Norsk Regnesentral

NORWEGIAN COMPUTING CENTER

Note

Comparing estimators for pair-copula constructions

Note no SAMBA/22/11

Author Ingrid Hobæk Haff

Date 10th August 2011

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modelling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Comparing estimators for pair-copula constructions**
Author	**Ingrid Hobæk Haff**
Date	10th August 2011
Publication number	SAMBA/22/11

Abstract

We compare two of the most used estimators for the parameters of a pair-copula construction (PCC), namely the semiparametric (SP) and the stepwise semiparametric (SSP) estimators. Whereas the former estimates all parameters simultaneously, the latter does it in several steps. Consequently, the speed of the SSP estimator is considerably higher, at the expense of its asymptotic efficiency. Based on an extensive simulation study, we find that the performance of the SSP estimator overall is rather good compared to its contender, at least as long as the dependence is not too strong. It loses some efficiency with respect to SP with increasing dependence, especially in the top levels of the structure. On the other hand, the SSP estimator seems to suffer less under reduced sample sizes and partial misspecification of the model. Finally, it is the only real alternative for large dimensional problems. Though it struggles with the top level parameters, the lower order dependencies of the resulting estimated PCC mimic the true distribution well.

Keywords	copulae; empirical distribution functions; hierarchical construction; stepwise estimation; vines; computational speed
Target group	
Availability	
Project	
Project number	
Research field	Finance, Insurance and Commodity Markets
Number of pages	33

1 Introduction

Multivariate modelling is a rapidly growing field in statistics. In particular, the interest for and use of copulae has positively boomed since the late 1990's. That has entailed the development of a number of hierarchical, copula-based structures, among those the pair-copula construction (PCC) proposed by Joe (1996). This structure has later been explored and considered by Bedford and Cooke (2001, 2002), Kurowicka and Cooke (2006) and Aas et al. (2009).

PCCs are treelike constructions, having pair-copulae as building-blocks. Delightfully simple as these structures are, they no doubt owe their popularity to high flexibility and the ability to model a wide scope of dependencies (Hobæk Haff et al., 2010; Joe et al., 2010). Accordingly, several estimators for PCC parameters have been proposed, e.g. Aas et al. (2009); Czado and Min (2010); Joe and Xu (1996); Kolbjørnsen and Stien (2008) and more recently, Hobæk Haff (2010). The aim of this work is to compare alternative estimation procedures. A PCC is in fact a multivariate copula. All standard copula parameter estimators are therefore applicable also to PCCs. We are mainly interested in estimation procedures that are less model specific than method of moments type estimators, typically inversion of Kendall's τ coefficients (Clayton, 1978; Genest, 1987; Genest and Rivest, 1993; Oakes, 1982).

Among the more general approaches, the most relevant are semiparametric (SP) estimation, inference function for margins estimation and maximum likelihood, of which the latter two depend on the specified margins. Our focus is on the dependence parameters. Moreover, the effect of margins on the estimation of copula parameters has already been extensively studied (Joe, 2005; Kim et al., 2007). Hence, out of the above three, we will only consider the semiparametric estimator, that will serve as a benchmark. Along with the contending stepwise semiparametric (SSP) estimator, it is one of the most commonly used estimators for PCCs.

As one would expect, the malleability of pair-copula constructions does not come without a price. Even low dimensional structures are rather extensive in parameters, and the number grows quickly with the number of variables. SP estimation consists in estimating all parameters simultaneously, which generally requires numerical optimisation. Hence, it is very likely that it will be numerically challenging and time consuming with increasing dimension. The SSP estimator, which is designed for PCCs, performs the estimation in several steps. That increases the speed considerably, but reduces its asymptotic efficiency.

There exist expressions for the asymptotic covariance matrices of these estimators (see for instance Genest et al. (1995); Hobæk Haff (2010)). However, they involve multiple integrals, which in practice are incalculable. Therefore, we base the comparison on an extensive simulation study.

The paper is organised as follows. Section 2 presents the model, i.e. PCCs, whereas the two estimators are introduced in Section 3. The results of the simulation study are exhibited in Section 4. Finally, we summarise and discuss the results in Section 5.

2 Model

Consider a d-variate stochastic vector $\boldsymbol{X} = (X_1, \ldots, X_d)^T$ from an absolutely continuous distribution $F_{1\ldots d}$ with strictly increasing margins $F_1, \ldots, F_d$. Using Sklar's theorem (Sklar, 1959), as well as the chain rule, the probability density function (pdf) of $\boldsymbol{X}$ may be expressed as

$$f_{1\ldots d}(x_1, \ldots, x_d) = \prod_{l=1}^{d} f_l(x_l) \cdot c_{1\ldots d}(F_1(x_1), \ldots, F_d(x_d)), \tag{2.1}$$

where $f_l,\ l = 1, \ldots, d$, are the corresponding marginal pdfs and $c_{1\ldots d}$ the copula density. Likewise, it can be factorised as

$$f_{1\ldots d}(x_1, \ldots, x_d) = f_1(x_1) f_{2|1}(x_2|x_1) \ldots f_{d|1\ldots d-1}(x_d|x_1, \ldots, x_{d-1}). \tag{2.2}$$

The related pair-copula construction (PCC) results from expressing the factors on the right hand side of (2.2) in terms of pair-copula densities and marginal pdfs, through the repeated use of (2.1). Applying (2.1) in two dimensions, the second factor $f_{2|1}(x_2|x_1)$ is given by

$$f_{2|1}(x_2|x_1) = f_2(x_2) c_{12}(F_1(x_1), F_2(x_2)).$$

Correspondingly,

$$\begin{aligned} f_{3|1,2}(x_3|x_1, x_2) &= f_{3|2}(x_3|x_2) c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \\ &= f_3(x_3) c_{23}(F_2(x_2), F_3(x_3)) c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)), \end{aligned}$$

where $c_{13|2}$ is the copula density for the distribution of the pair (X_1, X_3) conditioning on X_2. As one continues with the remaining factors of (2.2), one finally obtains a product of marginal pdfs and pair-copula densities, i.e. a PCC.

All components of the structure, both marginal and pair-copula densities, can be chosen completely freely, i.e. from different families. The resulting distribution is guaranteed to be valid. Despite its simple building blocks, the pair-copula construction is therefore an exceptionally flexible model, able to portray a wide range of dependence structures (Joe et al., 2010).

There are numerous ways of factorising $f_{1\ldots d}$ and of substituting the factors with pair-copula densities, each resulting in a valid PCC. A large subset of these belongs to the family of regular vines, introduced by Bedford and Cooke (2001, 2002), which again comprises canonical (C) and drawable (D) vines. Five dimensional examples of the latter two are shown in Figure 2.1. For a more thorough introduction to vines and pair-copula constructions, see for instance Aas et al. (2009).

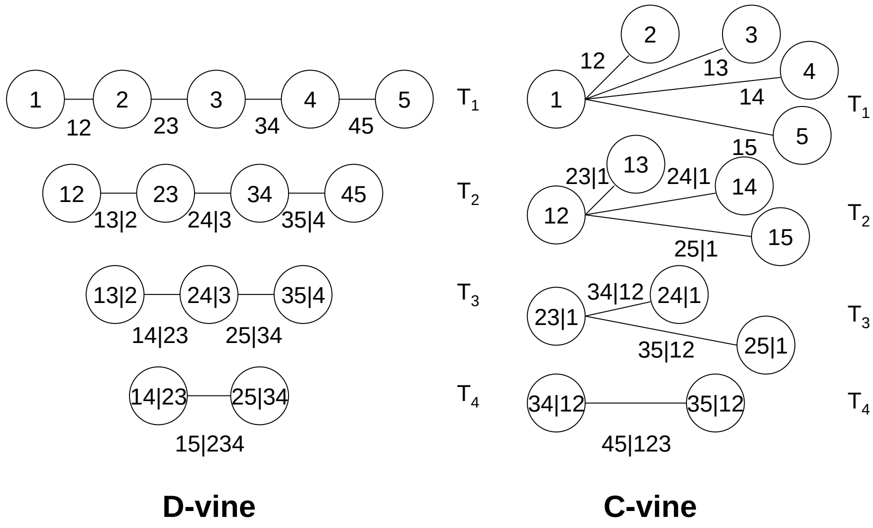


Figure 2.1. Five-dimensional D-vine (to the left) and C-vine (to the right).

For simplicity, we will hereafter restrict our attention to D-vines. The corresponding pdf is given by

$$f_{1..d}(x_1, \dots, x_d) = \prod_{l=1}^{d} f_l(x_l) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|v_{ij}} \left(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}) \right), \tag{2.3}$$

where v_{ij} denotes the index set $\{i+1, \dots, i+j-1\}$. In the double product over the pair-copula densities, j runs over the levels of the structure, for instance T_1 to T_5 in Figure 2.1, and i over the copulae at each level. Note that excepting the ground level, the arguments of the pair-copulae are conditional distributions, whose number of conditioning variables increases by one with each level. For inference to be possible in practice, one has to make the assumption that these so-called conditional pair-copulae depend on the conditioning variables only through their arguments, thus obtaining a *simplified* PCC. Take for example $c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))$, which is given by

$$c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) = \frac{f_{13|2}(x_1, x_3|x_2)}{f_{1|2}(x_1|x_2) f_{3|2}(x_3|x_2)}.$$

If $C_{13|2}$ is a Gaussian copula with parameter ρ, one must assume that ρ is constant over all values of x_1. This is generally not the case, but Hobæk Haff et al. (2010) showed that this is not a very restricting assumption, and that a general PCC may be very well approximated by a simplified one.

