On the simplified pair-copula construction – simply useful or too simplistic?

Ingrid Hobæk Haff[*,a,c], Kjersti Aas[a,c], Arnoldo Frigessi[b,c]

[a] *Norwegian Computing Center, PB 114 Blindern, NO-0373 Oslo, Norway*
[b] *Department of Biostatistics, University of Oslo, PB 1122 Blindern, NO-0317 Oslo, Norway*
[c] *Statistics for Innovation (sfi)[2], Oslo, Norway*

Abstract

Due to their high flexibility, yet simple structure, pair-copula constructions (PCCs) are becoming increasingly popular for constructing continuous multivariate distributions. However, inference requires the simplifying assumption that all the pair-copulae depend on the conditioning variables merely through the two conditional distribution functions that constitute their arguments, and not directly. In terms of standard measures of dependence, we express conditions under which a specific pair-copula decomposition of a multivariate distribution is of this simplified form. Moreover, we show that the simplified PCC in fact is a rather good approximation, even when the simplifying assumption is far from being fulfilled by the actual model.

Key words: copulae, vines, multivariate distributions, hierarchical structures

[*] Corresponding author

Email addresses: ingrid@nr.no (Ingrid Hobæk Haff), Kjersti.Aas@nr.no (Kjersti Aas), arnoldo.frigessi@medisin.uio.no (Arnoldo Frigessi)

1. Introduction

The rapidly increasing availability of multi-dimensional data for complex systems has lead to a renewed interest in multivariate modelling, and copulae in particular. This has resulted in a long and varied list of parametric bivariate copulae, perfectly adequate for bivariate models. However, in higher dimensions, the selection of parametric copulae is still rather limited [Genest et al., 2009].

Recent developments in this area tend toward hierarchical, copula-based structures. The perhaps most promising of these is the pair-copula construction (PCC). Originally proposed by Joe [1996], it has been further explored and discussed by Bedford and Cooke [2001, 2002], Kurowicka and Cooke [2006] and in an inferential context, by Aas et al. [2009]. Lately, a number of publications on PCCs have also appeared in the literature, especially in financial applications. These include Fischer et al. [2007], Chollete and Valdesogo [2008], Heinen and Valdesogo [2008], Schirmacher and Schirmacher [2008] and Czado et al. [2009]. Bayesian inference on PCCs is the topic of Czado and Min [2008], while Joe et al. [2009] explore tail dependence in such constructions. Kolbjørnsen and Stien [2008] present a non-parametric petroleum related application of PCCs.

The growing interest for the PCC is probably due to the combination of their simple structure and high flexibility. While built exclusively from pair-copulae, they can model a wide range of complex dependencies. In fact, the studies of Berg and Aas [2009] and Fischer et al. [2007], comparing PCCs with other multivariate models, e.g. hierarchical Archimedean constructions [Joe, 1997; Savu and Trede, 2006], concluded with the superiority of PCCs.

Nevertheless, the PCC has some shortcomings. A general multivariate model can be decomposed exactly in a hierarchical construction based on pair-copulae with conditional cumulative distribution functions (cdfs) as arguments, as for example

$$C_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3), \tag{1}$$

where $C_{ij|k}$ is the copula corresponding to the conditional cdf $F_{ij|k}$ of X_i and X_j given X_k, and $F_{i|k}$ the cdf of X_i given X_k. For inference to be fast, flexible and robust, however, one must assume that these pair-copulae are independent of the conditioning variables, except through the conditional distributions, i.e.

$$C_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)). \tag{2}$$

Hence, although the general pair-copula decomposition (1) can represent all absolutely continuous multivariate distributions with strictly increasing marginal distributions, realistically, one must resort to the simplified version (2).

In this paper, we explore the limitations of the simplified PCC. In particular, we express conditions under which a multivariate model is of the simplified form, in terms of standard measures of dependence. The simplified PCC is also a good approximation, as we will demonstrate in a simple example.

The rest of this paper is organised as follows. In Section 2, we present the problem precisely. Section 3 provides illustrative examples. Section 4 exhibits properties of the simplified PCC. Approximation with a simplified PCC is the subject of Section 5. Finally, Section 6 contains some concluding remarks.

2. The simplified PCC

Consider three random variables X_1, X_2, X_3 having the joint cdf $F_{123}(x_1, x_2, x_3)$. Assuming that $F_{123}(x_1, x_2, x_3)$ is absolutely continuous with strictly increasing marginal distributions $F_1(x_1)$, $F_2(x_2)$ and $F_3(x_3)$, the corresponding probability density function (pdf) $f_{123}(x_1, x_2, x_3)$ is factorised as

$$f_{123}(x_1, x_2, x_3) = f_3(x_3) f_{2|3}(x_2|x_3) f_{1|23}(x_1|x_2, x_3). \tag{3}$$

To obtain a PCC, one rewrites (3) in terms of copula densities. Let $c_{23}(F_2(x_2), F_3(x_3))$ be the density of the copula $C_{23}(F_2(x_2), F_3(x_3))$, corresponding to the distribution $F_{23}(x_2, x_3)$ of the pair X_2, X_3. The bivariate density $f_{23}(x_2, x_3)$ is then given by (McNeil et al. [2006], pp. 197)

$$f_{23}(x_2, x_3) = c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2) f_3(x_3).$$

Hence, the second factor on the right hand side of (3) is

$$f_{2|3}(x_2|x_3) = c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2).$$

The third factor of (3) can be expressed through $c_{12|3}$, the copula density belonging to $F_{12|3}(x_1, x_2|x_3)$, and c_{13}, as follows. First, we have that

$$c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3) = \frac{f_{12|3}(x_1, x_2|x_3)}{f_{1|3}(x_1|x_3) f_{2|3}(x_2|x_3)}.$$

Thus, we may write

$$\begin{aligned} f_{1|23}(x_1|x_2, x_3) &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3) f_{1|3}(x_1|x_3) \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3) c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1). \end{aligned}$$

Inserting this into (3), we obtain the full PCC expansion

$$\begin{aligned} f_{123}(x_1, x_2, x_3) =& f_1(x_1) f_2(x_2) f_3(x_3) \\ &\cdot c_{13}(F_1(x_1), F_3(x_3)) c_{23}(F_2(x_2), F_3(x_3)) \\ &\cdot c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3). \end{aligned} \tag{4}$$

Note that in general, the copula density $c_{12|3}$ **depends on the conditioning variable** x_3, not only through its arguments $F_{1|3}(x_1|x_3)$ and $F_{2|3}(x_2|x_3)$, but also directly through x_3. Moreover, (4) is one out of three possible decompositions in three dimensions. The number of decompositions grows rapidly with the dimension. There are as many as 240 different PCCs for a five-dimensional density, half of which are so-called regular vines [Bedford and Cooke, 2001, 2002]. We will come back to D-vines, a subset of regular vines of the form

$$f_{1\ldots d}(x_1,\ldots,x_d) = \prod_{k=1}^{d} f_k(x_k) \prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}), \tag{5}$$

where $v_{ij} = \{i+1,\ldots,i+j-1\}$, and correspondingly $\boldsymbol{x}_{v_{ij}} = (x_{i+1},\ldots,x_{i+j-1})$.

The building blocks of a PCC [Aas et al., 2009] are pair-copulae, whose two arguments are conditional distributions [Bedford and Cooke, 2001, 2002], except at the ground level, where there is no conditioning. The number of conditioning variables in these distributions increases with the level in the structure, from 1 to $d-2$, where d is the dimension. For instance, in five dimensions, the top level copula of one of the many possible decompositions has the two arguments $F_{1|234}(x_1|x_2,x_3,x_4)$ and $F_{5|234}(x_5|x_2,x_3,x_4)$.

