constraints: $\lambda_L = \lambda_U$ (symmetric) or $2\lambda_L = \lambda_U$ (asymmetric).

As performance measure we considered average and 5% trimmed average of the estimated parameter values or corresponding Kendall's τ value for each pair-copula. Further, bias, variance and mean squared error of these quantities are estimated. All performance measures are based on 100 data sets simulated from each Kendall's τ scenario and C-vine copula model. The detailed results for $n = 500$ are contained in Schepsmeier (2010) in Appendix C.

We now summarize the results of Schepsmeier (2010) for C-vine MLE's:
strength of dependence:
MLE's of Kendall's τ are about equally well estimated for high and low dependence. Parameter values are better estimated for smaller values of Kendall's τ. For the t-copula the degree of freedom parameter ν is slightly overestimated. Overestimation increases as ν increases as to be expected since t_ν is close to the normal distribution for values $\nu \geq 25$. For the BB1/BB7 the asymmetric case performs worse for high dependencies.
mixed versus non-mixed:
There no significant difference in the performance.
conditional versus non-conditional:
The performance slightly decreases as the number of conditioning variables increases. Therefore additionally five dimensional mixed C-vine copula models are simulated and analyzed.
sequential versus MLE estimates:
Except for the degrees of freedom parameter ν in the t-copula, which are estimated by a bivariate maximum likelihood, the sequential estimates $\hat{\boldsymbol{\theta}}^S$ are close to the d-variate optimized maximum likelihood estimates $\hat{\boldsymbol{\theta}}^{MLE}$.

Similar to the simulation tests described above and run by Schepsmeier (2010) we run four new simulation tests to investigate the performance of the implemented MLE algorithm with respect to the **number of observations**. For more stable results we used 500 simulated data sets instead of 100, for a non-mixed four dimensional C-vine with BB1 pair-copulas and a mixed C-vine with copulas as in Model 7 in Table 4. Table 5 summarizes the performance by reporting the average estimated relative MSE of the parameters $\widehat{MSE}_{rel}(\hat{\boldsymbol{\alpha}})$ and the corresponding relative MSE of the Kendall's τs $\widehat{MSE}_{rel}(\hat{\tau})$. In the left panel simulated data sets of $n = 500$ are used while the right panels are for $n = 2000$. From this we see that the relative MSE is generally small and decreases as n increases.

In summary the simulation shows that satisfactory performance of ML estimation procedure for mixed/non-mixed C-vine copula models with a large catalogue of pair-copula families under small/high pairwise (conditional) dependence is possible for moderate sample sizes.

6 Small sample performance of the C-vine selection method

In the previous section we fixed the C-vine structure as well as the pair-copula families. In a second simulation study, performed in R, we investigated the C-vine and copula selection methods described in Section 4 with respect to their fitting accuracy using the absolute difference of pairwise Kendall's τ-matrices as an evaluation criterion. If one does not fix the C-vine structure but selects it by our method a different C-vine structure to the true one can emerge. Thus different pair-copulas occur which have to be selected too. Again we restrict to a four dimensional C-vine copula set. As already mentioned in Section 5 six pair-copula families have to be chosen in this case, which is done sequentially by the AIC selection method of Section 4.
As evaluation criterion for our model selection method we utilized absolute differences of pairwise Kendall's τs between the estimated Kendall's τ-matrix of simulated data $\boldsymbol{U}_0 = (\boldsymbol{u}_{0,1}, \dots, \boldsymbol{u}_{0,d}), \boldsymbol{u}_{0,j} = (u_{0,1,j}, \dots, u_{0,n_0,j})^T$ for $j = 1, \dots, d$, of a true model, denoted by $\hat{\tau}_0$, and the estimated Kendall's τ-matrix of simulated data $\boldsymbol{U}_1 = (\boldsymbol{u}_{1,1}, \dots, \boldsymbol{u}_{1,d}), \boldsymbol{u}_{1,j} = (u_{1,1,j}, \dots, u_{1,n_1,j})^T$ for $j = 1, \dots, d$, of the selected model, denoted by $\hat{\tau}_1$. The procedure is as follows:

1. Select the structure, pair-copula families and pair-copula parameters

non-mixed C-vine (BB1)					
True values		n = 500		n = 2000	
$\boldsymbol{\alpha} = (\theta, \delta)$	$\tau(\boldsymbol{\alpha})$	$\widehat{MSE}_{rel}(\hat{\boldsymbol{\alpha}})$	$\widehat{MSE}_{rel}(\hat{\tau})$	$\widehat{MSE}_{rel}(\hat{\boldsymbol{\alpha}})$	$\widehat{MSE}_{rel}(\hat{\tau})$
$\theta_{12} = 0.80$	$\tau_{12} = 0.80$	3.36%	0.01%	0.97%	$< 0.00\%$
$\delta_{12} = 3.56$		0.44%		0.11%	
$\theta_{13} = 0.80$	$\tau_{13} = 0.80$	5.67%	0.01%	1.51%	$< 0.00\%$
$\delta_{13} = 3.56$		0.60%		0.14%	
$\theta_{14} = 0.80$	$\tau_{14} = 0.80$	7.67%	0.01%	2.46%	$< 0.00\%$
$\delta_{14} = 3.56$		0.80%		0.18%	
$\theta_{23\|1} = 0.29$	$\tau_{23\|1} = 0.20$	9.85%	2.12%	2.38%	0.46%
$\delta_{23\|1} = 1.09$		0.18%		0.04%	
$\theta_{24\|1} = 0.29$	$\tau_{24\|1} = 0.20$	9.78%	2.17%	2.52%	0.50%
$\delta_{24\|1} = 1.09$		0.18%		0.05%	
$\theta_{34\|12} = 0.29$	$\tau_{34\|12} = 0.20$	10.83%	1.89%	2.68%	0.56%
$\delta_{34\|12} = 1.09$		0.18%		0.05%	
mixed C-vine (BB7,t5,F,N,t10,BB1)					
True values		n = 500		n = 2000	
$\boldsymbol{\alpha}$	τ	$\widehat{MSE}_{rel}(\hat{\boldsymbol{\alpha}})$	$\widehat{MSE}_{rel}(\hat{\tau})$	$\widehat{MSE}_{rel}(\hat{\boldsymbol{\alpha}})$	$\widehat{MSE}_{rel}(\hat{\tau})$
$\theta_{12} = 1.18$	$\tau_{12} = 0.80$	0.32%	$< 0.00\%$	0.05%	$< 0.00\%$
$\delta_{12} = 8.40$		0.52%		0.11%	
$\rho_{13} = 0.95$	$\tau_{13} = 0.80$	$< 0.00\%$	$< 0.00\%$	$< 0.00\%$	$< 0.00\%$
$\nu_{13} = 5.00$		0.68%		0.14%	
$\theta_{14} = 18.10$	$\tau_{14} = 0.80$	0.31%	$< 0.00\%$	0.08%	$< 0.00\%$
$\rho_{23\|1} = 0.95$	$\tau_{23\|1} = 0.80$	0.01%	0.03%	$< 0.00\%$	0.01%
$\rho_{24\|1} = 0.95$	$\tau_{24\|1} = 0.80$	0.01%	0.03%	$< 0.00\%$	0.01%
$\nu_{24\|1} = 10.00$		21.84%		4.24%	
$\theta_{34\|12} = 6.48$	$\tau_{34\|12} = 0.80$	11.83%	0.02%	3.14%	0.01%
$\delta_{34\|12} = 1.17$		1.23%		0.29%	

Table 5: Top: Average estimated relative MSE of parameters and Kendall's τ for non-mixed C-vine with BB1 pair-copulas, assuming $\lambda^L = \lambda^U$ based on 500 simulated data sets, (n=500 (left) and n=2000 (right)). Bottom: Average estimated relative MSE of parameters and Kendall's τ for mixed C-vine (BB7,t5,F,N,t10,BB1), assuming $\lambda^L = 0.2$ for the bivariate copula families based on 500 simulated data sets, (n=500 (left) and n=2000 (right)).

of the true C-vine model.

2. Simulate data $\boldsymbol{U}_0$ of the true model with length n_0.

3. Perform the model selection algorithm to choose the order of roots in a C-vine, the pair-copula families and estimate the copula parameters.

4. Simulate data $\boldsymbol{U}_1$ of the selected model with length n_1.

5. Sum up the absolute differences of the estimated Kendall's τ-matrices $\hat{\tau}_0$ and $\hat{\tau}_1$.

6. Repeat step 1. - 4. R times.

Note that n_2 should be much larger that n_1 ($n_2 \gg n_1$) to ensure that the influence of the simulation error is negligible. Next we calculated the replication average of the mean absolute pairwise Kendall's τ differences by

$$\bar{\tau} = \frac{1}{R} \sum_{r=1}^{R} \sum_{\{s,t\} \in \binom{\{1,\dots,d\}}{2}} |\hat{\tau}(\boldsymbol{u}_{0,s}^r, \boldsymbol{u}_{0,t}^r) - \hat{\tau}(\boldsymbol{u}_{1,s}^r, \boldsymbol{u}_{1,t}^r)|,$$

where $\binom{\{1,\dots,d\}}{2} := \{\{s,t\} : s \neq t, s,t \in \{1,\dots,d\}\}$ is the set of unordered pairs in $\{1,\dots,d\}$.
Our simulation study deals with the eight models 1-8 of Table 4 with four scenarios each described in Table 3 and Section 5, all in all 32 different models. The described procedure was run for $R = 100, n_1 = 1,000$ and $n_2 = 10,000$. In Table 6 we give the calculated values of $\bar{\tau}$ based on sequential pair-copula estimates of the selected model (left panel) and joint ML estimates (right panel), respectively.