Assume now that we have n independent observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ from the model (2.3), and that we wish to estimate the corresponding parameters. The log-likelihood function is given by

$$\begin{aligned} l(\boldsymbol{\alpha}, \boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = & \sum_{k=1}^{n} \sum_{l=1}^{d} \log f_l(x_{lk}; \boldsymbol{\alpha}) + \\ & \sum_{k=1}^{n} \sum_{j=1}^{d-1} \sum_{i=1}^{d-j} \log c_{i,i+j|v_{ij}} \Big(F_{i|v_{ij}}(x_{ik}|\boldsymbol{x}_{v_{ij},k}; \boldsymbol{\alpha}, \boldsymbol{\theta}), \\ & \qquad F_{i+j|v_{ij}}(x_{i+j,k}|\boldsymbol{x}_{v_{ij},k}; \boldsymbol{\alpha}, \boldsymbol{\theta}); \boldsymbol{\theta} \Big), \end{aligned} \tag{2.4}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ are the marginal and dependence parameters, respectively. The computation of (2.4) requires calculation of the conditional distributions that are arguments of the pair-copulae. Using the simplifying assumption, one may express them as functions of two other conditionals with one conditioning variable less. As shown by Joe (1997)

$$F_{i|v\cup j}(x_i|\boldsymbol{x}_{v\cup j}) = \left. \frac{\partial C_{ij|v}(u_i, u_j)}{\partial u_j} \right|_{u_i = F_{i|v}(x_i|\boldsymbol{x}_v), u_j = F_{j|v}(x_j|\boldsymbol{x}_v)}, \tag{2.5}$$

where i, j are distinct indices, and v is a non-empty set of indices, that contains neither i nor j. Likewise, $F_{i|v}$ and $F_{j|v}$ can be expressed as bivariate functions of conditional distributions with a conditioning set reduced by one, and so on. Hence, all the necessary conditional distributions in (2.4) are nested functions of the margins $F_1, \ldots, F_d$. The log-likelihood function can therefore be written as

$$l(\boldsymbol{\alpha}, \boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = l_M(\boldsymbol{\alpha}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) + l_C(\boldsymbol{\theta}; \boldsymbol{u}_1(\boldsymbol{\alpha}), \ldots, \boldsymbol{u}_n(\boldsymbol{\alpha})), \tag{2.6}$$

where $\boldsymbol{u}_k(\boldsymbol{\alpha}) = (F_1(x_{1k}; \boldsymbol{\alpha}), \ldots, F_d(x_{dk}; \boldsymbol{\alpha}))$.

3 Estimators

The estimators we wish to compare are two of the most commonly used ones for pair-copula constructions, namely the semiparametric and the stepwise semiparametric estimators.

3.1 The semiparametric estimator

The semiparametric (SP) estimator was introduced by Genest et al. (1995), and for censored data by Shih and Louis (1995). Later, it has been generalised by Tsukahara (2005). Dependence on the chosen margins is removed by replacing the parametric cdfs $F_i(\cdot, \boldsymbol{\alpha})$ in (2.6) with non-parametric ones $F_{in}(\cdot)$. If one is interested in estimating measures, such as Kendall's τ, Spearman's ρ or tail dependence coefficients, that are functions only of the dependence parameters, the SP estimator is a very natural choice. Not only is it more robust to misspecified margins, it avoids specifying margins altogether.

Define the pseudo observations

$$u_{ikn} = F_{in}(x_{ik}) = \frac{1}{n+1} \sum_{j=1}^{n} I(x_{ik} \leq x_{ij}),\ i = 1, \ldots, d,\ k = 1, \ldots, n,$$

where $I(\cdot)$ is the indicator function, and the pseudo log-likelihood function as

$$l_{C,P}(\boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = l_C(\boldsymbol{\theta}; \boldsymbol{u}_{1n}, \ldots, \boldsymbol{u}_{nn}), \tag{3.1}$$

with l_C as defined in (2.6) and $\boldsymbol{u}_{kn} = (u_{1kn}, \ldots, u_{dkn})$. This is the sum over all log-copula densities, plugging in the pseudo observations. The SP estimator is now simply the maximiser of $l_{C,P}$ with respect to $\boldsymbol{\theta}$, i.e.

$$\hat{\boldsymbol{\theta}}^{SP} = \underset{\boldsymbol{\theta}}{argmax}\{l_{C,P}(\boldsymbol{\theta}; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\}.$$

This estimation procedure is thoroughly described for D-vines in Aas et al. (2009), whereas the large sample properties of the estimator for PCCs are described for instance in Hobæk Haff (2010).

Although the pseudo log-likelihood function can be expressed directly as a function of the empirical margins, it is in practice computed iteratively, level by level. At a given level, the necessary pair-copula arguments are computed by applying the functions (2.5) to the arguments from the preceding level. Thus, they are successive transformations of the pseudo observations. More specifically, the arguments of level l are the result of $l-1$ such transformations.

Previous studies have shown that the SP estimator performs well for copulae compared to alternative estimators, such as maximum likelihood, inference function for margins and the minimum distance estimator (Kim et al., 2007; Tsukahara, 2005). However, these studies have focussed on bivariate examples. As mentioned earlier, the number of parameters of a pair-copula construction grows quickly with the number of variables. For instance, a five dimensional D-vine consisting of t-copulae has 20 parameters (not counting the margins). With an additional dimension, the number of parameters increases to 30. Thus, for a medium to high number of variables, the optimisation of the pseudo log-likelihood function becomes numerically demanding and time consuming.

3.2 The stepwise semiparametric estimator

By performing the estimation in several steps, one can speed up the procedure considerably. This is the main idea of the stepwise semiparametric (SSP) estimator, which is designed for PCCs. It was suggested in Aas et al. (2009), and more formally introduced by Hobæk Haff (2010). The latter also explores its large sample characteristics.

The SSP estimator is very similar to the SP one. The difference is that the PCC parameters are estimated level by level, plugging in parameters from previous levels at each step. Let

$$l_{C,l} = \sum_{k=1}^{n} \sum_{j=1}^{l} \sum_{i=1}^{d-j} \log c_{i,i+j|v_{ij}}.$$

This is the sum of all log-copula densities up to and including level l, over all observations. Note that $l_{C,d-1} = l_C$ for the top level $l = d-1$. Now, let $\boldsymbol{\theta}_l$ denote the parameters of all pair-copulae at level l. Then, $l_{C,l}$ is a function of the parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l$ from level 1 up to l, but not of $\boldsymbol{\theta}_{l+1}, \ldots, \boldsymbol{\theta}_{d-1}$ from the following levels. Analogously to (3.1), define the l-level pseudo log-likelihood function $l_{C,P,l}$ as

$$l_{C,P,l}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = l_{C,l}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l; \boldsymbol{u}_{1n}, \ldots, \boldsymbol{u}_{nn}), \tag{3.2}$$

i.e. by substituting the parametric margins for the empirical ones. At the top level, (3.2) is simply the pseudo log-likelihood function from (3.1), when seen as a function of $\boldsymbol{\theta}$. Nonetheless, the SSP top level estimates are different from the SP ones. The estimation procedure is as follows:

- maximise $l_{C,P,1}(\boldsymbol{\theta}_1; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ over $\boldsymbol{\theta}_1$ to obtain $\hat{\boldsymbol{\theta}}_1^{SSP}$.
- for level $l = 2, \ldots, d-1$

maximise $l_{C,P,l}(\hat{\boldsymbol{\theta}}_1^{SSP}, \ldots, \hat{\boldsymbol{\theta}}_{l-1}^{SSP}, \boldsymbol{\theta}_l; \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ over $\boldsymbol{\theta}_l$
to obtain $\hat{\boldsymbol{\theta}}_l^{SSP}$.

For a detailed estimation algorithm and description of how to compute $l_{C,P,l}$, $l = 1, \ldots, d-1$, see Hobæk Haff (2010). Note that when none of the copulae constituting the structure share parameters, as in the models we use for the comparison of estimators (Section 4), the optimisation is done for each pair-copula, individually. Also, the necessary pair-copula arguments at a given level are computed by iterative transformations of the pseudo observations, just as for the SP estimator (see Section 3.1). The difference is that the parameters plugged into the transformations are final estimates, while they are part of the simultaneous optimisation in SP estimation.

The SSP estimator does not take into account information the next levels might have on the parameters. Obviously, it is asymptotically less efficient than the SP estimator. Nonetheless, it is sensible to use SSP in any case, to get start values for the SP estimation. The question is, how much precision does one really gain by a subsequent SP estimation? Is it worth the extra time spent?

4 Comparison

We want to study how the SP and SSP estimators perform for pair-copula constructions, both in terms of computing time and finite sample bias and variance. As earlier mentioned, we concentrate on the subgroup of D-vines. More specifically, we base all experiments, but one, on five dimensional D-vines (as depicted in Figure 2.1), varying the copula types. As explained earlier, the effect of the choice of margins on the dependence parameter estimates has already been extensively studied. Therefore we let the margins be uniform $U[0, 1]$ in all the experiments.

The object of the study is to explore how the estimators' performance is affected by the type and degree of dependence, the number of observations n and the correctness of the model. We also include one large dimensional example.

4.1 Type of dependence

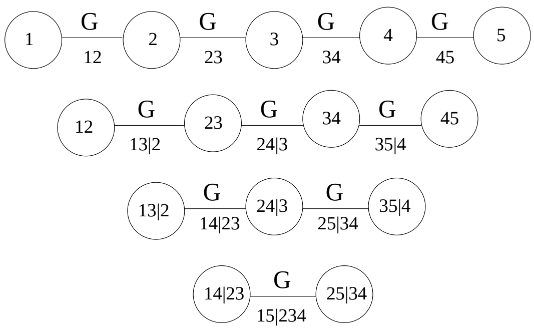
By type, we mean presence, or not, of tail dependence and dependence asymmetry. We account for this by considering three copula families, namely

- the Gaussian (no tail dependence),
- the Student's t (tail dependence)
- and the Clayton copula (lower, but not upper tail dependence),

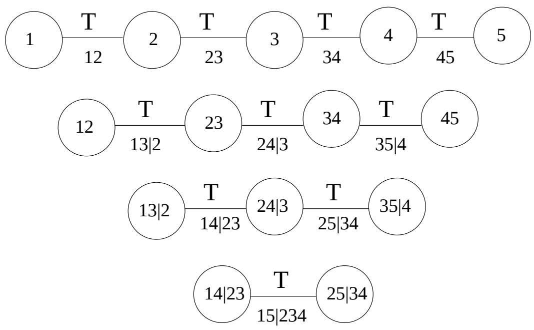
combined in four different models. The first three of these consist of only one of the copula types, whereas the last is a mix of all three types (see Figure 4.1).

Moreover, we have varied the degree of dependence by means of different parameter values. More specifically, we let the correlation parameter ρ of the Gaussian and Student's t copulae take the values $\{0.2, 0.5, 0.8\}$, corresponding to low, medium and high dependence, respectively. To facilitate comparison, we chose to consider the values $\{0.294, 1, 2.88\}$ for the parameter α of the Clayton copulae, that give the same Kendall's τ coefficients as the ρ values above. Furthermore, we fixed the number of degrees of freedom of the Student's t copulae to $\nu = 6$, which ensures a rather strong tail dependence.