The reason for leaving the full PCC, where all pair-copulae are allowed to depend directly on the conditioning variables is purely practical. At the second level of the construction, it may still be possible to estimate a copula that depends additionally on the single conditioning variable, using some sort of smoothing technique. However, at higher levels, where the number of conditioning variables increases, this becomes very difficult in a parametric setting, and impossible in a non-parametric one.

Inference with a PCC therefore requires the assumption that the pair-copulae are independent of the conditioning variables, except through the conditional distributions. In the three-dimensional case (4), this amounts to

$$\begin{aligned} f_{123}(x_1,x_2,x_3) = &f_1(x_1)f_2(x_2)f_3(x_3) \\ &\cdot c_{13}(F_1(x_1),F_3(x_3))c_{23}(F_2(x_2),F_3(x_3)) \\ &\cdot c_{12|3}(F_{1|3}(x_1|x_3),F_{2|3}(x_2|x_3)). \end{aligned} \tag{6}$$

Making this assumption, one obtains what we will hereafter denote the **simplified PCC**, as opposed to the general one, given in (4).

3. Examples

What kinds of distributions can the simplified PCC represent? More specifically, which are the necessary characteristics of the joint distribution for the simplified PCC to be correct, and how limiting are these conditions? First, we will illustrate these questions with some examples. For the sake of simplicity and interpretability, all examples, but one, are three-dimensional distributions, but extensions to arbitrary dimensions can be constructed.

Example 3.1. Consider the distribution given by

$$\left[\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \middle| X_3 = x_3\right] \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, x_3\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),$$

where

$$X_3 \sim Gamma^{-1}\left(\frac{\nu}{2}, \frac{2}{\nu}\right),$$

i.e. X_3 has the pdf

$$f_3(x_3) = \frac{\nu^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})x_3^{\frac{\nu+2}{2}}} \exp\left\{-\frac{\nu}{2x_3}\right\}$$

Hence, the unconditional distribution of (X_1, X_2) is the bivariate t-distribution with correlation ρ and ν degrees of freedom (McNeil et al. [2006], pp. 75). The joint density of $X_1,\ X_2$ and X_3 is

$$\begin{aligned} f_{123}(x_1,&x_2,x_3) \\ &= \frac{\nu^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}+1}\pi\Gamma(\frac{\nu}{2})\sqrt{1-\rho^2}x_3^{\frac{\nu+4}{2}}} \exp\left\{-\frac{1}{2x_3}\left(\frac{x_1^2+x_2^2-2\rho x_1 x_2}{1-\rho^2}+\nu\right)\right\}, \\ &x_1, x_2 \in \mathbb{R}, x_3 > 0. \end{aligned}$$

To assess whether decomposition (4) of this multivariate distribution is of the simplified form, one must compute the copula density $c_{12|3}$, linking $F_{1|3}(x_1|x_3)$ and $F_{2|3}(x_2|x_3)$. We know that $[X_i|X_3 = x_3] \sim \mathcal{N}(0, x_3), i = 1, 2$. Thus, the copula density $c_{12|3}$ is given by

$$\begin{aligned} c_{12|3}(F_{1|3}&(x_1|x_3), F_{2|3}(x_2|x_3); x_3) \\ &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{1|3}(x_1|x_3)f_{2|3}(x_2|x_3)} \\ &= \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho}{2(1-\rho^2)}\left(\rho\frac{x_1^2}{x_3}+\rho\frac{x_2^2}{x_3}-2\frac{x_1x_2}{x_3}\right)\right\}. \end{aligned} \tag{7}$$

Defining

$$u_{i|3} = F_{i|3}(x_i|x_3) = \Phi\left(\frac{x_i}{\sqrt{x_3}}\right), i = 1, 2,$$

where Φ is the cumulative standard normal distribution function, we have

$$\frac{x_i}{\sqrt{x_3}} = \Phi^{-1}(u_{i|3}), i = 1, 2.$$

Hence, (7) becomes

$$\begin{aligned} &c_{12|3}(u_{1|3}, u_{2|3}; x_3) \\ &= \frac{1}{\sqrt{1-\rho^2}} \\ &\quad \cdot \exp\left\{-\frac{\rho}{2(1-\rho^2)}\left(\rho\Phi^{-1}(u_{1|3})^2+\rho\Phi^{-1}(u_{2|3})^2-2\Phi^{-1}(u_{1|3})\Phi^{-1}(u_{2|3})\right)\right\}. \end{aligned}$$

This copula density is independent of x_3 (except through $u_{1|3}$ and $u_{2|3}$), and therefore of the simplified form. We recognise it as the density of a Gaussian copula with correlation ρ. Note that copulae are invariant to location and scale. That is why the conditioning variable does not affect the linking copula when it only enters the scale of the conditional distribution. It may be shown that the two other decompositions are of the simplified form as well (see Appendix A.1).

Example 3.2. Now consider the distribution given by

$$\left[\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \middle| X_3 = x_3 \right] \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & x_3 \\ x_3 & 1 \end{pmatrix} \right),$$

where

$$X_3 \sim Beta(\alpha, \beta),$$

i.e. X_3 has the pdf

$$f_3(x_3) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x_3^{\alpha-1}(1-x_3)^{\beta-1}.$$

The joint density of $X_1,\ X_2$ and X_3 is

$$\begin{aligned} f_{123}(x_1, x_2, x_3) = & \frac{\Gamma(\alpha+\beta)}{2\pi\Gamma(\alpha)\Gamma(\beta)} \frac{x_3^{\alpha-1}(1-x_3)^{\beta-1}}{\sqrt{1-x_3^2}} \\ & \cdot \exp\left\{ -\frac{1}{2(1-x_3^2)} \left(x_1^2 + x_2^2 - 2x_1x_2x_3 \right) \right\}, \\ & x_1, x_2 \in \mathbb{R}, 0 \le x_3 \le 1. \end{aligned}$$

This example is very similar to Example 3.1, except that X_3 is now the correlation in the conditional distribution of the other two variables, instead of the variance. In this case, $c_{12|3}$ is the density of the Gaussian copula with association parameter x_3. That density is obviously not of the simplified form, since the conditioning variable is a copula parameter.

Example 3.3. Consider the distribution

$$\left[\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \middle| X_3 = x_3 \right] \sim t_{x_3}^2 \left(\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right),$$

where

$$X_3 \sim Pareto(\theta, 1),$$

i.e. X_3 has the pdf

$$f_3(x_3) = \frac{\theta}{x_3^{\theta+1}},\ x_3 > 1,$$

which is the Pareto distribution of the first kind. Moreover, $t_\nu^2(\mathbf{R})$ is the bivariate t-distribution with ν degrees of freedom and correlation matrix $\mathbf{R}$. The joint density of $X_1,\ X_2$ and X_3 is

$$f_{123}(x_1,x_2,x_3) = \frac{\theta}{\pi\sqrt{1-\rho^2}} \frac{\Gamma\left(\frac{x_3+2}{2}\right)}{x_3^{\theta+1}\Gamma\left(\frac{x_3}{2}\right)} \left(1+\frac{x_1^2+x_2^2-2\rho x_1 x_2}{(1-\rho^2)x_3}\right)^{-\frac{x_3+2}{2}}, \quad x_1,x_2\in\mathbb{R}, x_3>1.$$

The copula $C_{12|3}$ is, in this case, a t-copula with correlation ρ and x_3 degrees of freedom. As in Example 3.2, the conditioning variable x_3 is one of the copula parameters. Hence, the simplifying assumption is invalid.