	sequential estimation				ML estimation			
Model	1	2	3	4	1	2	3	4
1	0.04	0.07	0.03	0.10	0.05	0.07	0.03	0.09
2	0.25	0.06	0.03	0.05	0.28	0.08	0.03	0.08
3	0.02	0.06	0.03	0.05	0.03	0.06	0.03	0.05
4	0.11	0.06	0.03	0.06	0.11	0.06	0.03	0.07
5	0.17	0.06	0.03	0.05	0.18	0.07	0.04	0.05
6	0.05	0.06	0.03	0.06	0.05	0.07	0.04	0.07
7	0.02	0.06	0.03	0.03	0.04	0.06	0.03	0.06
8	0.09	0.05	0.04	0.06	0.09	0.06	0.04	0.06

Table 6: Estimated mean of absolute pairwise Kendall's τ differences for Model 1-8 of Table 4 with 4 scenarios (3) each. Left panel: based on sequential estimates, right panel: based on joint MLE after selection step.

Beside two larger values in Scenario 1 of Model 2 and 5, all mean absolute pairwise Kendall's τ differences are quite small, i.e. our model selection

method performs quite well. The additional joint ML step for the copula parameters after fitting the C-vine model with sequential estimates does not improve the fitting of the Kendall's τ matrix. For the two outliers we have no obvious explanation. One reason may be the numerical restrictions in the implementation of the MLE and selection algorithm for high Kendall's τs.

7 Application: US-Exchange rates

We apply now our mixed C-vine model to 8 time series of US-exchange rates from different countries from July 22, 2005 until July 17, 2009. Therefore we have 1007 daily values available for each country considered. The US-exchange rates are quoted in the home currency, e.g. 1 US-Dollar = 0.8466 Euro. For simplification we use the following abbreviations: EUR (Euro-area), UK (United Kingdom), CAN (Canada), AUS (Australia), BRA (Brazil), JPN (Japan), SZ (Switzerland) and IN (India).
First of all for each marginal exchange rate series an appropriate ARMA(P,Q)-GARCH(p,q) model, developed by Bollerslev (1986), is determined. Ljung-Box tests (Ljung and Box (1978)) for serial independence applied to the estimated standardized residuals show that ARMA(1,1)-GARCH(1,1) models are sufficient to remove the time dependence in each of the individual US-exchange rates. The corresponding p-value of the Ljung-Box tests as well as additional QQ-Plots can be found in Chapter 5 of Schepsmeier (2010). The QQ-Plots provide evidence that the residuals are fat tailed and t- or skewed t-innovations in the GARCH(1,1) part is needed. The resulting standardized residuals of these models are transformed using the empirical probability integral transformation and a scaling factor $\frac{n}{n-1}$ to copula data on $(0,1)^8$. Figure 2 shows scatter plots and the estimated Kendall's τ for the copula data. We can detect some stronger dependencies between the standardized residual US-exchange rates, especially between the EURO and UK, EURO and AUS, EURO and SZ. Further, we can see that some standardized residual US-exchange rates are almost independent as for example AUS and JPN and IN and SZ. A next interesting fact is, that almost all dependencies are positive beside the pair BRA - JPN and CAN - JPN.

We apply now the sequential procedure to select an appropriate C-vine copula for the US-exchange rate copula data. Table 7 gives the empirical Kendall's τ matrix and the sum of their absolute values, denoted by S (compare to (4.1)). From this we see that EUR is the first root variable. Given this first root variable and the sequential C-vine identification procedure from previous chapter the next root variable AUS followed by SZ, BRA, CAN and