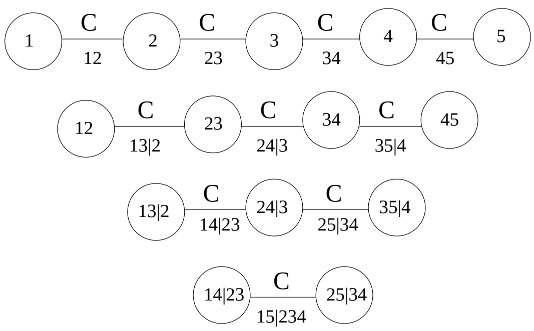
In all experiments, we have generated $n = 5000$ independent samples of the model in question, which is a relatively large sample size compared to most applications, estimated the parameters using SP and SSP, and repeated this $N = 1000$



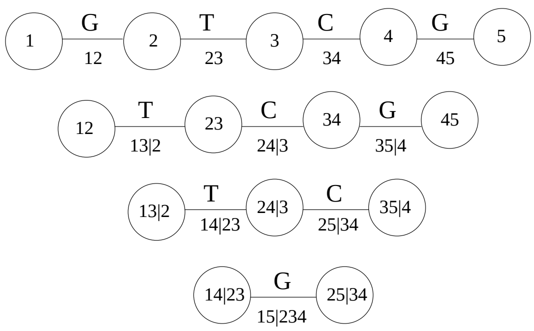
(a) Model 1



(b) Model 2



(c) Model 3



(d) Model 4

Figure 4.1. Models used for the comparison. The first three, (a), (b), (c), consist of only one of the three types of copulae, Gaussian (G), Student's t (T) and Clayton (C), respectively, while the last, (d), is a mix of the three.

times. The measures we have compared are

- the finite sample bias $Bias(\hat{\theta}_i, \theta_i) = \frac{1}{N}\sum_{k=1}^{N}(\hat{\theta}_{ik} - \theta_i)$,
- the mean squared error $MSE(\hat{\theta}_i, \theta_i) = \frac{1}{N}\sum_{k=1}^{N}(\hat{\theta}_{ik} - \theta_i)^2$
- and the CPU time,

the former two being computed for each parameter θ_i, $i = 1, \dots, |\boldsymbol{\theta}|$. All computations were programmed in R, version 2.11.1, and run on a computer with 72 GB RAM, 8 CPU kernels and hyperthreading, that allows for 16 simultaneous threads.

The results are summarised in Table 4.1 for Models 4.1a, 4.1b and 4.1c, and in Table 4.2 for Model 4.1d. In the former table, the results are averaged over each of the four levels of the structure, while they are presented per copula in the latter. Starting with the models consisting of copulae of the same type (Table 4.1), the bias and MSE of two estimators decrease with the degree of dependence. The reason is probably that the log-likelihood function steepens at stronger dependence. Further, the two measures appear to be rather constant over the four levels at low and medium dependence, while they increase with the level at high dependence. As explained in Section 3, the arguments of pair-copulae above the ground level are obtained though iterative transformations of the original pseudo observations. These transformations are functions of copulae at lower levels, and therefore depend on their parameters. The increasing estimator variance with level number may indicate a stronger sensitivity to repeated transformations when the degree of dependence is high. Note that the SP and SSP estimates for the Gaussian PCC are virtually the same. This is as anticipated, since both estimators are semiparametrically efficient for that particular model (Hobæk Haff, 2010). Of course, the SSP estimator is asymptotically less efficient than the SP estimator for the Student's t and Clayton vines. This is reflected in higher bias and MSE. The difference is however mostly moderate to small, which indicates that the SSP estimator performs well relative to SP. Furthermore, the computing time of the former is much lower. The factor ranges from order 10^{-1} to 10^{-3} in favour of SSP. As one would expect, the most marked time gain is obtained for the Student's t vine, whose parameter vector is twice the size of the other two models'.

Figures 4.2 and 4.3 show bias and MSE ratios for the parameter estimates of Model 4.1b and 4.1c as functions of the values of ρ and α, respectively, averaged over each of the four levels of the structure. These plots strengthen the impressions from Table 4.1. For the Student's t vines (Model 4.1b), the bias and MSE ratios appear to increase with the degree of dependence. Moreover, the increase is largest at the top levels. This means that the SP estimator performs better relative to the SSP one with growing dependence, especially for the higher level parameters. The reason may be the earlier suggested stronger sensitivity to repeated transformations, which are likely to affect the SSP estimator more, due to its sequential nature. For the Clayton vines (Model 4.1c) on the other hand, the bias and MSE ratios seem to grow with the dependence up to a certain point, before they start decreasing again. There is no apparent reason for this different behaviour. It could relate to characteristics of the Clayton copula, but may also be artifacts.

For the mixed vine of Model 4.1d (Table 4.2), the results are rather similar. The overall degree of dependence is medium. As expected, the bias and MSE are rather constant over the different levels for copulae of the same type. Again the SSP estimator performs quite well compared to the SP estimator in terms of efficiency, while its CPU time is drastically lower.

4.2 Number of observations n

The experiments in Section 4.1 were based on a large number of observations. To see how the two estimators perform on samples of smaller size, we repeated the simulations from Models 4.1a, 4.1b and 4.1c, reducing n, first to 500, and then to 50. The results from the simulations with medium dependence are summarised in Table 4.3. As the sample size decreases, we expect the estimators' variance to increase and finally explode. For $n = 500$, the finite sample bias and MSE of the correlations ρ and the parameters α are actually not discouragingly high. Estimation of the degrees of freedom parameters ν apparently requires a larger sample size. As n decreases to 50, none of the parameters are well estimated. Once more, the efficiency of the SSP estimator is rather good compared to the SP estimator, except for ν. Actually, the SSP estimator appears to suffer less under reduced sample size than its competitor. That could relate to the fact that the numerical optimisations are performed for each copula individually, as opposed to all at once. Also, the computing time of the former is lower.

4.3 Robustness

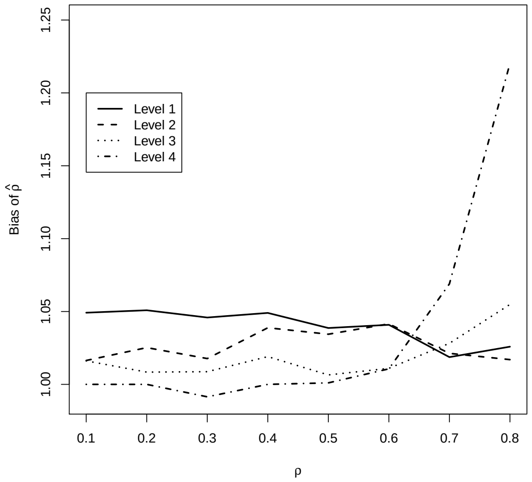
Since they are based on the empirical margins, both the SP and the SSP estimators are robust to deviations from the chosen marginal distributions. However, their properties do rely on the correctness of the specified dependence structure.

As described in Aas et al. (2009), selecting the pair-copulae constituting the PCC for a given dataset is a levelwise procedure. At the ground level, it consists in a set of bivariate copula selection problems, that may be handled for instance with copula goodness-of-fit tests (Berg, 2009; Genest et al., 2009). However, as one proceeds into the structure, one must condition on the pair-copulae chosen at preceding levels to be able to compute the necessary conditional distributions,

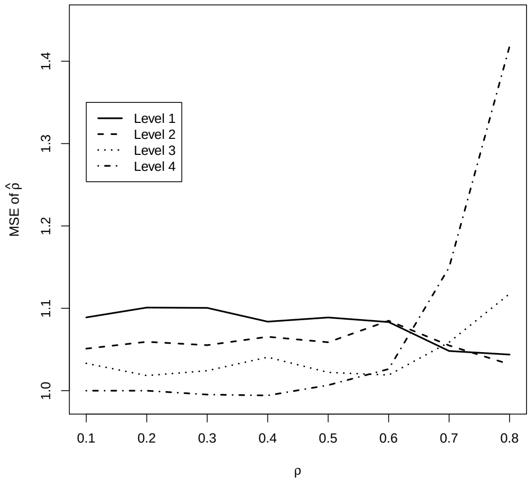
Degree of dependence			SP					
			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.49 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	$3.96 \cdot 10^{-3}$	$2.44 \cdot 10^{-5}$
		2	$1.09 \cdot 10^{-2}$	$1.87 \cdot 10^{-4}$	$8.31 \cdot 10^{-3}$	$1.08 \cdot 10^{-4}$	$4.16 \cdot 10^{-3}$	$2.75 \cdot 10^{-5}$
		3	$1.09 \cdot 10^{-2}$	$1.86 \cdot 10^{-4}$	$8.82 \cdot 10^{-3}$	$1.20 \cdot 10^{-4}$	$4.26 \cdot 10^{-3}$	$2.88 \cdot 10^{-5}$
		4	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.41 \cdot 10^{-3}$	$1.13 \cdot 10^{-4}$	$9.48 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
		CPU	$1.79 \cdot 10^{-2}$		$5.33 \cdot 10^{-2}$		$4.05 \cdot 10^{-1}$	
Student's t	ρ	1	$1.18 \cdot 10^{-2}$	$2.18 \cdot 10^{-4}$	$9.31 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$4.64 \cdot 10^{-3}$	$3.42 \cdot 10^{-5}$
		2	$1.19 \cdot 10^{-2}$	$2.19 \cdot 10^{-4}$	$9.30 \cdot 10^{-3}$	$1.36 \cdot 10^{-4}$	$4.71 \cdot 10^{-3}$	$3.43 \cdot 10^{-5}$
		3	$1.19 \cdot 10^{-2}$	$2.18 \cdot 10^{-4}$	$9.26 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$4.74 \cdot 10^{-3}$	$3.57 \cdot 10^{-5}$
		4	$1.20 \cdot 10^{-2}$	$2.27 \cdot 10^{-4}$	$9.75 \cdot 10^{-3}$	$1.49 \cdot 10^{-4}$	$1.09 \cdot 10^{-2}$	$1.53 \cdot 10^{-4}$
	ν	1	$4.88 \cdot 10^{-1}$	$3.79 \cdot 10^{-1}$	$3.96 \cdot 10^{-1}$	$2.54 \cdot 10^{-1}$	$3.86 \cdot 10^{-1}$	$2.41 \cdot 10^{-1}$
		2	$4.81 \cdot 10^{-1}$	$3.76 \cdot 10^{-1}$	$4.29 \cdot 10^{-1}$	$3.07 \cdot 10^{-1}$	$3.10 \cdot 10^{-1}$	$1.56 \cdot 10^{-1}$
		3	$4.96 \cdot 10^{-1}$	$4.04 \cdot 10^{-1}$	$4.58 \cdot 10^{-1}$	$3.51 \cdot 10^{-1}$	$3.08 \cdot 10^{-1}$	$1.52 \cdot 10^{-1}$
		4	$5.09 \cdot 10^{-1}$	$4.28 \cdot 10^{-1}$	$5.17 \cdot 10^{-1}$	$4.50 \cdot 10^{-1}$	$3.01 \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$
		CPU	1.66		1.97		1.08	
Clayton	α	1	$1.66 \cdot 10^{-2}$	$4.28 \cdot 10^{-4}$	$2.65 \cdot 10^{-2}$	$1.09 \cdot 10^{-3}$	$5.91 \cdot 10^{-2}$	$5.41 \cdot 10^{-3}$
		2	$1.62 \cdot 10^{-2}$	$4.21 \cdot 10^{-4}$	$2.37 \cdot 10^{-2}$	$8.94 \cdot 10^{-4}$	$5.68 \cdot 10^{-2}$	$4.85 \cdot 10^{-3}$
		3	$1.67 \cdot 10^{-2}$	$4.31 \cdot 10^{-4}$	$2.63 \cdot 10^{-2}$	$1.10 \cdot 10^{-3}$	$1.24 \cdot 10^{-1}$	$1.95 \cdot 10^{-2}$
		4	$1.67 \cdot 10^{-2}$	$4.46 \cdot 10^{-4}$	$2.78 \cdot 10^{-2}$	$1.24 \cdot 10^{-3}$	$6.12 \cdot 10^{-1}$	$3.83 \cdot 10^{-1}$
		CPU	$1.88 \cdot 10^{-2}$		$6.49 \cdot 10^{-2}$		$6.07 \cdot 10^{-2}$	