Example 3.4. Five-dimensional example. Let the five variables $X_1,\dots,X_5$ have a multivariate Burr distribution with identical marginals (Kotz et al. [2000], pp. 609). Their joint density is

$$f_{12345}(x_1,\dots,x_5) = \frac{\alpha^5\beta^5\prod_{i=1}^5(\theta+i-1)x_i^{\beta-1}}{(\alpha\sum_{i=1}^5 x_i^\beta+1)^{\theta+5}},\ x_i>0,\ i=1,\dots,5.$$

This distribution, which is also called the multivariate Pareto distribution of the fourth kind, was first discussed by Takahasi [1965]. It is a generalisation of the multivariate Pareto distribution of the first kind, introduced by Mardia [1962, 1964].

One possible decomposition is the D-vine

$$\begin{aligned}
&f_{12345}(x_1,\dots,x_5)\\
&= f_1(x_1)f_2(x_2)f_3(x_3)f_4(x_4)f_5(x_5)\\
&\cdot c_{12}(F_1(x_1),F_2(x_2))c_{23}(F_2(x_2),F_3(x_3))c_{34}(F_3(x_3),F_4(x_4))c_{45}(F_4(x_4),F_5(x_5))\\
&\cdot c_{13|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2);x_2)\\
&\cdot c_{24|3}(F_{2|3}(x_2|x_3),F_{4|3}(x_4|x_3);x_3)\\
&\cdot c_{35|4}(F_{3|4}(x_3|x_4),F_{5|4}(x_5|x_4);x_4)\\
&\cdot c_{14|23}(F_{1|23}(x_1|x_2,x_3),F_{4|23}(x_4|x_2,x_3);x_2,x_3)\\
&\cdot c_{25|34}(F_{2|34}(x_2|x_3,x_4),F_{5|34}(x_5|x_3,x_4);x_3,x_4)\\
&\cdot c_{15|234}(F_{1|234}(x_1|x_2,x_3,x_4),F_{5|234}(x_5|x_2,x_3,x_4);x_2,x_3,x_4).
\end{aligned} \tag{8}$$

The above construction is of the simplified form if the copula densities on the last six lines of (8) are functions of the conditioning variables merely through their arguments. This is the case. More specifically, the densities are given by (derived in Appendix A.2)

$$\begin{aligned}
&c_{i,i+j|i+1,\dots,i+j-1}(u_{i|i+1,\dots,i+j-1},u_{i+j|i+1,\dots,i+j-1};x_{i+1},\dots,x_{i+j-1})\\
&= \frac{\theta+j}{\theta+j-1}(1-u_{i|i+1,\dots,i+j-1})^{-\frac{\theta+j}{\theta+j-1}}(1-u_{i+j|i+1,\dots,i+j-1})^{-\frac{\theta+j}{\theta+j-1}}\\
&\quad\cdot((1-u_{i|i+1,\dots,i+j-1})^{-\frac{1}{\theta+j-1}}+(1-u_{i+j|i+1,\dots,i+j-1})^{-\frac{1}{\theta+j-1}}-1)^{-(\theta+j+1)},
\end{aligned}$$

where $u_{k|i+1,\ldots,i+j-1} = F_{k|i+1,\ldots,i+j-1}(x_k|x_{i+1},\ldots,x_{i+j-1}),\ k = i, i+j$, which is the density of the Clayton survival copula with parameter $\frac{1}{\theta+j-1}$, and clearly of the simplified form. In fact, according to Cook and Johnson [1981], the copula corresponding to the joint distribution of $X_1,\ldots,X_5$ is a five-dimensional Clayton survival copula. Thus, the multivariate Clayton survival copula can be represented by a simplified PCC.

There are $\frac{5!}{2} = 60$ other D-vine decompositions of the joint pdf [Aas et al., 2009]. All these are equivalent, since the distribution has permutable variables, and are therefore simplified PCCs.

4. Properties of the simplified PCC

We have illustrated how some distributions have one or more decompositions of the simplified form, while others do not. Under which conditions can a distribution be represented by a simplified PCC?

The most commonly used measure of dependence is the linear (or Pearson's) correlation coefficient. Although it may be useful and interpretable for elliptical distributions (as long as it exists), it is not a measure of concordance [Embrechts et al., 2003]. Therefore, in general, we do not expect the conditional linear correlation to be appropriate for describing the conditions under which the simplifying assumption is valid.

Measures of concordance, such as Kendall's tau, Spearman's rho and the coefficients of tail dependence, provide a more natural description of dependence in copula models. As we shall see next, they form the basis for some results concerning the validity of the simplifying assumption for a given decomposition of the joint distribution.

Proposition 1. *Let $X_1,\ldots,X_d$ be random variables having the joint probability density $f_{1\ldots d}(x_1,\ldots,x_d)$. Without loss of generality, consider decomposition* (5)*. If, for any linking copula $C_{i,i+j|v_{ij}}$, the corresponding Kendall's tau $\tau(X_i, X_j|\boldsymbol{x}_{v_{ij}})$ is a function of the conditioning variables $\boldsymbol{x}_{v_{ij}}$, the decomposition is not of the simplified form.*

Proof. The result follows immediately from the expression for Kendall's tau in terms of the copula function $C_{i,i+j|v_{ij}}$ [Nelsen, 1999];

$$\tau(X_i, X_j|\boldsymbol{x}_{v_{ij}}) = \\ 4\int_0^1\int_0^1 C_{i,i+j|v_{ij}}(u_{i|v_{ij}}, u_{i+j|v_{ij}}; \boldsymbol{x}_{v_{ij}})dC_{i,i+j|v_{ij}}(u_{i|v_{ij}}, u_{i+j|v_{ij}}; \boldsymbol{x}_{v_{ij}}) - 1,$$

where $u_{k|v_{ij}} = F_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}}), k = i, j$. Obviously, $\tau(X_i, X_j|\boldsymbol{x}_{v_{ij}})$ cannot be a function of the conditioning variables unless the copula function is. □

Remark 1. (i) Proposition 1 may instead be formulated in terms of Spearman's rho, Blomqvist's beta, or any measure of monotone association.

(ii) The converse statement is not true. Even if none of the Kendall's tau coefficients corresponding to linking copulae is a function of the conditioning variables, the simplifying assumption may be invalid for the decomposition in question. To illustrate this, we return to Example 3.3. Kendall's tau for the t-distributed pair X_1, X_2, conditioned on $X_3 = x_3$, is given by

$$\tau(X_1, X_2|X_3 = x_3) = \frac{2}{\pi}\arcsin(\rho),$$

which is not a function of the conditioning variable X_3. Despite this, the decomposition is not of the simplified form.

(iii) For simplicity, the chosen decomposition (5) is a D-vine. However, the result is valid for any PCC.