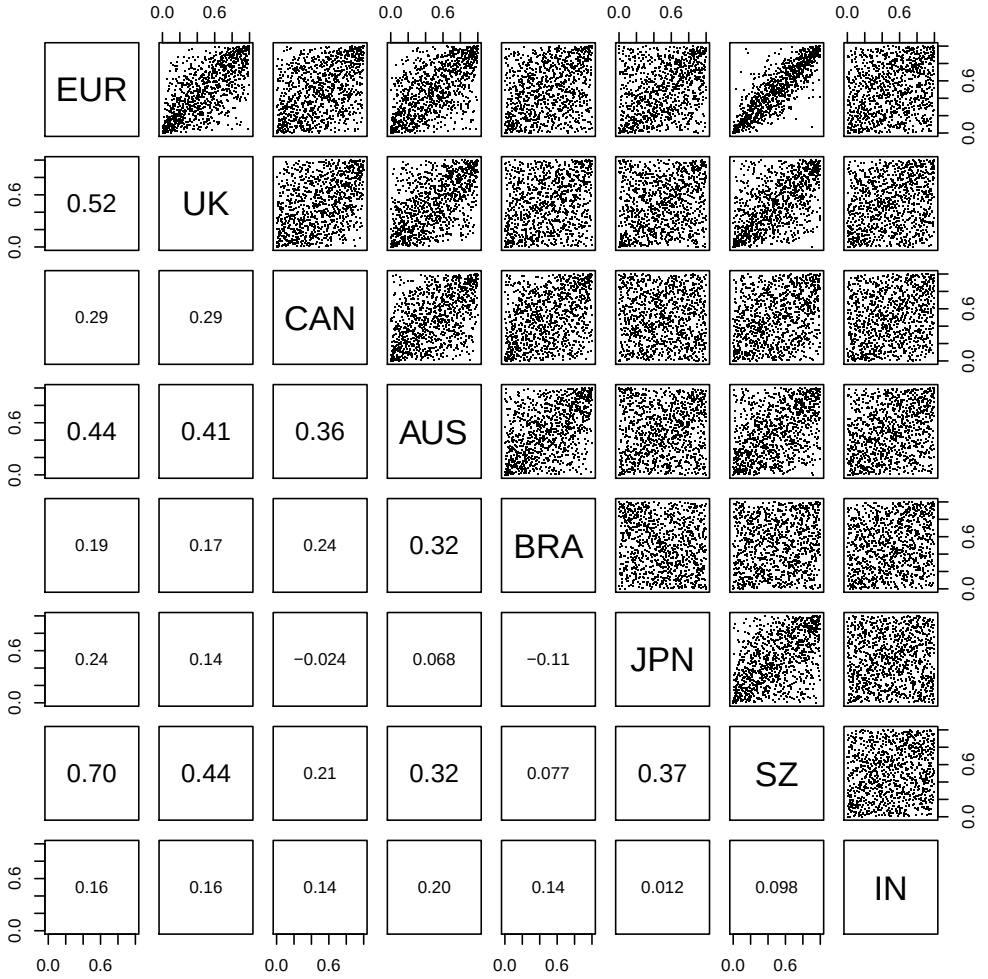


Figure 2: Pairsplot of the copula data (top, right) formed from transformed standardized residual US-exchange rates and the corresponding estimated Kendall's τ (bottom, left)

finally IN can be identified. Table 8 displays now the empirical Kendall's τ matrix of the variables given EUR as first root and copula families and parameters as in Table 9 (i=0), and sum over the absolute entries of each row. As possible copula families we allowed the eight copula families utilized in the simulation study. Since the sequential selection procedure identifies the pair-copula types and provides sequential estimates $\hat{\theta}^S$, we use those as starting values to determine the corresponding MLE $\hat{\theta}^{MLE}$. The resulting mixed

		EUR	UK	CAN	AUS	BRA	JPN	SZ	IN	S
1	EUR	1.00	0.51	0.29	0.44	0.19	0.24	0.69	0.16	**3.54**
2	UK	0.51	1.00	0.28	0.41	0.17	0.13	0.43	0.15	3.12
3	CAN	0.29	0.28	1.00	0.35	0.24	-0.02	0.20	0.14	2.56
4	AUS	0.44	0.41	0.35	1.00	0.31	0.06	0.32	0.19	3.12
5	BRA	0.19	0.17	0.24	0.31	1.00	-0.11	0.07	0.14	2.27
6	JPN	0.24	0.13	-0.02	0.06	-0.11	1.00	0.37	0.01	1.97
7	SZ	0.69	0.43	0.20	0.32	0.07	0.37	1.00	0.09	3.21
8	IN	0.16	0.15	0.14	0.19	0.14	0.01	0.09	1.00	1.92

Table 7: Empirical Kendall's τ matrix and the sum over the absolute entries of each row for the exchange rate data set

		EUR,AUS	EUR,SZ	EUR,BRA	EUR,CAN	EUR,IN	EUR,JPN	EUR,UK	S
1	EUR,AUS	1.00	-0.14	0.25	0.23	0.13	-0.11	0.17	**2.03**
2	EUR,SZ	-0.14	1.00	-0.21	-0.12	-0.09	0.34	0.00	1.90
3	EUR,BRA	0.25	-0.21	1.00	0.18	0.11	-0.20	0.04	1.99
4	EUR,CAN	0.23	-0.12	0.18	1.00	0.07	-0.15	0.10	1.85
5	EUR,IN	0.13	-0.09	0.11	0.07	1.00	-0.04	0.60	1.50
6	EUR,JPN	-0.11	0.34	-0.20	-0.15	-0.04	1.00	-0.03	1.87
7	EUR,UK	0.17	0.00	0.04	0.10	0.60	-0.03	1.00	1.40