Degree of dependence			SSP					
			Low		Medium		High	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$1.09 \cdot 10^{-2}$	$1.84 \cdot 10^{-4}$	$8.47 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	$3.98 \cdot 10^{-3}$	$2.45 \cdot 10^{-5}$
		2	$1.09 \cdot 10^{-2}$	$1.87 \cdot 10^{-4}$	$8.31 \cdot 10^{-3}$	$1.08 \cdot 10^{-4}$	$4.16 \cdot 10^{-3}$	$2.75 \cdot 10^{-5}$
		3	$1.09 \cdot 10^{-2}$	$1.86 \cdot 10^{-4}$	$8.82 \cdot 10^{-3}$	$1.20 \cdot 10^{-4}$	$4.27 \cdot 10^{-3}$	$2.90 \cdot 10^{-5}$
		4	$1.09 \cdot 10^{-2}$	$1.85 \cdot 10^{-4}$	$8.41 \cdot 10^{-3}$	$1.13 \cdot 10^{-4}$	$9.49 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
		CPU	$1.63 \cdot 10^{-3}$		$2.98 \cdot 10^{-3}$		$1.67 \cdot 10^{-2}$	
Student's t	ρ	1	$1.24 \cdot 10^{-2}$	$2.40 \cdot 10^{-4}$	$9.67 \cdot 10^{-3}$	$1.47 \cdot 10^{-4}$	$4.76 \cdot 10^{-3}$	$3.57 \cdot 10^{-5}$
		2	$1.22 \cdot 10^{-2}$	$2.32 \cdot 10^{-4}$	$9.62 \cdot 10^{-3}$	$1.44 \cdot 10^{-4}$	$4.79 \cdot 10^{-3}$	$3.54 \cdot 10^{-5}$
		3	$1.20 \cdot 10^{-2}$	$2.22 \cdot 10^{-4}$	$9.32 \cdot 10^{-3}$	$1.38 \cdot 10^{-4}$	$5.00 \cdot 10^{-3}$	$3.99 \cdot 10^{-5}$
		4	$1.20 \cdot 10^{-2}$	$2.27 \cdot 10^{-4}$	$9.76 \cdot 10^{-3}$	$1.50 \cdot 10^{-4}$	$1.33 \cdot 10^{-2}$	$2.17 \cdot 10^{-4}$
	ν	1	$5.13 \cdot 10^{-1}$	$4.26 \cdot 10^{-1}$	$5.13 \cdot 10^{-1}$	$4.22 \cdot 10^{-1}$	$5.31 \cdot 10^{-1}$	$4.70 \cdot 10^{-1}$
		2	$5.06 \cdot 10^{-1}$	$4.15 \cdot 10^{-1}$	$5.03 \cdot 10^{-1}$	$4.26 \cdot 10^{-1}$	$5.15 \cdot 10^{-1}$	$4.38 \cdot 10^{-1}$
		3	$5.10 \cdot 10^{-1}$	$4.30 \cdot 10^{-1}$	$4.94 \cdot 10^{-1}$	$4.08 \cdot 10^{-1}$	$5.06 \cdot 10^{-1}$	$4.38 \cdot 10^{-1}$
		4	$5.11 \cdot 10^{-1}$	$4.35 \cdot 10^{-1}$	$5.18 \cdot 10^{-1}$	$4.47 \cdot 10^{-1}$	$5.23 \cdot 10^{-1}$	$4.68 \cdot 10^{-1}$
		CPU	$1.22 \cdot 10^{-2}$		$1.41 \cdot 10^{-2}$		$6.38 \cdot 10^{-3}$	
Clayton	α	1	$1.72 \cdot 10^{-2}$	$4.57 \cdot 10^{-4}$	$2.90 \cdot 10^{-2}$	$1.32 \cdot 10^{-3}$	$5.76 \cdot 10^{-2}$	$5.29 \cdot 10^{-3}$
		2	$1.66 \cdot 10^{-2}$	$4.40 \cdot 10^{-4}$	$2.77 \cdot 10^{-2}$	$1.21 \cdot 10^{-3}$	$5.72 \cdot 10^{-2}$	$5.06 \cdot 10^{-3}$
		3	$1.70 \cdot 10^{-2}$	$4.47 \cdot 10^{-4}$	$2.83 \cdot 10^{-2}$	$1.27 \cdot 10^{-3}$	$1.24 \cdot 10^{-1}$	$1.93 \cdot 10^{-2}$
		4	$1.67 \cdot 10^{-2}$	$4.48 \cdot 10^{-4}$	$2.81 \cdot 10^{-2}$	$1.29 \cdot 10^{-3}$	$6.96 \cdot 10^{-1}$	$4.93 \cdot 10^{-1}$
		CPU	$1.74 \cdot 10^{-3}$		$1.35 \cdot 10^{-3}$		$3.90 \cdot 10^{-4}$	

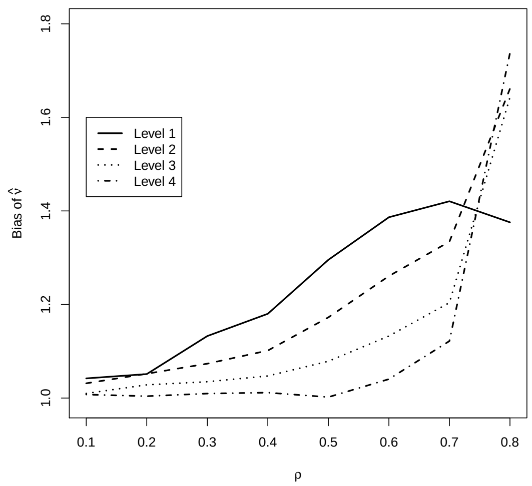
Table 4.1. Results from simulations of $n = 5000$ observations from Model 4.1a, 4.1b and 4.1c (see Figure 4.1), i.e. consisting of only one copula type, with low ($\rho = 0.2, \alpha = 0.294$), medium ($\rho = 0.5, \alpha = 1$) and high ($\rho = 0.8, \alpha = 2.88$) dependence.



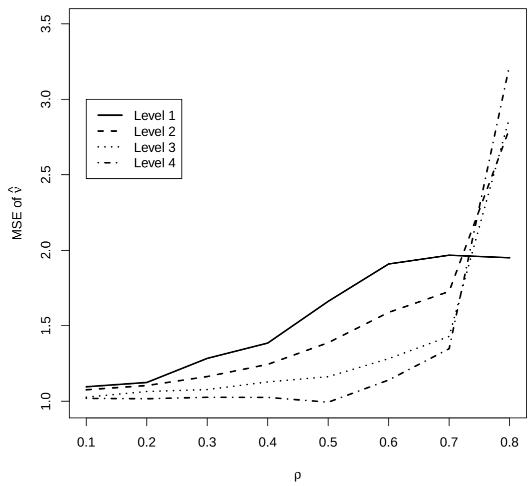
(a) Model 2: Bias ratio of $\hat{\rho}$



(b) Model 2: MSE ratio of $\hat{\rho}$



(c) Model 2: Bias ratio of $\hat{\nu}$

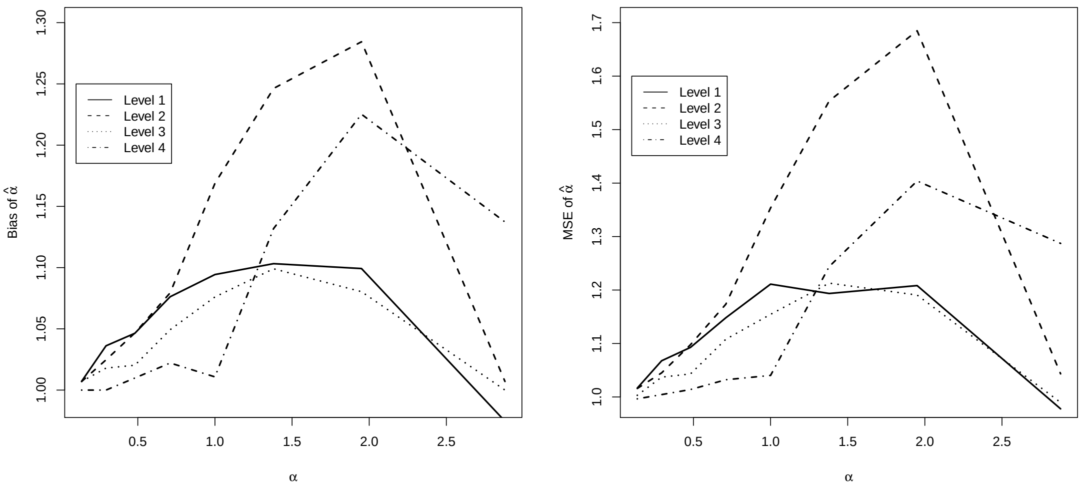


(d) Model 2: MSE ratio of $\hat{\nu}$

Figure 4.2. Bias and MSE ratios (to the left and right, respectively) of the parameter estimates $\hat{\rho}$ (top row) and $\hat{\nu}$ (bottom row) of Model 4.1b as a function of ρ, averaged over each of the four levels.