Proposition 2. *Let $X_1, \dots, X_d$ be random variables having the joint probability density $f_{1\dots d}(x_1, \dots, x_d)$. Without loss of generality, consider decomposition* (5)*. If, for any linking copula $C_{i,i+j|v_{ij}}$, the corresponding upper tail dependence coefficient $\lambda_U(X_i, X_j|\boldsymbol{x}_{v_{ij}})$ is a function of the conditioning variables $\boldsymbol{x}_{v_{ij}}$, the decomposition is not of the simplified form.*

Proof. The result follows from the expression for the upper tail dependence coefficient in terms of the copula function $C_{i,i+j|v_{ij}}$ [Embrechts et al., 2003]:

$$\lambda_U(X_i, X_j|\boldsymbol{x}_{v_{ij}}) = \lim_{u\nearrow 1}\frac{1-2u+C_{i,i+j|v_{ij}}(u,u;\boldsymbol{x}_{v_{ij}})}{1-u},$$

which is not a function of the conditioning variables unless the copula function is. □

Remark 2. (i) Proposition 2 may be written in terms of the lower tail dependence coefficient $\lambda_L(X_i, X_j|\boldsymbol{x}_{v_{ij}}) = \lim_{u\searrow 0}\frac{C_{i,i+j|v_{ij}}(u,u;\boldsymbol{x}_{v_{ij}})}{u}$.

(ii) The converse statement is not true. Consider Example 3.2. $C_{12|3}$ is a Gaussian copula with association parameter x_3, for which the simplifying assumption obviously is not valid. However,

$$\lambda_L(X_1, X_2|X_3 = x_3) = \lambda_U(X_1, X_2|X_3 = x_3) = 0$$

are not functions of the conditioning variable X_3.

The results presented so far provide conditions for a decomposition **not** to be of the simplified form. For a particular family of distributions, we can be more positive.

Proposition 3. *Let $X_1, \dots, X_d$ be random variables having the joint probability density $f_{1\dots d}(x_1, \dots, x_d)$. Consider decomposition* (5)*. Assume that all conditional distributions entering the decomposition have the form*

$$F_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}}) = g_{k|v_{ij}}\left(\frac{x_k - a_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}\right),\ k = i, i+j, \tag{9}$$

where $a_{k|v_{ij}}$, $b_{k|v_{ij}}$ are some continuous functions on $\mathbb{R}$, with $b_{k|v_{ij}} > 0$, and $g_{k|v_{ij}}$ is a cumulative distribution function on $\mathbb{R}$, hence that $F_{k|v_{ij}}$ belongs to a location-scale family with location and scale parameters that are functions of the conditioning variables $\boldsymbol{x}_{v_{ij}}$. Then, the decomposition is of the simplified form if and only if the bivariate conditional probability densities corresponding to the pair-copulae are of the form

$$\begin{aligned}
&f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}}) \\
&= \frac{1}{b_{i|v_{ij}}(\boldsymbol{x}_{v_{ij}}) b_{i+j|v_{ij}}(\boldsymbol{x}_{v_{ij}})}\, h_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})), \quad (10)
\end{aligned}$$

where $h_{k|v_{ij}}$ is some continuous, non-negative function on $[0,1]^2$.

Proof. Consider the linking copula $C_{i,i+j|v_{ij}}$ from the decomposition. Its density is given by

$$\begin{aligned}
&c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}) \\
&\qquad = \frac{f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}})}{f_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}) f_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})}. \quad (11)
\end{aligned}$$

According to (9), the univariate densities in the denominator are given by

$$\begin{aligned}
f_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}}) &= \frac{d}{dx_k} F_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}}) \\
&= \frac{d}{dx_k} g_{k|v_{ij}}\left(\frac{x_k - a_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}\right) \\
&= \frac{d}{dz} g_{k|v_{ij}}\Bigg|_{z=\frac{x_k - a_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}} \cdot \frac{1}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}, \; k = i, i+j.
\end{aligned}$$

Since $g_{k|v_{ij}}$ is strictly increasing, we have

$$\frac{x_k - a_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})} = g^{-1}_{k|v_{ij}}\left(F_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}})\right).$$

Now, define $\tilde{g}_{k|v_{ij}} \equiv g'_{k|v_{ij}} \circ g^{-1}_{k|v_{ij}}$, where $g'_{k|v_{ij}}(z) = \frac{d}{dz} g_{k|v_{ij}}(z)$. Then,

$$f_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}}) = \frac{\tilde{g}_{k|v_{ij}}\left(F_{k|v_{ij}}(x_k|\boldsymbol{x}_{v_{ij}})\right)}{b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}})}, \; k = i, i+j.$$

Inserting this into (11), we obtain

$$\begin{aligned}
&c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}) \\
&= \frac{b_{i|v_{ij}}(\boldsymbol{x}_{v_{ij}}) b_{i+j|v_{ij}}(\boldsymbol{x}_{v_{ij}})}{\tilde{g}_{i|v_{ij}}\left(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}})\right) \tilde{g}_{i+j|v_{ij}}\left(F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})\right)} \cdot f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}}).
\end{aligned}$$

If $f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}})$ is of the form (10), we have

$$
\begin{aligned}
c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}),&F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}) \\
&= \frac{h_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}))}{\tilde{g}_{i|v_{ij}}\left(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}})\right)\tilde{g}_{i+j|v_{ij}}\left(F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})\right)},
\end{aligned}
$$

which clearly satisfies the simplifying assumption.

Conversely, if $c_{i,i+j|v_{ij}}$ satisfies the simplifying assumption, i.e.

$$
\begin{aligned}
c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}),&F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}) \\
&= c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})),
\end{aligned}
$$

we have

$$
\begin{aligned}
&f_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}}) \\
&\qquad = c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}); \boldsymbol{x}_{v_{ij}}) \\
&\qquad\quad \cdot f_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}) f_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}) \\
&\qquad = \frac{1}{b_{i|v_{ij}}(\boldsymbol{x}_{v_{ij}}) b_{i+j|v_{ij}}(\boldsymbol{x}_{v_{ij}})} c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})) \\
&\qquad\quad \cdot \tilde{g}_{i|v_{ij}}\left(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}})\right)\tilde{g}_{i+j|v_{ij}}\left(F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}})\right),
\end{aligned}
$$

which is of the form (10). □

Example 4.1. Elliptical distributions. A consequence of Proposition 3 is that elliptical distributions can be represented by a PCC of the simplified form, as long as their scale matrix is positive definite.

Consider the random variables $X_1, \dots, X_d$ from a multivariate elliptical distribution with location vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and characteristic generator ϕ, i.e. $\boldsymbol{X} = (X_1, \dots, X_d)^T \sim E_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$. When $\boldsymbol{\Sigma}$ is positive definite, the joint pdf is defined, and of the form [Cambanis et al., 1981]

$$
f_{1\dots d}(x_1, \dots, x_d) = \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} g_{1\dots d}\left((\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right),
$$

for some positive function $g_{1\dots d}$ on $\mathbb{R}$. Moreover, all marginal distributions are elliptical distributions with the same characteristic generator. Hence, $\boldsymbol{X}_{ij} = (X_i, \boldsymbol{X}_{v_{ij}}^T, X_{i+j})^T \sim E_{j+1}(\boldsymbol{\mu}_{ij}, \boldsymbol{\Sigma}_{ij}, \phi)$, with

$$
\boldsymbol{\mu}_{ij} = \begin{pmatrix} \mu_i \\ \boldsymbol{\mu}_{v_{ij}} \\ \mu_{i+j} \end{pmatrix}
$$

$$
\boldsymbol{\Sigma}_{ij} = \begin{pmatrix} \Sigma_{ii} & \boldsymbol{\Sigma}_{i,v_{ij}}^T & \Sigma_{i,i+j} \\ \boldsymbol{\Sigma}_{i,v_{ij}} & \boldsymbol{\Sigma}_{v_{ij},v_{ij}} & \boldsymbol{\Sigma}_{i+j,v_{ij}}^T \\ \Sigma_{i,i+j} & \boldsymbol{\Sigma}_{i+j,v_{ij}} & \Sigma_{i+j,i+j} \end{pmatrix},
$$

where v_{ij} is as defined in (5).