Table 8: Empirical Kendall's τ matrix and the sum over the absolute entries of each row for the exchange rate data set given EUR as first root

C-vine copula model we denote by M1. The sequential and ML estimates for M1 are provided in Table 9. Note that the variable i indicates the number of variables in the conditioning set and the pair-copula family type chosen. Here we use N(Gaussian), t(t-), C(Clayton), G(Gumbel) and F(Frank) as abbreviations. Surprisingly, the flexibility of two parameter bivariate copula families such as the BB1 and BB7 was not required for this data set.

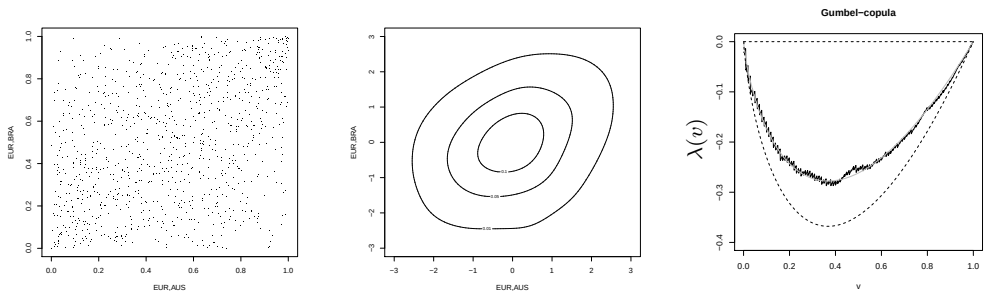
In particular only a few strong non-symmetric dependencies among the standardized residuals are detected, i.e. BRA and AUS given EU and CAN and BRA given EU, AUS and SZ. For the first conditional dependency we provide exemplarily the exploratory copula selection analysis in Table 10. This includes also the goodness-of-fit statistic and p-value of the Cramer-von Mises test. Similar results are given for all pair-copula terms up to conditioning on 3 variables can be found in Schepsmeier (2010).

From Table 9 we see that often very low dependence is estimated. Therefore

		M1			M4	
i	**Parameter**	**Copula**	$\hat{\boldsymbol{\theta}}^{S}$	$\hat{\boldsymbol{\theta}}^{MLE}$	$\hat{\boldsymbol{\theta}}^{S}$	$\hat{\boldsymbol{\theta}}^{MLE}$
0	$\rho_{EUR,AUS}$	t	0.64	0.63	0.64	0.64
	$\nu_{EUR,AUS}$		10.16	13.29		
	$\rho_{EUR,SZ}$	t	0.89	0.89	0.89	0.87
	$\nu_{EUR,SZ}$		4.00	4.68		
	$\rho_{EUR,BRA}$	t	0.30	0.31	0.30	0.31
	$\nu_{EUR,BRA}$		6.39	7.42		
	$\rho_{EUR,CAN}$	t	0.46	0.46	0.45	0.46
	$\nu_{EUR,CAN}$		8.44	8.75		
	$\rho_{EUR,IN}$	N	0.25	0.26	0.25	0.26
	$\rho_{EUR,JPN}$	t	0.36	0.37	0.37	0.34
	$\nu_{EUR,JPN}$		4.38	5.88		
	$\rho_{EUR,UK}$	t	0.72	0.72	0.73	0.72
	$\nu_{EUR,UK}$		8.49	9.21		
1	$\rho_{SZ,AUS\|EUR}$	t	-0.24	-0.23	-0.22	-0.25
	$\nu_{SZ,AUS\|EUR}$		11.36	16.49		
	$\theta_{BRA,AUS\|EUR}$	G	1.33	1.30	0.37	0.39
	$\rho_{CAN,AUS\|EUR}$	N	0.35	0.35	0.35	0.36
	$\theta_{IN,AUS\|EUR}$	F	1.19	1.19	0.20	0.19
	$\rho_{JPN,AUS\|EUR}$	t	-0.18	-0.18	-0.17	-0.22
	$\nu_{JPN,AUS\|EUR}$		6.86	7.89		
	$\rho_{UK,AUS\|EUR}$	t	0.26	0.26	0.27	0.27
	$\nu_{UK,AUS\|EUR}$		11.77	12.29		
2	$\theta_{BRA,SZ\|EUR,AUS}$	F	-1.48	-1.52	-0.27	-0.24
	$\theta_{CAN,SZ\|EUR,AUS}$	F	-0.75	-0.79	-0.13	-0.11
	$\rho_{IN,SZ\|EUR,AUS}$	t	-0.11	-0.11	-0.11	-0.11
	$\nu_{IN,SZ\|EUR,AUS}$		50.00	135.05		
	$\rho_{JPN,SZ\|EUR,AUS}$	t	0.48	0.49	0.51	0.50
	$\nu_{JPN,SZ\|EUR,AUS}$		11.91	12.28		
	$\theta_{UK,SZ\|EUR,AUS}$	F*	0.26	0.26	0.04	-0.02
3	$\theta_{CAN,BRA\|EUR,AUS,SZ}$	G	1.10	1.10	0.16	0.15
	$\rho_{IN,BRA\|EUR,AUS,SZ}$	N	0.08	0.09	0.08	0.09
	$\theta_{JPN,BRA\|EUR,AUS,SZ}$	F	-0.87	-0.88	-0.17	-0.15
	$\rho_{UK,BRA\|EUR,AUS,SZ}$	N*	-0.05	-0.05	-0.04	-0.04
4	$\rho_{IN,CAN\|EUR,AUS,BRA,SZ}$	N*	0.02	0.05	0.02	0.05
	$\rho_{JPN,CAN\|EUR,AUS,BRA,SZ}$	N	-0.13	-0.12	-0.13	-0.13
	$\rho_{UK,CAN\|EUR,AUS,BRA,SZ}$	N	0.08	0.09	0.08	0.10
5	$\rho_{JPN,IN\|EUR,CAN,AUS,BRA,SZ}$	N*	0.02	0.02	0.03	0.02
	$\rho_{UK,IN\|EUR,CAN,AUS,BRA,SZ}$	N*	0.05	0.05	0.05	0.05
6	$\rho_{UK,JPN\|EUR,CAN,AUS,BRA,SZ,IN}$	N*	-0.03	-0.02	-0.03	-0.03