		SP		SSP	
Copula	Par.	Bias	MSE	Bias	MSE
12	ρ	$7.66 \cdot 10^{-3}$	$9.04 \cdot 10^{-5}$	$8.63 \cdot 10^{-3}$	$1.17 \cdot 10^{-4}$
23	ρ	$8.08 \cdot 10^{-3}$	$1.01 \cdot 10^{-4}$	$9.42 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$
	ν	$2.94 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	$5.23 \cdot 10^{-1}$	$4.43 \cdot 10^{-1}$
34	α	$2.55 \cdot 10^{-2}$	$1.04 \cdot 10^{-3}$	$2.79 \cdot 10^{-2}$	$1.26 \cdot 10^{-3}$
45	ρ	$7.18 \cdot 10^{-3}$	$8.04 \cdot 10^{-5}$	$8.27 \cdot 10^{-3}$	$1.07 \cdot 10^{-4}$
13|2	ρ	$9.13 \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$	$9.60 \cdot 10^{-3}$	$1.45 \cdot 10^{-4}$
	ν	$4.15 \cdot 10^{-1}$	$2.71 \cdot 10^{-1}$	$4.84 \cdot 10^{-1}$	$3.80 \cdot 10^{-1}$
24|3	α	$2.81 \cdot 10^{-2}$	$1.25 \cdot 10^{-3}$	$2.87 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$
35|4	ρ	$7.84 \cdot 10^{-3}$	$9.68 \cdot 10^{-5}$	$8.82 \cdot 10^{-3}$	$1.21 \cdot 10^{-4}$
14|23	ρ	$9.80 \cdot 10^{-3}$	$1.53 \cdot 10^{-4}$	$1.01 \cdot 10^{-2}$	$1.62 \cdot 10^{-4}$
	ν	$4.86 \cdot 10^{-1}$	$3.80 \cdot 10^{-1}$	$5.14 \cdot 10^{-1}$	$4.23 \cdot 10^{-1}$
25|34	α	$2.76 \cdot 10^{-2}$	$1.19 \cdot 10^{-3}$	$2.79 \cdot 10^{-2}$	$1.22 \cdot 10^{-3}$
15|234	ρ	$8.15 \cdot 10^{-3}$	$1.07 \cdot 10^{-4}$	$8.19 \cdot 10^{-3}$	$1.09 \cdot 10^{-4}$
CPU		$3.04 \cdot 10^{-1}$		$8.10 \cdot 10^{-4}$	

Table 4.2. Results from simulations of $n = 5000$ observations from Model 4.1d (see Figure 4.1), i.e. consisting of different copula types, with medium ($\rho = 0.5, \alpha = 1$) dependence.



(a) Model 3: Bias ratio of $\hat{\alpha}$

(b) Model 3: MSE ratio of $\hat{\alpha}$

Figure 4.3. Bias and MSE ratios (to the left and right, respectively) of the parameter estimates $\hat{\alpha}$ for Model 4.1c as a function of α, averaged over each of the four levels.

Degree of dependence			SP			
			n=500		n=50	
Model	Par.	Level	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$2.73 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.09 \cdot 10^{-2}$	$1.25 \cdot 10^{-2}$
		2	$2.71 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.10 \cdot 10^{-2}$	$1.30 \cdot 10^{-2}$
		3	$2.69 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.24 \cdot 10^{-2}$	$1.35 \cdot 10^{-2}$
		4	$2.68 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.75 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$
		CPU	$2.10 \cdot 10^{-4}$		$1.70 \cdot 10^{-4}$	
Student's t	ρ	1	$3.00 \cdot 10^{-2}$	$1.43 \cdot 10^{-3}$	$9.90 \cdot 10^{-2}$	$1.50 \cdot 10^{-2}$
		2	$3.04 \cdot 10^{-2}$	$1.44 \cdot 10^{-3}$	$9.82 \cdot 10^{-2}$	$1.52 \cdot 10^{-2}$
		3	$2.99 \cdot 10^{-2}$	$1.42 \cdot 10^{-3}$	$1.04 \cdot 10^{-1}$	$1.69 \cdot 10^{-2}$
		4	$3.13 \cdot 10^{-2}$	$1.53 \cdot 10^{-3}$	$1.05 \cdot 10^{-1}$	$1.78 \cdot 10^{-2}$
	ν	1	1.48	4.91	20.3	2030
		2	1.56	6.03	24.6	2430
		3	1.79	8.57	27.9	2930
		4	2.10	12.0	30.0	3020
		CPU	$9.44 \cdot 10^{-3}$		$1.35 \cdot 10^{-3}$	
Clayton	α	1	$8.21 \cdot 10^{-2}$	$1.06 \cdot 10^{-2}$	$2.93 \cdot 10^{-1}$	$1.42 \cdot 10^{-1}$
		2	$7.97 \cdot 10^{-2}$	$1.00 \cdot 10^{-2}$	$2.75 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$
		3	$8.46 \cdot 10^{-2}$	$1.09 \cdot 10^{-2}$	$2.88 \cdot 10^{-1}$	$1.33 \cdot 10^{-1}$
		4	$9.41 \cdot 10^{-2}$	$1.31 \cdot 10^{-2}$	$3.41 \cdot 10^{-1}$	$1.81 \cdot 10^{-1}$
		CPU	$4.50 \cdot 10^{-4}$		$1.00 \cdot 10^{-5}$	

Degree of dependence			SSP			
			n=500		n=50	
Model	Par.	Level	Bias	MSE	Bias	MSE
Gaussian	ρ	1	$2.72 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.18 \cdot 10^{-2}$	$1.30 \cdot 10^{-2}$
		2	$2.72 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.16 \cdot 10^{-2}$	$1.31 \cdot 10^{-2}$
		3	$2.69 \cdot 10^{-2}$	$1.15 \cdot 10^{-3}$	$9.22 \cdot 10^{-2}$	$1.34 \cdot 10^{-2}$
		4	$2.68 \cdot 10^{-2}$	$1.16 \cdot 10^{-3}$	$9.74 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$
		CPU	$2.00 \cdot 10^{-5}$		$1.00 \cdot 10^{-5}$	
Student's t	ρ	1	$3.13 \cdot 10^{-2}$	$1.54 \cdot 10^{-3}$	$1.01 \cdot 10^{-1}$	$1.56 \cdot 10^{-2}$
		2	$3.11 \cdot 10^{-2}$	$1.51 \cdot 10^{-3}$	$9.88 \cdot 10^{-2}$	$1.52 \cdot 10^{-2}$
		3	$3.04 \cdot 10^{-2}$	$1.46 \cdot 10^{-3}$	$1.03 \cdot 10^{-1}$	$1.67 \cdot 10^{-2}$
		4	$3.16 \cdot 10^{-2}$	$1.54 \cdot 10^{-3}$	$1.04 \cdot 10^{-1}$	$1.86 \cdot 10^{-2}$
	ν	1	2.60	142.00	89.0	24900
		2	2.35	82.2	97.0	27100
		3	2.98	203.00	111.00	31300
		4	2.62	52.7	127.00	35900
		CPU	$3.00 \cdot 10^{-5}$		$< 1.00 \cdot 10^{-18}$	
Clayton	α	1	$9.01 \cdot 10^{-2}$	$1.28 \cdot 10^{-2}$	$3.23 \cdot 10^{-1}$	$1.74 \cdot 10^{-1}$
		2	$9.03 \cdot 10^{-2}$	$1.27 \cdot 10^{-2}$	$2.83 \cdot 10^{-1}$	$1.30 \cdot 10^{-1}$
		3	$8.97 \cdot 10^{-2}$	$1.24 \cdot 10^{-2}$	$2.88 \cdot 10^{-1}$	$1.27 \cdot 10^{-1}$
		4	$1.01 \cdot 10^{-1}$	$1.52 \cdot 10^{-2}$	$3.55 \cdot 10^{-1}$	$1.79 \cdot 10^{-1}$
		CPU	$3.00 \cdot 10^{-5}$		$1.00 \cdot 10^{-5}$	

Table 4.3. Results from simulations of $n = 500$ and $n = 50$ observations from Model 4.1a, 4.1b and 4.1c (see Figure 4.1), i.e. consisting of only one copula type, with medium ($\rho = 0.5, \alpha = 1$) dependence.

that are pair-copula arguments (as explained in Section 3). One would therefore expect the difficulty of finding adequate copulae to increase with the level number.

We wish to explore how the two estimators perform when the specified model deviates from the true model. In order to do that, we have perturbed the original models (from 4.1). More specifically, we have mixed the copulae of each of the two models 4.1b and 4.1c with the copulae of Model 4.1a, using the same degree of dependence and a fixed mixing probability p. For instance, the pair (u_1, u_2) of observations is drawn from a Student's t-copula (Clayton copula) with probability $1 - p$ and from a Gaussian copula with probability p. In all experiments, we used $n = 5000$ and medium dependence, letting the mixing probability take each of the values $\{0.05, 0.1, 0.2\}$. The results are summarised by level in Table 4.4.

Both the SP and the SSP estimators of the correlation parameters of Model 4.1b perform almost as well as in the non-perturbed case (Table 4.1). The reason for this is probably that the correlations of the model we are mixing with have the same values. The corresponding degrees of freedom estimates are however decreasingly accurate as the mixing probability grows. That is also the case for the parameters of the Clayton copulae of Model 4.1c, as one would expect. Moreover, the computing times of both estimators grow with the mixing probability, due to the need for extra iterations before convergence. Finally, the difference between the performance of the SP and the SSP estimators is reduced with an increasing degree of perturbation, maybe because the latter uses information only from preceding levels, and not from the following. Hence, when the model assumptions are not completely accurate, the gain from using the SP estimator seems to be smaller.

4.4 Large dimension d

As mentioned earlier, the SP estimator is computationally too demanding and time consuming for high dimensional problems. Our belief is that the SSP might be a good alternative in such cases. We have therefore tried it on a 50 dimensional D-vine of Student's copulae with $\rho = 0.2$, corresponding to low dependence. This model has as many as 2450 parameters. Optimising over so many parameters simultaneously would not only be highly time consuming, but also numerically dubious. We have therefore only considered the SSP estimator for this model. Moreover, we let $n = 5000$ and $N = 1000$ as in the experiments of Section 4.1.