In order to use Proposition 3, we need the pdf of the bivariate conditional distribution of $\left[(X_i, X_{i+j})^T | \boldsymbol{X}_{v_{ij}} = \boldsymbol{x}_{v_{ij}}\right]$, as well as its marginals. As shown by Cambanis et al. [1981], this is a bivariate elliptical distribution $E_2(\boldsymbol{\mu}_{ij|v_{ij}}, \boldsymbol{\Sigma}_{ij|v_{ij}}, \tilde{\phi})$, with

$$\boldsymbol{\mu}_{ij|v_{ij}} = \begin{pmatrix} \mu_i + \boldsymbol{\Sigma}_{i,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} (\boldsymbol{x}_{v_{ij}} - \boldsymbol{\mu}_{v_{ij}}) \\ \mu_{i+j} + \boldsymbol{\Sigma}_{i+j,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} (\boldsymbol{x}_{v_{ij}} - \boldsymbol{\mu}_{v_{ij}}) \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{ij|v_{ij}} = \begin{pmatrix} \Sigma_{ii} - \boldsymbol{\Sigma}_{i,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{i,v_{ij}} & \Sigma_{i,i+j} - \boldsymbol{\Sigma}_{i,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{i+j,v_{ij}} \\ \Sigma_{i,i+j} - \boldsymbol{\Sigma}_{i,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{i+j,v_{ij}} & \Sigma_{i+j,i+j} - \boldsymbol{\Sigma}_{i+j,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{i+j,v_{ij}} \end{pmatrix}.$$

The characteristic generator $\tilde{\phi}$ is different from the the original ϕ, except for the multivariate normal distribution. For instance, in the multivariate t-distribution, the number of degrees of freedom changes from ν to $\nu + j - 1$. The conditional marginals are

$$\begin{aligned} &[X_k | \boldsymbol{X}_{v_{ij}} = \boldsymbol{x}_{v_{ij}}] \\ &\sim E_1\left(\mu_k + \boldsymbol{\Sigma}_{k,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} (\boldsymbol{x}_{v_{ij}} - \boldsymbol{\mu}_{v_{ij}}), \Sigma_{kk} - \boldsymbol{\Sigma}_{k,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{k,v_{ij}}\right), \ k = i, i+1, \end{aligned}$$

which is of the form (9), with

$$\begin{aligned} a_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}}) &= \mu_k + \boldsymbol{\Sigma}_{k,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} (\boldsymbol{x}_{v_{ij}} - \boldsymbol{\mu}_{v_{ij}}) \\ b_{k|v_{ij}}(\boldsymbol{x}_{v_{ij}}) &= \sqrt{\Sigma_{kk} - \boldsymbol{\Sigma}_{k,v_{ij}}^T \boldsymbol{\Sigma}_{v_{ij},v_{ij}}^{-1} \boldsymbol{\Sigma}_{k,v_{ij}}} = b_{k|v_{ij}}. \end{aligned}$$

Finally, we know that the pdf of bivariate conditional distribution is given by

$$\begin{aligned} &f_{i,i+j|v_{ij}}(x_i, x_{i+i} | \boldsymbol{x}_{v_{ij}}) \\ &= \frac{1}{|\boldsymbol{\Sigma}_{ij|v_{ij}}|^{\frac{1}{2}}} g_{i,i+j|v_{ij}} \left(\left((x_i, x_{i+j})^T - \boldsymbol{\mu}_{ij|v_{ij}} \right)^T \boldsymbol{\Sigma}_{ij|v_{ij}}^{-1} \left((x_i, x_{i+j})^T - \boldsymbol{\mu}_{ij|v_{ij}} \right) \right) \\ &= \frac{1}{|\boldsymbol{\Sigma}_{ij|v_{ij}}|^{\frac{1}{2}}} g_{i,i+j|v_{ij}} \left(\left(\frac{(x_i - a_{i|v_{ij}}(\boldsymbol{x}_{v_{ij}}))}{b_{i|v_{ij}}} \right)^2 + \left(\frac{(x_{i+j} - a_{i+j|v_{ij}}(\boldsymbol{x}_{v_{ij}}))}{b_{i+j|v_{ij}}} \right)^2 \right. \\ &\quad \left. + \left(\frac{(x_i - a_{i|v_{ij}}(\boldsymbol{x}_{v_{ij}}))}{b_{i|v_{ij}}} \right) \left(\frac{(x_{i+j} - a_{i+j|v_{ij}}(\boldsymbol{x}_{v_{ij}}))}{b_{i+j|v_{ij}}} \right)^2 \right), \end{aligned}$$

which is of the form (10). Hence, by Proposition 3, the distribution can be expressed as a simplified PCC.

Remark 3. (i) All our examples, except Example 3.3, are of the form (9).

(ii) It follows directly from Proposition 3 that if two variables X_i, X_{i+j}, linked by the copula $C_{i,i+j|v_{ij}}$, are marginally independent of the conditioning variables, i.e. $F_{k|v_{ij}}(x_k | \boldsymbol{x}_{v_{ij}}) = F_k(x_k)$, $k = i, i+j$, they must also be jointly independent of the conditioning variables, such that $F_{i,i+j|v_{ij}}(x_i, x_{i+j}|\boldsymbol{x}_{v_{ij}}) = F_{i,i+j}(x_i, x_{i+j})$. In Example 3.2, this is not fulfilled. The two variables X_1, X_2 are marginally independent of X_3. However, their conditional correlation, given X_3, is X_3.

Remark 4. Modelling global behaviour of many variables through their local interactions is one of the important points on the agenda of modern stochastic science. The simplified PCC goes in this direction, since it requires the modelling of pair-copulae only to describe complex multi-component interactions. Among the several theories, Gibbs fields play an important role. According to the Hammersley-Clifford theorem [Hammersley and Clifford, 1971; Besag, 1974], the density of a continuous d-variate distribution may, under some regularity conditions, be written as a Gibbs distribution, i.e.

$$f_{1\dots d}(x_1, \dots, x_d) = \frac{1}{K} \exp\left\{ -\sum_{q \in Q} V_q(\mathbf{x}_q) \right\},$$

where K is a normalising constant, Q is the set of all cliques, as defined from the graph theory, involving the d variables, and V_q are some real-valued potential functions, depending only on the variables in clique q. In practice, inference on both parametric and non-parametric potential functions becomes very complex, sometimes even unmanageable, for interactions between more than two variables. Applications of Gibbs models are therefore mostly limited to bivariate interactions

$$f_{1\dots d}(x_1, \dots, x_d) = \frac{1}{K} \exp\left\{ -\sum_{i=1}^{d} V_i(x_i) - \sum_{i<j} V_{ij}(x_i, x_j) \right\}.$$

Rewriting the simplified form of the D-vine decomposition (5), we obtain

$$\begin{aligned} f_{1\dots d}&(x_1, \dots, x_d) = \\ &\exp\left\{ -\sum_{k=1}^{d} \left(-\log\left(f_k(x_k)\right)\right) \right. \\ &\qquad \left. -\sum_{j=1}^{d-1} \sum_{i=1}^{n-j} \left(-\log\left(c_{i,i+j|v_{ij}}(F_{i|v_{ij}}(x_i|\boldsymbol{x}_{v_{ij}}), F_{i+j|v_{ij}}(x_{i+j}|\boldsymbol{x}_{v_{ij}}))\right)\right) \right\}, \end{aligned}$$

which is a Gibbs model. Although it is composed solely of singletons (the marginal distributions) and pair-wise potentials in two transformed variables, it can represent models involving interactions between more than two variables. In fact all the examples presented in Section 2 involve triple interactions. Hence, while possessing the same simplicity of construction as a pair-wise interaction model, the simplified PCC can capture more complex dependencies.