Table 9: Sequential and ML estimates for C-vine copula models M1 and M4 (* = independence copula in M2).

Copula $c_{BRA,AUS|EUR}$ ($\hat{\tau} = 0.25$)



	Gaussian	t	Clayton	Gumbel	Frank
Cramer-von Mises stat.	0.033	0.035	0.262	0.019	0.047
p-value	0.049	0.036	0.000	**0.448**	0.003

	Gaussian	t	Clayton	Gumbel	Frank	BB1	BB7
Vuong	1	1	-6	1	1	**2**	0
Clarke	-6	**6**	-3	4	-1	1	-1

Possible copula(s):	Selected copula:
G, BB1, BB7	G

Table 10: Pairs-plot, normalized contour plot, λ-function and goodness-of-fit test scores for the conditional copula $c_{BRA,AUS|EUR}$ of the US-exchange rate data.

for each pair-copula a test for independence is performed to decide whether the corresponding pair-copula has to be replaced with the independence copula or not. As independence test we used a test based on Kendall's τ discussed in Genest and Favre (2007) utilizing the asymptotic distribution of empirical Kendall's τ under independence (Kendall (1938)). If the test fails to reject the null hypothesis of independence an independence copula ($c(u_1, u_2) \equiv 1$) is chosen. Otherwise the sequential selection procedure is left unchanged. We call the resulting model M2. Comparing M1 to M2 we see that copula indices are the same up to a conditioning set of size 5. The remaining root variable for M2 is then JPN, while IN was chosen for M1. In Table 9 those pairs chosen to be the independence copula in M2 are indicated by an asterisk *. For brevity the corresponding parameter estimates for M2 are not shown. It should be noted that M2 has only 33 parameters compared to 39 for M1. To investigate if a mixed C-vine copula is necessary for this data set we also fit a C-vine copula with the same structure as M1 but only with t pair-copulas. This model is referred as M3. Finally we also want to investigate if a C-vine specification is needed at all. Therefore we set all pair-copulas in M1 to a Gaussian copula and denote this model

Model	**Model type**	**Model selection**
M1	mixed C-vine	sequential selection without independence test
M2	mixed C-vine	sequential selection with independence test
M3	C-vine t-copula	same as M1 but all pair-copulas are t-copulas
M4	Gauss copula	same as M1 but all pair-copulas are Gaussian copulas

Table 11: Summary of models investigated

by M4. Table 11 summarizes the specification of the four models investigated.