Figure 4.4 displays the bias and MSE of the parameter estimates $\hat{\rho}$ and $\hat{\nu}$, averaged over each level. These are rather low up to level 20 for $\hat{\rho}$ and up to level 30 for $\hat{\nu}$, after which they explode. This is due to numerical problems with the repeated transformations of the original data. After a certain level, the computed estimates

Degree of dependence			SP					
			$p = 0.05$		$p = 0.1$		$p = 0.2$	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	ρ	1	$9.17 \cdot 10^{-3}$	$1.32 \cdot 10^{-4}$	$9.37 \cdot 10^{-3}$	$1.36 \cdot 10^{-4}$	$9.24 \cdot 10^{-3}$	$1.34 \cdot 10^{-4}$
		2	$9.45 \cdot 10^{-3}$	$1.40 \cdot 10^{-4}$	$9.43 \cdot 10^{-3}$	$1.40 \cdot 10^{-4}$	$9.39 \cdot 10^{-3}$	$1.36 \cdot 10^{-4}$
		3	$9.28 \cdot 10^{-3}$	$1.35 \cdot 10^{-4}$	$9.49 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$	$9.57 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$
		4	$1.03 \cdot 10^{-2}$	$1.59 \cdot 10^{-4}$	$9.88 \cdot 10^{-3}$	$1.52 \cdot 10^{-4}$	$9.87 \cdot 10^{-3}$	$1.55 \cdot 10^{-4}$
	ν	1	$4.94 \cdot 10^{-1}$	$4.10 \cdot 10^{-1}$	$7.73 \cdot 10^{-1}$	$9.20 \cdot 10^{-1}$	1.67	3.47
		2	$5.25 \cdot 10^{-1}$	$4.77 \cdot 10^{-1}$	$7.48 \cdot 10^{-1}$	$9.16 \cdot 10^{-1}$	1.49	2.92
		3	$5.75 \cdot 10^{-1}$	$5.76 \cdot 10^{-1}$	$7.68 \cdot 10^{-1}$	$9.99 \cdot 10^{-1}$	1.43	2.79
		4	$5.95 \cdot 10^{-1}$	$6.22 \cdot 10^{-1}$	$8.05 \cdot 10^{-1}$	1.11	1.45	2.97
		CPU	$1.84 \cdot 10^{-1}$		$2.22 \cdot 10^{-1}$		$8.03 \cdot 10^{-1}$	
Clayton	α	1	$4.90 \cdot 10^{-2}$	$3.29 \cdot 10^{-3}$	$8.83 \cdot 10^{-2}$	$8.95 \cdot 10^{-3}$	$1.65 \cdot 10^{-1}$	$2.86 \cdot 10^{-2}$
		2	$3.82 \cdot 10^{-2}$	$2.11 \cdot 10^{-3}$	$6.50 \cdot 10^{-2}$	$5.19 \cdot 10^{-3}$	$1.21 \cdot 10^{-1}$	$1.57 \cdot 10^{-2}$
		3	$4.34 \cdot 10^{-2}$	$2.67 \cdot 10^{-3}$	$7.43 \cdot 10^{-2}$	$6.67 \cdot 10^{-3}$	$1.34 \cdot 10^{-1}$	$1.92 \cdot 10^{-2}$
		4	$6.67 \cdot 10^{-2}$	$5.66 \cdot 10^{-3}$	$1.15 \cdot 10^{-1}$	$1.47 \cdot 10^{-2}$	$1.95 \cdot 10^{-1}$	$3.96 \cdot 10^{-2}$
		CPU	$5.92 \cdot 10^{-3}$		$9.79 \cdot 10^{-3}$		$3.30 \cdot 10^{-2}$	
Degree of dependence			SSP					
			$p = 0.05$		$p = 0.1$		$p = 0.2$	
Model	Par.	Level	Bias	MSE	Bias	MSE	Bias	MSE
Student's t	ρ	1	$9.55 \cdot 10^{-3}$	$1.42 \cdot 10^{-4}$	$9.63 \cdot 10^{-3}$	$1.44 \cdot 10^{-4}$	$9.47 \cdot 10^{-3}$	$1.41 \cdot 10^{-4}$
		2	$9.72 \cdot 10^{-3}$	$1.48 \cdot 10^{-4}$	$9.73 \cdot 10^{-3}$	$1.48 \cdot 10^{-4}$	$9.62 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$
		3	$9.38 \cdot 10^{-3}$	$1.38 \cdot 10^{-4}$	$9.61 \cdot 10^{-3}$	$1.46 \cdot 10^{-4}$	$9.62 \cdot 10^{-3}$	$1.44 \cdot 10^{-4}$
		4	$1.03 \cdot 10^{-2}$	$1.61 \cdot 10^{-4}$	$9.93 \cdot 10^{-3}$	$1.53 \cdot 10^{-4}$	$9.89 \cdot 10^{-3}$	$1.56 \cdot 10^{-4}$
	ν	1	$5.87 \cdot 10^{-1}$	$5.89 \cdot 10^{-1}$	$7.75 \cdot 10^{-1}$	1.02	1.50	3.24
		2	$5.91 \cdot 10^{-1}$	$6.21 \cdot 10^{-1}$	$7.92 \cdot 10^{-1}$	1.07	1.47	3.02
		3	$6.19 \cdot 10^{-1}$	$6.72 \cdot 10^{-1}$	$8.30 \cdot 10^{-1}$	1.14	1.48	3.05
		4	$6.09 \cdot 10^{-1}$	$6.34 \cdot 10^{-1}$	$8.34 \cdot 10^{-1}$	1.15	1.50	3.10
		CPU	$1.22 \cdot 10^{-3}$		$1.60 \cdot 10^{-3}$		$5.71 \cdot 10^{-3}$	
Clayton	α	1	$3.17 \cdot 10^{-2}$	$1.58 \cdot 10^{-3}$	$4.20 \cdot 10^{-2}$	$2.58 \cdot 10^{-3}$	$7.57 \cdot 10^{-2}$	$6.80 \cdot 10^{-3}$
		2	$3.83 \cdot 10^{-2}$	$2.17 \cdot 10^{-3}$	$6.19 \cdot 10^{-2}$	$4.89 \cdot 10^{-3}$	$1.16 \cdot 10^{-1}$	$1.45 \cdot 10^{-2}$
		3	$5.23 \cdot 10^{-2}$	$3.71 \cdot 10^{-3}$	$9.26 \cdot 10^{-2}$	$9.80 \cdot 10^{-3}$	$1.67 \cdot 10^{-1}$	$2.90 \cdot 10^{-2}$
		4	$8.20 \cdot 10^{-2}$	$8.06 \cdot 10^{-3}$	$1.45 \cdot 10^{-1}$	$2.25 \cdot 10^{-2}$	$2.52 \cdot 10^{-1}$	$6.51 \cdot 10^{-2}$
		CPU	$3.57 \cdot 10^{-4}$		$4.20 \cdot 10^{-4}$		$1.13 \cdot 10^{-3}$	

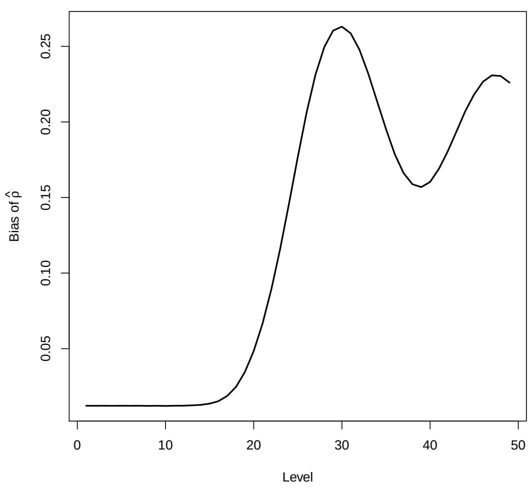
Table 4.4. Results from simulations of $n = 5000$ observations from Model 4.1b and 4.1c mixed with Model 4.1a (see Figure 4.1), also with medium dependence, for different mixing probabilities p.

in practice tend towards independence, i.e. partial correlations of approximately 0 and a high number of degrees of freedom. The subsequent decrease (and increase for $\hat{\rho}$) of the bias and MSE seems unintuitive. The correlation estimates $\hat{\rho}$ appear to fluctuate around 0, whereas the estimates $\hat{\nu}$ are almost halved on average. Actually, there is not that much difference between a bivariate t-copula with 100 and 200 degrees of freedom. Hence, this may just be an artifact. Note that what we are trying to estimate at these levels, is conditional dependence with a very high number of conditioning variables. It is not that surprising that it is difficult to estimate such higher order dependencies. The question is of course, how different is the estimated distribution from the true distribution?

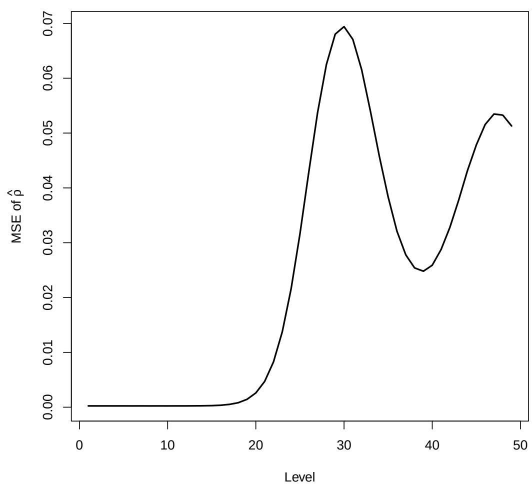
In a pair-copula construction of dimension d, only $d-1$ of the pairwise dependencies are modelled unconditionally. For instance, in this 50-dimensional D-vine, the dependence between U_1 and U_{40} is modelled through $C_{1,40|2\ldots39}$, i.e. conditionally on 38 variables, as well as conditionals in lower levels, involving a subset of the 40 variables in question. In most applications, it is the unconditional dependence, or at least lower order dependencies, one is interested in. As long as the bottom levels are well estimated, one may hope that the imprecise estimates for the top levels do not affect the lower order dependencies too much.

To study this, we have generated $3 \times n = 5000$ from one of the estimated Student's t-vines, and compared certain characteristics of these samples to those of samples from the true distribution. More specifically, we have computed all (unconditional) pairwise empirical Kendall's τ coefficients based on $50,000$ samples from the true distribution, which should be rather close to the true coefficients, as well as for the simulations from the estimated distribution. We have also calculated the 90%, 95%, 97.5% and 99% quantiles of $u_i + u_j$, for all pairs (i, j). These are displayed in Figure 4.5 and 4.6, respectively. The values are averaged over the level the corresponding conditional dependence belongs to. For instance, the dependence between U_1 and U_{40} is modelled conditionally at level 39. The quantiles and Kendall's τ coefficients for this pair therefore contribute to the mean at level 39 in the plots. As expected, the Kendall's τ coefficients and quantiles corresponding to the lower levels of the structure appear to be close to the true values. However, they are fairly good also for the pairs modelled in the top levels, even though the conditional dependence between these pairs has been highly underestimated. Hence, the parameter estimates for the upper levels do not seem to have that much effect on the lower order dependencies. This is an argument for truncating large structures after a certain level, letting the top level copulae be independence copulae Brechmann et al. (2010).