5. Approximating with the simplified PCC

As we have seen, it is not possible to represent all multivariate distributions by a pair-copula decomposition of the simplified form. Any distribution might however be *approximated* by a simplified PCC. Next, we will study the quality of such an approximation in a simple case, more specifically Example 3.2, which illustrates the situation well.

Recall that X_1, X_2, conditioned on $X_3 = x_3$, are bivariate normal with correlation x_3, while X_3 is beta distributed with parameters (α, β). We wish to approximate the general decomposition (4) with the simplified one (6). As X_1 and X_2 are marginally independent of X_3, $c_{13} = c_{23} = 1$. Hence,

$$\begin{aligned} f_{123}(x_1, x_2, x_3) =& f_1(x_1) f_2(x_2) f_3(x_3) \\ & \cdot c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3). \end{aligned}$$

The copula $C_{12|3}$ is a Gaussian copula with parameter x_3. The approximation $\hat{f}_{123}$ of f_{123} is obtained simply by replacing this copula with one that is independent of x_3. More specifically, we let

$$\begin{aligned} \hat{f}_{123}(x_1, x_2, x_3) =& f_1(x_1) f_2(x_2) f_3(x_3) \\ & \cdot \hat{c}_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)), \end{aligned} \tag{12}$$

where $\hat{c}_{12|3}$ is the density of a Gaussian copula with a constant parameter ρ. To measure the quality of this approximation, we compute its Kullback-Leibler divergence from the true distribution (as did Nikoloulopoulos and Karlis [2008] in their copula model comparison):

$$\begin{aligned} & D_{KL}(f_{123}, \hat{f}_{123}) \\ &= \int_0^1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{123}(x_1, x_2, x_3) \log \frac{f_{123}(x_1, x_2, x_3)}{\hat{f}_{123}(x_1, x_2, x_3)} dx_1 dx_2 dx_3 \\ &= \int_0^1 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{123}(x_1, x_2, x_3) \log \frac{c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3); x_3)}{\hat{c}_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))} dx_1 dx_2 dx_3 \\ &= \log(1 - \rho^2) + \frac{\rho^2}{1 - \rho^2} - \frac{\rho}{1 - \rho^2} \frac{\alpha}{\alpha + \beta} - \mathrm{E}(\log(1 - X_3^2)). \end{aligned} \tag{13}$$

The value of ρ that minimises the Kullback-Leibler divergence (13), that the maximum likelihood estimator converges to, the so-called "least false" parameter value, is $\rho = \alpha/(\alpha+\beta)$. Let ρ take this value, which is also the expected value of X_3 and the unconditional correlation between X_1 and X_2. The expression (13) then reduces to

$$D_{KL}(f_{123}, \hat{f}_{123}) = \log\left(1 - \left(\frac{\alpha}{\alpha + \beta}\right)^2\right) - \mathrm{E}(\log(1 - X_3^2)).$$

It remains to compute the expectation $\mathrm{E}(\log(1 - X_3^2))$. We start by replacing $\log(1-x_3^2)$ with its Taylor expansion $-\sum_{n=1}^{\infty} x_3^{2n}/n$. Moreover, the nth moment of X_3 is given by $\mathrm{E}(X_3^n) = \prod_{i=0}^{n-1}(\alpha+i)/(\alpha+\beta+i)$. The resulting expression for the Kullback-Leibler divergence is

$$D_{KL}(f_{123}, \hat{f}_{123}) = \log\left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^2\right) + \sum_{n=1}^{\infty} \frac{1}{n} \prod_{i=0}^{2n-1} \frac{\alpha+i}{\alpha+\beta+i}. \tag{14}$$

We expect the approximation (12) to get worse when the standard deviation of X_3 increases. Therefore, we have computed the Kullback-Leibler divergence as a function of $\mathrm{Sd}(X_3) = \sqrt{\mathrm{Var}(X_3)} = \sqrt{(\alpha\beta)/((\alpha+\beta)^2(\alpha+\beta+1))}$, keeping $\mathrm{E}(X_3)$ fixed, for a set of expected values. Varying the standard deviation, the distribution of X_3 ranges from the uniform distribution $U[0,1]$ to a rather peaked beta distribution. We are only considering values of α and β for which the beta distribution is either uniform or unimodal. The resulting Kullback-Leibler divergences are displayed in Figure 1. As expected, they increase with the standard deviation of X_3. Moreover, they grow with the expected value $\mathrm{E}(X_3)$. To better illustrate this, we have also plotted the Kullback-Leibler divergence as a function of $\mathrm{E}(X_3)$ for a set of standard deviations, in Figure 2. We would expect the approximation to be best when the dependence X_1 and X_2 is not too strong. As the correlation between X_1 and X_2 is $\mathrm{E}(X_3)$, it is therefore not surprising that the Kullback-Leibler divergence (14) increases with $\mathrm{E}(X_3)$. Furthermore, note that (14) may be written as

$$D_{KL}(f_{123}, \hat{f}_{123}) = \mathrm{E}(-\log(1 - X_3^2)) - \left(-\log\left(1 - \left(\frac{\alpha}{\alpha+\beta}\right)^2\right)\right).$$

Since $-\log(1-x^2)$ is a convex function, Jensen's inequality ensures not only that this difference is always non-negative, but also that it is an increasing function of the expected value $\mathrm{E}(X_3)$.

The Kullback-Leibler divergence enables the comparison between approximations resulting from diverse parameter sets. However, it is not a interpretable measure of the absolute quality of the approximation. Moreover, the Kullback-Leibler divergence describes a weighted average fit, down-weighting the tails due to the log-transformation. For many applications involving copulae, the main focus is actually the tails of the distribution. In such cases, the Kullback-Leibler divergence is not the most appropriate measure. Therefore, we have also computed a quantile of the sum $Y = X_1 + X_2$. In an application, this could typically be the Value-at-Risk of a portfolio of financial assets (with equal weights). For the true distribution, the $\xi \cdot 100\%$ quantile y_ξ is given by

$$\begin{aligned}
\xi &= \mathrm{P}(Y \le y_\xi) \\
&= \int_{-\infty}^{y_\xi} f_Y(y)dy \\
&= \int_{-\infty}^{y_\xi} \int_0^1 f_{Y|3}(y|x_3) f_3(x_3) dx_3 dy \\
&= \int_0^1 f_3(x_3) \int_{-\infty}^{y_\xi} f_{Y|3}(y|x_3) dy dx_3,
\end{aligned}$$

where the last equality follows from Fubini's theorem. We know that $[Y|X_3 = x_3] \sim \mathcal{N}(0, 2+2x_3)$. Using the variable substitution $z = y/\sqrt{2+2x_3}$, we obtain

$$
\begin{aligned}
\mathrm{P}(Y \leq y_\xi) &= \int_0^1 f_3(x_3) \int_{-\infty}^{\frac{y_\xi}{\sqrt{2+2x_3}}} \phi(z) dz dx_3 \\
&= \int_0^1 f_3(x_3) \Phi\left(\frac{y_\xi}{\sqrt{2+2x_3}}\right) dx_3 \\
& \mathrm{E}\left(\Phi\left(\frac{y_\xi}{\sqrt{2+2X_3}}\right)\right),
\end{aligned}
$$

where ϕ is the standard normal probability density. Thus, the quantile is the solution to the equation

$$
\xi = \mathrm{E}\left(\Phi\left(\frac{y_\xi}{\sqrt{2+2X_3}}\right)\right). \tag{15}
$$