Model comparison is performed based on the AIC (Akaike (1974)) and BIC (Schwarz (1978)) criteria. The AIC and BIC values has been computed also using the sequential estimates $\hat{\boldsymbol{\theta}}^S$ as a proxy for $\hat{\boldsymbol{\theta}}^{MLE}$, to see how close the sequential estimates are to the MLE's with respect to these criteria. Table 12 gives the results. Thus the model M2 is the clearly preferred over all remaining models showing that a mixed C-vine copula is fitting best. The same conclusions can be drawn if one uses $\hat{\boldsymbol{\theta}}^S$ as proxy for $\hat{\boldsymbol{\theta}}^{MLE}$, which is not surprising since $\hat{\boldsymbol{\theta}}^S$ is consistent for $\boldsymbol{\theta}$.

	M1	M2	M3	M4
Log-likelihood seq.	2203.4	2198.5	2212.1	2077.0
Log-likelihood MLE	2207.5	2202.4	2217.5	2088.0
# of parameters	39	33	56	28
AIC seq.	-4328.8	-4331.0	-4312.2	-4098.0
AIC MLE	-4337.0	-4338.8	-4323.0	-4120.0
BIC seq.	-4137.1	-4168.8	-4037.2	-3960.4
BIC MLE	-4145.3	-4176.6	-4047.8	-3982.4

Table 12: Log-likelihood, number of parameters, AIC and BIC for models M1-M4 using ML or sequential estimates.

In Table 12 we ignore the fact that only M3 and M4 are nested models, while the other pairs are not. Therefore we conducted appropriate Vuong and Clarke tests with Schwarz correction and the corresponding results are given in Table 13. These also support decisively the conclusion that M2 is the preferred model. In summary a mixed C-vine copula is needed to model the dependencies among the standardized US-exchange rate residuals.

Null Hypothesis	Method		Alternative M2	M3	M4
M1	Vuong	statistics	-4.7	4.94	3.16
		p-value	< 0.000	< 0.000	0.002
		decision	$M2 > M1$	$M1 > M3$	$M1 > M4$
	Clarke	statistics	389	616	556
		p-value	< 0.000	< 0.000	0.001
		decision	$M2 > M1$	$M1 > M3$	$M1 > M4$
M2	Vuong	statistics		6.15	3.76
		p-value		< 0.000	< 0.000
		decision		$M2 > M3$	$M2 > M4$
	Clarke	statistics		629	578
		p-value		< 0.000	< 0.000
		decision		$M2 > M3$	$M2 > M4$
M3	Vuong	statistics			1.28
		p-value			0.202
		decision			$M3 = M4$
	Clarke	statistics			473
		p-value			0.029
		decision			$M4 > M3$

Table 13: Pairwise non-nested model comparison using Vuong and Clarke tests with Schwarz correction.

8 Summary and Outlook

In this paper we introduced the class of mixed C-vine copulas and provided sequential and ML estimation procedures for the unknown parameters. Two extensive simulation studies showed very satisfactory behavior of the ML estimation for many different mixed and non-mixed C-vine copulas and our model selection method, respectively.
In contrast to earlier papers on vines we considered the problem of jointly estimating the C-vine structure together with the choice of bivariate copula families. We developed a data driven sequential approach, which follows the heuristic of wanting to achieve parsimonious and simple model specifications. Following these considerations C-vine structures are selected so that most of the dependence as measured by the sum absolute empirical pairwise Kendall's τ values occurs early in the conditioning procedure. The computational complexity of the log-likelihood is increased as the number of conditioning variables is increased, since the required conditional cdf's require higher recursions. For the selection of the appropriate pair-copula families we followed standard test approaches involving goodness-of-fit tests

for bivariate copulas, Vuong and Clarke tests suitable for non-nested models and finally explorative tools based on scatter and contour plots as well as the empirical λ-function of Genest and Rivest (1993).
Finally we considered an application involving US-exchange rates. Here mixed C-vines are used to model the dependencies among standardized residuals, which are formed from univariate time series models. This allows to have pseudo copula data available which is approximately i.i.d. Therefore a standard two-step estimation procedure is followed, which however allows via simulation to consider value at risk of arbitrary portfolios. The advantage of this two part joint model (marginal + copula model) is that simulation has to be conducted only once for all portfolio's considered, while allowing for very flexible non Gaussian dependencies in contrast for example to a CCC model of Bollerslev (1990).
We like to note that our experience with joint estimation methods of models with regression marginals (Lanzendörfer (2009)), with AR(1) marginals (Czado et al. (2010)) and GARCH(1,1) (Hofmann and Czado (2010)) coupled with D-vine copula models have shown that the loss in efficiency when two step estimation is done is small. In these papers a Bayesian approach was followed, which allows for credible intervals for parameters and quantities of interest depending on parameters. These interval estimates are difficult to obtain in a ML setup. The mentioned papers in this paragraph are extensions of the Bayesian approach followed by Min and Czado (2010a),
In the future we like to investigate the severeness of the restrictions imposed by the heuristic C-vine structure search. For this a discrepancy measure between the fitted C-vine structure and the true model has to be considered. A general statistical discrepancy measure would be the Kulback-Leibler distance, which however is difficult to obtain in these complex models. Therefore simpler discrepancy measures have to be developed. Another future area of research is the development of search algorithms for D-vines and more generally for regular vines. Using the simplified pair-copula construction we loose the direct influence of the conditional variables on the copula density and thus on the copula parameter estimation. One way would be to allow the parameter of the pair-copula to depend on the value of conditioning variables in a regression setup, i.e. the use of these values as covariates. This will be the topic of future research.