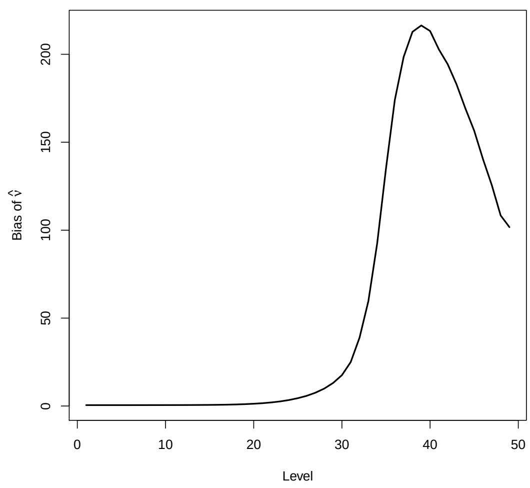
Since the bias and MSE curves were rather different from what we had anticipated, we repeated the above experiments with a t-copula of dimension $d = 50$, all pairwise correlations set to 0.2 and $\nu = 6$. That is actually a special case of a D-vine with t-copulae. At a given level l, the correlation parameters are the corresponding partial correlations, and the number of degrees of freedom is $\nu + l - 1 = 5 + l$. We simulated from this D-vine as described earlier, and estimated its parameters using the SSP estimator. Since the correlation parameters decrease with each level, whereas the degrees of freedom increase, it is more natural to consider the relative bias and MSE in this case, i.e. $Bias(\hat{\theta}_i, \theta_i)/\theta_i$ and $MSE(\hat{\theta}_i, \theta_i)/\theta_i^2$. These are shown in Figure 4.7. We have also computed the average Kendall's τ coefficients and quantiles per level, based on simulations from estimated distributions, displayed in Figure 4.8 and 4.9, respectively. This is an unnecessarily cumbersome way of estimating the parameters of a t-copula. The purpose of this experiment was only to investigate whether either the simulation or the estimation routines are flawed on an example for which we know the true Kendall's τ coefficients and quantiles. We note that the results are reassuring. The relative bias and MSE, increased steadily with the level of the structure, as one would expect. Moreover, all the pairwise Kendall's τ coefficients fluctuated around the true value of about 0.128, and likewise for the quantiles.



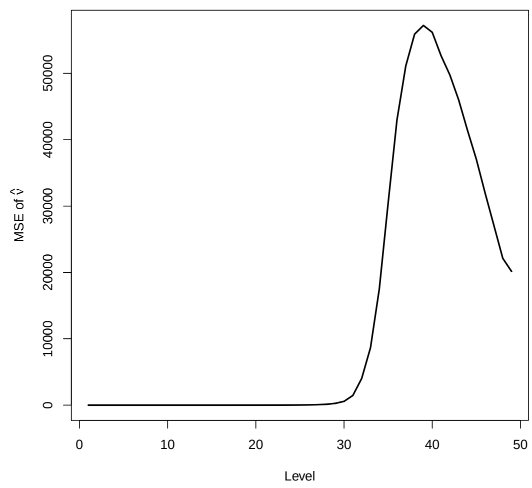
(a) Bias of $\hat{\rho}$



(b) MSE of $\hat{\rho}$



(c) Bias of $\hat{\nu}$



(d) MSE of $\hat{\nu}$

Figure 4.4. Bias and MSE of the parameter estimates of the Student's t-vine in 50 dimensions, with low dependence ($\rho = 0.2$), averaged over each level.

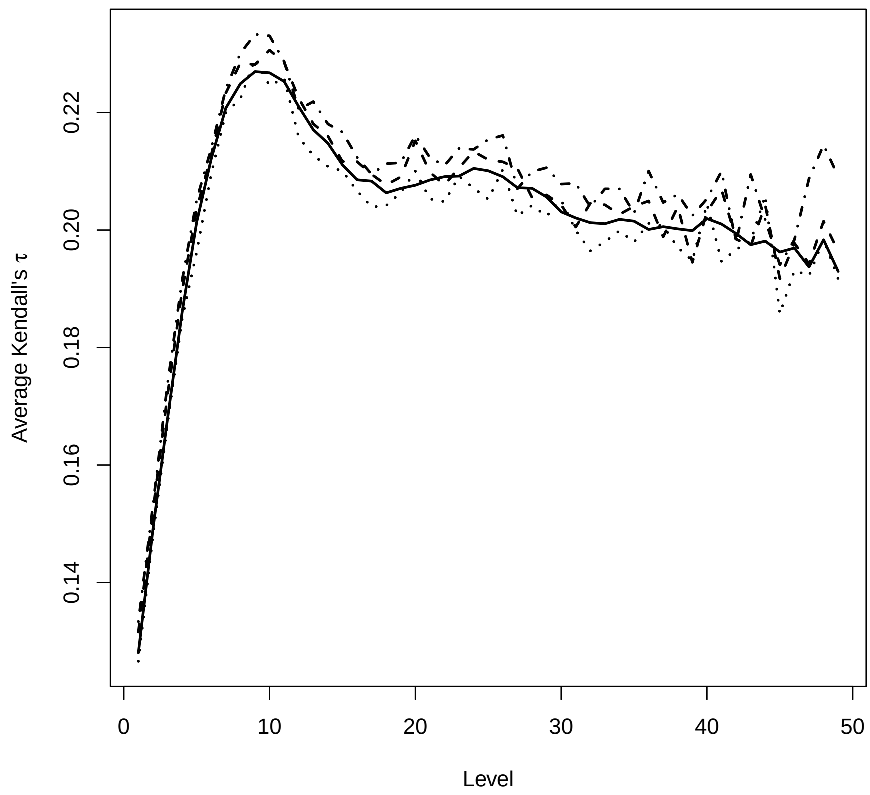


Figure 4.5. Empirical pairwise Kendall's τ coefficients for $50,000$ samples from the true distribution (connected line) and three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modelled conditional dependence belongs to.

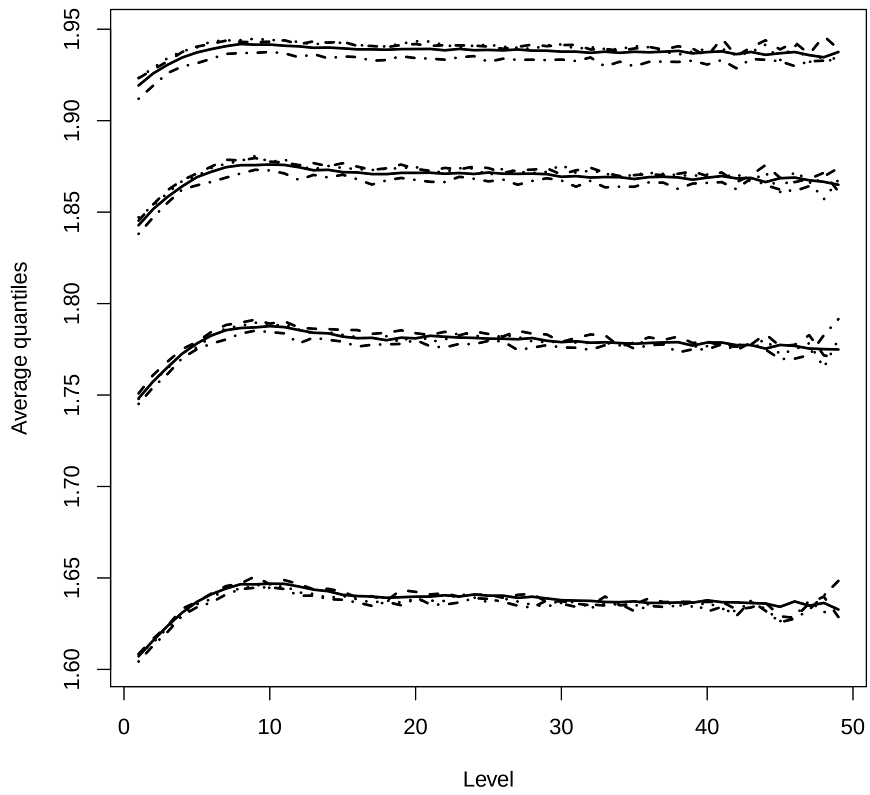
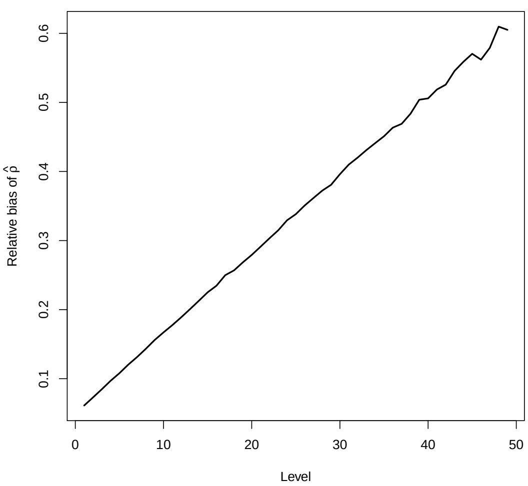
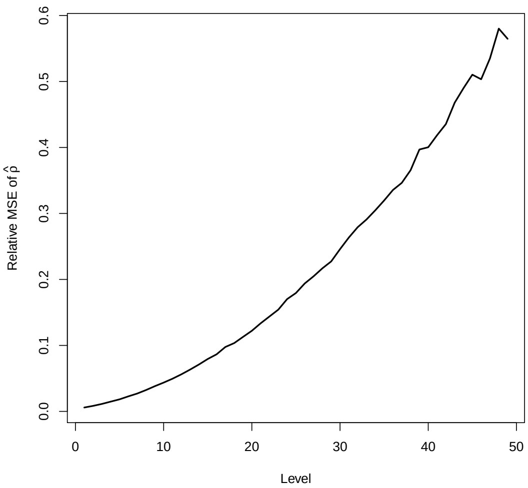


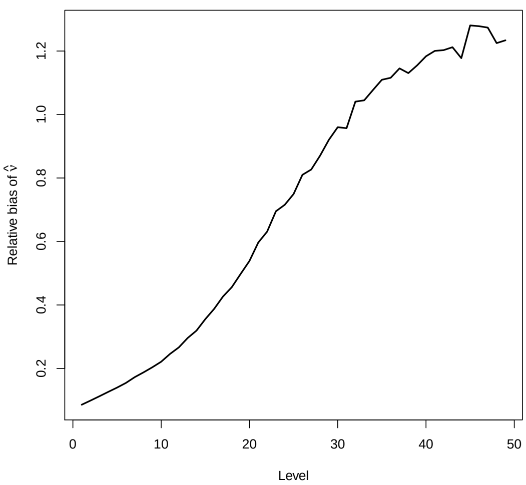
Figure 4.6. Estimated 90%, 95%, 97.5% and 99% quantiles for $50,000$ samples from the true distribution (connected lines) and three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modelled conditional dependence belongs to.