In the approximated model, i.e. the simplified PCC, the distribution of the sum $X_1 + X_2$ is simply the normal distribution with mean 0 and standard deviation $\sqrt{2+2\rho}$. The corresponding $\xi \cdot 100\%$ quantile is given by

$$
\hat{y}_\xi = \sqrt{2+2\rho}\ \Phi^{-1}(\xi).
$$

Figure 3 shows the relative difference between the 95% quantile from the true and from the approximated model, i.e. $(y_\xi - \hat{y}_\xi)/y_\xi$, as a function of the standard deviation $\mathrm{Sd}(X_3)$. The various curves correspond to different expected values $\mathrm{E}(X_3)$. They all lie entirely under 0, which means that in this example the simplified PCC consequently overestimates the quantile, thus being conservative in a risk management sense. Moreover, the relative error increases with the standard deviation, as expected. Most importantly, the error is only one per thousand for the worst approximation (corresponding to a uniformly distributed X_3). Hence, in this case, the simplified PCC is a rather good approximation.

6. Concluding remarks

In their general form, PCCs can represent most continuous multivariate distributions. However, for all practical purposes, one must resort to simplified PCCs, made of pair-copulae that depend on the conditioning variables merely through their arguments.

The simple structure, composed solely of pair-copulae, resembles Gibbs fields with bivariate interactions. Nevertheless, simplified PCCs can represent interactions between more than two variables.

Conditions for a specific decomposition of a multivariate model **not** to be of the simplified form can be expressed in terms of standard measures of dependence, more specifically Kendall's tau, Spearman's rho and the coefficients of tail dependence. If all conditional distributions entering the PCC belong to a location-scale family with location and scale parameters that are functions of the conditioning variables, one can also formulate conditions for the converse. Among others, one may show that elliptical distributions can be represented by simplified PCCs as long as their scale matrix is positive definite. However, further research is necessary to fully understand what the simplifying assumption signifies for the dependency.

Not all multivariate distributions can be represented by a simplified PCC. However, one can always use it as an approximation. We have shown that it may in fact be a good one, even when the simplifying assumption is far from being fulfilled. In the example we presented, the pair-copulae constituting the approximated, simplified PCC were of the same type as the building blocks of the exact, general PCC. This need not be the case. One should simply choose the best-fitting pair-copulae. In many cases, it is also probable that the approximation is better for some of the possible decompositions than for others. This is a matter we have not addressed in this paper, and may be a subject for future work.

7. Acknowledgements

This work is funded by Statistics for Innovation, (sfi)^2. We thank Odd Kolbjørnsen for pointing the problem that originated this work, and Christian Genest for helpful discussions. We would also like to thank the participants of the Second Vine Copula Workshop in Delft, December 2008, for important discussions on the theme of vines. Finally, we thank the referees for helping us improve the paper with their good comments and suggestions.

References

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

Abramowitz, M., Stegun, I. A., 1972. Handbook of Mathematical Function. Dover, New York.

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

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

Berg, D., Aas, K., 2009. Models for construction of multivariate dependence. European Journal of Finance. Accepted for publication in 2009.

Besag, J., 1974. Spatial interaction and statistical analysis of lattice systems. Journal of the Royal Statistical Society, Series B 36, 192–236.

Cambanis, S., Huang, S., Simons, G., 1981. On the theory of elliptically contoured distributions. Journal of Multivariate Analysis 11, 368–385.

Chollete, L.and Heinen, A., Valdesogo, A., 2008. Modeling international financial returns with a multivariate regime switching copula. MPRA paper 8114, University of Munich, Department of Economics.

Cook, R. D., Johnson, M. E., 1981. A family of distributions for modelling non-elliptically symmetric multivariate data. Journal of the Royal Statistical Society, Series B 43 (2).

Czado, C., Min, A., 2008. Bayesian inference for multivariate copulas using pair-copula constructions, in revision at the Journal of Financial Econometrics.

Czado, C., Min, A., Baumann, T., Dakovic, R., 2009. Pair-copula constructions for modeling exchange rate dependence, submitted to the Journal of Applied Econometrics.

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

Fischer, M., Köck, C., Schlüter, S., Weigert, F., 2007. Multivariate copula models at work: Outperforming the "desert island copula"? Tech. Rep. 79, Universität Erlangen-Nürnberg, Lehrstuhl für Statistik und Ökonometrie.

Genest, C., Gerber, H. U., Goovaerts, M. J., Laeven, R. J. A., 2009. Editorial to the special issue on modeling and measurement of multivariate risk in insurance and finance. Insurance: Mathematics and Economics 44 (2).

Hammersley, J. M., Clifford, P., 1971. Markov field on finite graphs and lattices, unpublished.

Heinen, A., Valdesogo, A., 2008. A canonical vine autoregressive model for large dimensions, working paper.

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

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

Joe, H., Li, H., Nikoloulopoulos, A., K., 2009. Tail dependence functions and vine copulas. Journal of Multivariate Analysis. Accepted for publication in 2009.

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

Kotz, S., Balakrishnan, N., Johnson, N., 2000. Multivariate Continuous Distributions, 2nd Edition. Wiley.

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

Mardia, K., 1964. Some results on the order statistics of the multivariate normal and pareto type i populations. Annals of Mathematical Statistics 35, 1815–1818.

Mardia, K. V., 1962. Multivariate pareto distributions. Annals of Mathematical Statistics 33, 1008–1015.

McNeil, A., Frey, R., Embrechts, P., 2006. Quantitative Risk Management: Concepts, Techniques, and Tools. Priceton University Press.

Nelsen, R., 1999. An Introduction to Copulas. Vol. 139 of Lecture notes in Statistics. Springer Verlag, New York.

Nikoloulopoulos, A., Karlis, D., 2008. Copula model evaluation based on parametric bootstrap. Computational Statistics and Data Analysis 52, 3342–3353.

Savu, C., Trede, M., May 2006. Hierarchical archimedian copulas. Presented at the International Conference on High Frequency Finance, Konstanz, Germany.

Schirmacher, D., Schirmacher, E., 2008. Multivariate dependence modeling using pair-copulas. Presented at The 2008 ERM Symposium, Chicago.

Takahasi, K., 1965. Note on the multivariate burr's distribution. Annals of the Institute of Statistical Mathematics 17, 257–260.

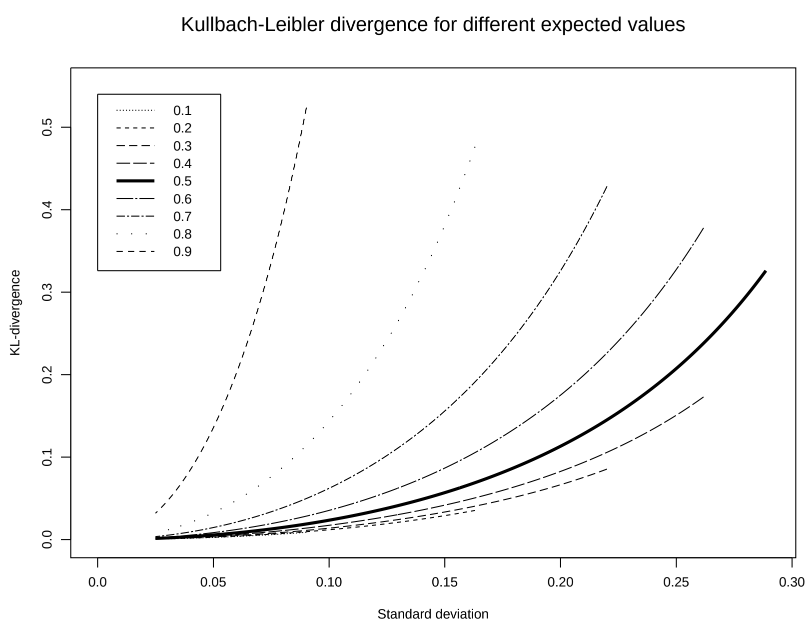

Figure 1: Kullback-Leibler divergence between the true and the approximated distribution as a function of $\mathrm{Sd}(X_3)$. The different curves correspond to different expected values $\mathrm{E}(X_3) = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$.