9 Acknowledgment

The numerical computations were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG. We like to thank the Editor and two

anonymous referees for their useful comments which helped to improve the manuscript.

References

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

Akaike, H. (1974). A new look at statistical model identification. *IEEE Transaction on Automatic Control 19*(6), 716–723.

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

Bedford, T. and R. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance 15*, 1466–4364.

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

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economic and Statistics 72*, 498–505.

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

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

Clarke, K. (2007). A Simple Distribution-Free Test for Nonnested Model Selection. *Political Analysis 15*, 347–363.

Czado, C. (2010). Pair-Copula Constructions of Multivariate Copulas. In P. e. a. Jaworski (Ed.), *Copula Theory and Its Applications, Lecture Notes in Statistics*, Volume 198, Berlin Heidelberg, pp. 93–109. Springer-Verlag.

Czado, C., F. Gärtner, and A. Min (2010). Joint Bayesian inference of D-vines with AR(1) margins. In H. Kurowicka, D. Joe (Ed.), *Dependence*

Modeling-Handbook on Vine Copulas, Singapore, pp. 330–359. World Scientific Publishing.

Embrechts, P., F. Lindskog, and A. McNail (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of heavy tailed distributions in finance*, pp. 329–384. Amsterdam: Elsevier/North-Holland.

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

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

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

Genest, C., J.-F. Quessy, and B. Rémillard (2006). Goodness-of-fit Procedures for Copula Model Based on the Probability Integral Transformation. *Scandinavian Journal of Statistics 33*, 337–366.

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

Genest, C., B. Rémillard, and D. Beaudoin (2009). Goodness-of-fit tests for copulas: a review and power study. *Insur. Math. Econ. 44*, 199–213.

Haff, I. (2010). Parameter estimation for pair-copula constructions. *preprint*.

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

Hofmann, M. and C. Czado (2010). Assessing the VaR of a portfolio using D-vine copula based multivariate GARCH models. submitted preprint.

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 (Ed.), *Distributions with Fixed Marginals and Related Topics*, Volume 28, Hayward, CA, pp. 120–141. Inst. Math. Statist.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman und Hall, London.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis 94*(2), 401–419.

Kendall, M. (1938). A new measure of rank correlation. *Biometrika 30*, 81–93.

Kruskal, W. (1958). Ordinal Measures of Association. *Journal of the American Statistical Association 53*(284), 814–861.

Kurowicka, D. and R. Cooke (2001). Conditional, partial and rank correlation for elliptical copula; dependence modeling in uncertainty analysis. *Proceedings of ESREL 2001, Turyn, Italy*.

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

Lanzendörfer, J. (2009). Joint estimation of parameters in multivariate normal regression with correlated errors using pair-copula constructions and an application to finance. Diplomarbeit, Center of Mathematical Sciences, Munich University of Technology, Garching bei München. Available under http://www-m4.ma.tum.de/Diplarb/da_txt.html.

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

Min, A. and C. Czado (2010a). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics 8*(4), 511–546.

Min, A. and C. Czado (2010b). Bayesian model selection for D-vine pair-copula constructions. *forthcoming in Canadian Journal of Statistics*.

Min, A. and C. Czado (2010c). SCOMDY models based on pair-copula constructions with application to exchange rates. *preprint*.

Nelsen, R. (2006). *An Introduction to Copulas*. Springer, New York.

Nikoloulopoulos, A., H. Joe, and H. Li (2011). Vine copulas with asymmetric tail dependence and applications to financial return data. *in press by Computational Statistics and Data Analysis*.

Schepsmeier, U. (2010). Maximum likelihood estimation of C-vine pair-copula constructions on bivariate copulas from different families. Diplomarbeit, Center of Mathematical Sciences, Munich University of Technology, Garching bei München. Available under http://www-m4.ma.tum.de/Diplarb/da_txt.html.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*(2), 461–464.

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

Smith, M., A. Min, C. Almeida, and C. Czado (2010). Modelling longitudinal data using a pair-copula decomposition of serial dependence. Working Paper, Zentrum Mathematik, Technische Universtät München, forthcoming in *The Journal of the American Statistical Association*.

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