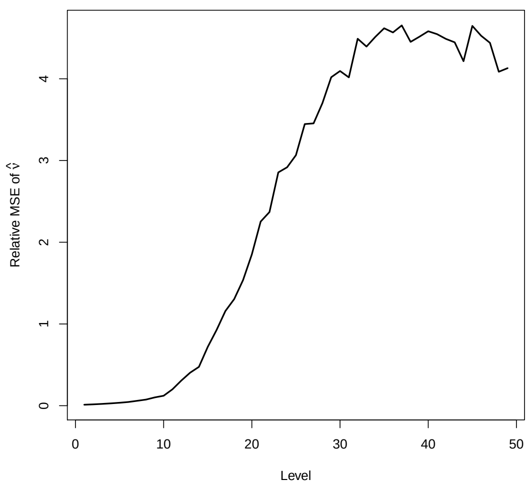
(a) Relative bias of $\hat{\rho}$



(b) Relative MSE of $\hat{\rho}$



(c) Relative bias of $\hat{\nu}$



(d) Relative MSE of $\hat{\nu}$

Figure 4.7. Bias and MSE of the parameter estimates of the Student's t-vine in 50 dimensions, with low dependence ($\rho = 0.2$), averaged over each level.

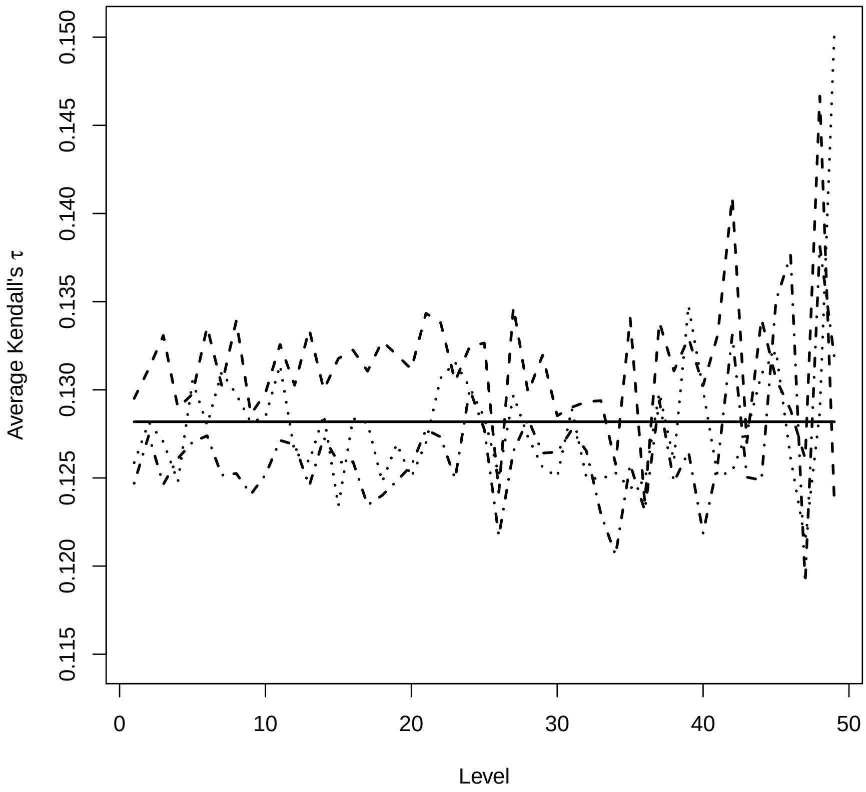


Figure 4.8. Empirical pairwise Kendall's τ coefficients for three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modelled conditional dependence belongs to, and the corresponding true coefficients (connected line).

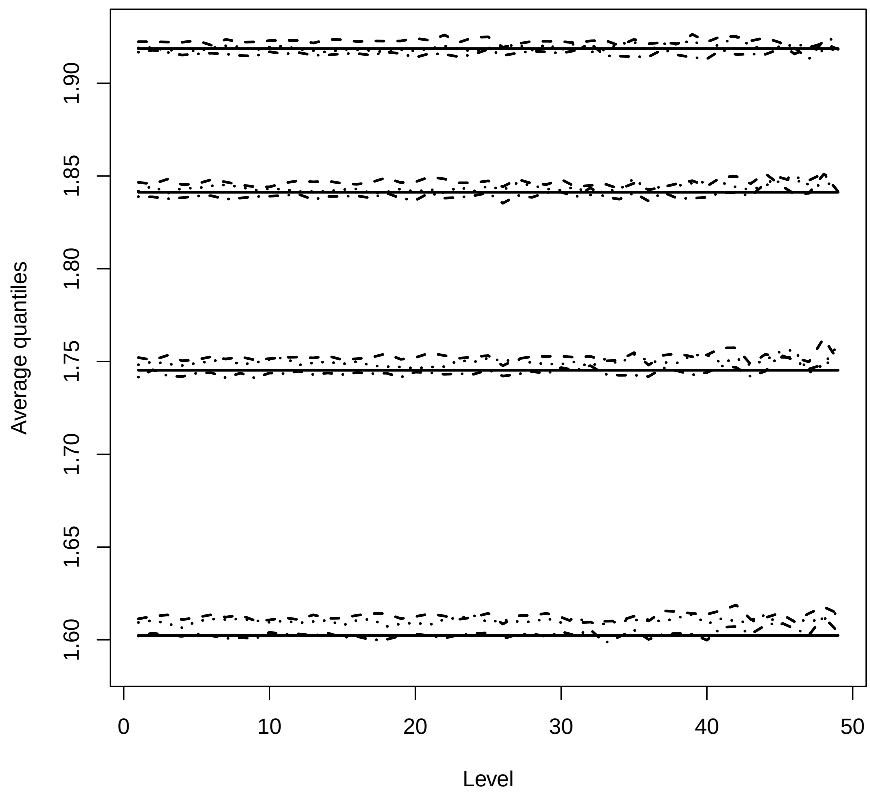


Figure 4.9. Estimated 90%, 95%, 97.5% and 99% quantiles for three samples generated from the estimated vine (dashed and dotted lines), averaged over the level the corresponding modelled conditional dependence belongs to, and the corresponding true quantiles (connected line).

5 Conclusion

There are various estimators for the parameters of a pair-copula construction. The aim of this work has been to study and compare two of the most common ones, namely the semiparametric and stepwise semiparametric estimators. Both of these are based on transforming the original data with their empirical distribution functions. Whereas the SP procedure consists in estimating all parameters simultaneously, the SSP procedure addresses one level at a time, plugging in estimates from preceding levels.

In order to compare the two estimators, we have carried out a simulation study based on D-vines, a subset of PCCs. Except for one example, all the considered models are five-dimensional. Moreover, they are based on at least one out of three copula types, namely the Gaussian, the Student's t and the Clayton copulae. Varying the degree of dependence and the sample size, we have studied the effect on the two estimators.

Generally, the finite sample bias and MSE of the SSP estimator are higher than its competitor's, reflecting its lower asymptotic efficiency. The difference between the two estimators increases, in favour of SP, with the degree of dependence, and also with the level when the dependence is strong. Most likely, this is due to SSP's greater sensitivity towards the repeated transformations of the pseudo observations. Nonetheless, the performance of the SSP estimator is overall rather good compared to SP. Moreover, the former is consistently faster than the latter, especially when the number of parameters is high, as one would expect.

When the sample size decreases, both estimators' variances increase. Based on 500 observations, the parameter estimates for the five-dimensional vines are still rather precise, whereas $n = 50$ is not enough. Also, the difference between the estimators becomes smaller with the sample size. That is also the case when the model is not correctly specified. Neither the SP nor the SSP estimators are particularly robust towards misspecification of the pair-copulae constituting the PCC, but the former appears to suffer more than the latter. Hence, the gain from using the more time consuming SP estimator is reduced for small samples and inaccurate model assumptions.

For high dimensional problems, the SP estimator is simply too expensive, and would probably be numerically unstable, whereas SSP estimation still is doable.

Used on a 50-dimensional Student's t-vine, the resulting estimates are quite good up to level $20-30$. After that, the finite sample bias and MSE explode. Estimation of such high order dependencies is unfortunately numerically highly challenging. However, this does not seem to affect the corresponding lower order dependencies. Despite the erroneous estimates for the top levels, the estimated distribution is in fact rather similar to the true one. This is an incentive to truncate large structures after a certain level, letting the copulae of the top levels be the independence copula.

Overall, the SSP estimator well suited to set start values for the SP one in small to medium sized problems. Furthermore, the extra time spent on SP estimation is not necessarily worthwhile, especially when the dependence is not too strong, the sample size is low, or the model is partly misspecified. When the number of parameters becomes large, SSP estimation is the only alternative in practice. Of course, we have not considered cases of extreme dependence, for which the SSP estimator is likely to be outperformed. Also, we have restricted our attention to D-vines. The relative behaviour of the two estimators for C-vines and other regular vines may be a subject for future work.

References

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

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

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

Berg, D. (2009). Copula goodness-of-fit testing: an overview and power comparison. *European Journal of Finance*, 15:675–701.

Brechmann, E. C., Czado, C., and Aas, K. (2010). Truncated regular vines in high dimensions with application to financial data. Submitted for publication.

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

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

Embrechts, P., McNeil, A. J., and Straumann, D. (1999). Correlation: Pitfalls and alternatives. *Risk*, 12:69–71.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika*, 74:549–555.

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

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

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

Hobæk Haff, I. (2010). Parameter estimation for pair-copula constructions. Technical Report SAMBA/36/10, Norsk Regensentral.

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.

Joe, H. (1996). *Distributions with Fixed Marginals and Related Topics*, chapter Families of m-variate distributions with given margins and m(m-1)/2 dependence parameters. IMS, Hayward, CA.

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

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

Joe, H., Li, H., and Nikoloulopoulos, A., K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101:252–270.

Joe, H. and Xu, J. (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, University of British Columbia, Department of Statistics.

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

Kolbjørnsen, O. and Stien, M. (2008). D-vine creation of non-gaussian random field. In *Procedings of the Eight International Geostatistics Congress*, pages 399–408. GECAMIN Ltd.

Kurowicka, D. and Cooke, R. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley, New York.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society, Series B*, 44:414–422.

Shih, J. and Louis, T. (1995). Inferences on the association parameter in copula models for survival data. *Biometrics*, 51:1384–1399.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris*, 8.

Tsukahara, H. (2005). Semiparametric estimation of copula models. *The Canadian Journal of Statistics*, 33:357–375.