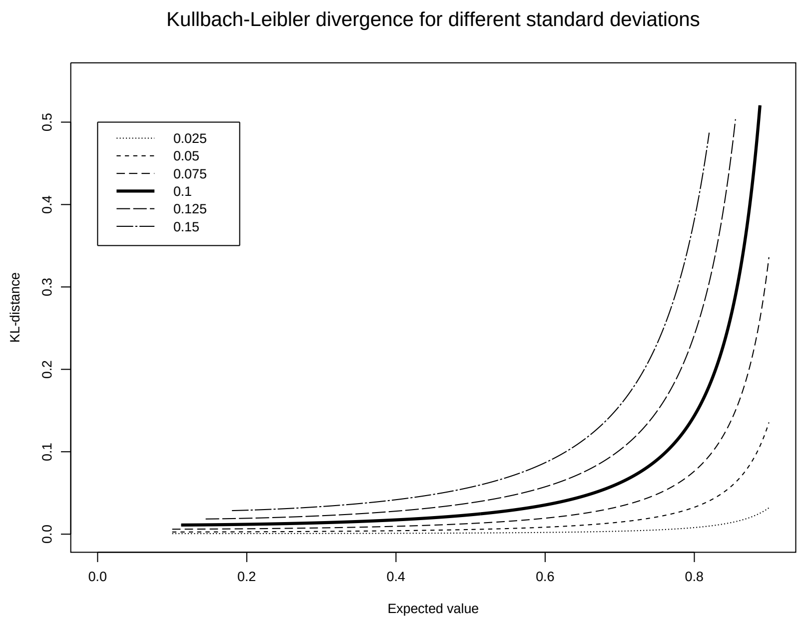

Figure 2: Kullback-Leibler divergence between the true and the approximated distribution as a function of $\mathrm{E}(X_3)$. The different curves correspond to different standard deviations $\mathrm{Sd}(X_3) = \{0.025, 0.050, 0.075, 0.100, 0.125, 0.150\}$.

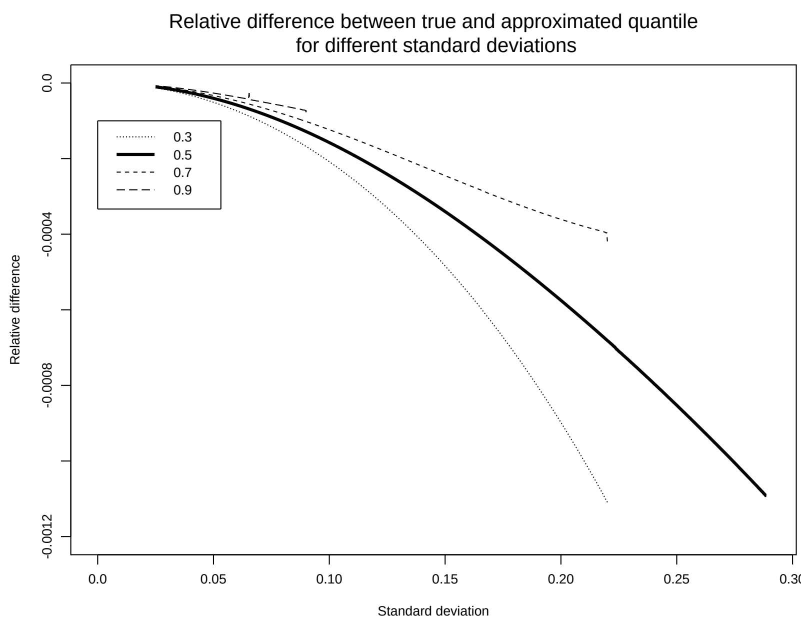

Figure 3: Relative difference between the true and the approximated 95% quantile of $X_1 + X_2$ as a function of $\mathrm{Sd}(X_3)$. The different curves correspond to different expected values $\mathrm{E}(X_3) = \{0.3, 0.5, 0.7, 0.9\}$.

A. Appendix

A.1. Computations for Example 3.1

In order to check whether the two remaining decompositions (conditioning on X_1 and X_2, respectively) are of the simplified form, we must compute the copula densities

$$c_{3-i,3|i}(F_{3-i|i}(x_{3-i})|x_i), F_{3|i}(x_3|x_i); x_i) = \frac{f_{3-i,3|i}(x_{3-i}, x_3)}{f_{3-i|i}(x_{3-i}|x_i) f_{3|i}(x_3|x_i)}, \ i = 1, 2. \quad (16)$$

The numerator of (16) is given by

$$\begin{aligned} f_{3-i,3|i}(x_{3-i}, x_3) &= \frac{f_{123}(x_1, x_2, x_3)}{f_i(x_i)} \\ &= \frac{(\nu + x_i^2)^{\frac{\nu+1}{2}}}{2^{\frac{\nu+2}{2}} \sqrt{\pi} \Gamma(\frac{\nu+1}{2}) \sqrt{1-\rho^2} x_3^{\frac{\nu+4}{2}}} \exp\left\{ -\frac{1}{2x_3} \left(\frac{x_i^2 + x_{3-i}^2 - 2\rho x_i x_{3-i}}{1-\rho^2} + \nu \right) \right\}. \end{aligned}$$

Since the unconditional distribution of (X_1, X_2) is the (standard) bivariate t-distribution with correlation ρ and ν degrees of freedom, the marginal distribution of $X_i, \ i = 1, 2$, is a standard t-distribution with ν degrees of freedom. Moreover, the conditional distribution of $[X_{3-i}|X_i = x_i], \ i = 1, 2$, is a t-distribution with location ρx_i, scale $\sqrt{\frac{(1-\rho^2)(\nu+x_i^2)}{\nu+1}}$ and $\nu + 1$ degrees of freedom. Hence,

$$f_{3-i|i}(x_{3-i}|x_i) = \frac{\Gamma(\frac{\nu+2}{2})}{\sqrt{\pi}\Gamma(\frac{\nu+1}{2})\sqrt{(1-\rho^2)(\nu + x_i^2)}} \left(1 + \frac{(x_{3-i} - \rho x_i)^2}{(1-\rho^2)(\nu + x_i^2)} \right)^{-\frac{\nu+2}{2}}, \ i = 1, 2.$$

Further, the second factor of the denominator of (16) is given by

$$\begin{aligned} f_{3|i}(x_3|x_i) &= \frac{f_{i|3}(x_i|x_3) f_3(x_3)}{f_i(x_i)} \\ &= \frac{(\nu + x_i^2)^{\frac{\nu+1}{2}}}{2^{\frac{\nu+1}{2}} \Gamma(\frac{\nu+1}{2}) x_3^{\frac{\nu+1}{2}+1}} \exp\left\{ -\frac{\nu + x_i^2}{2x_3} \right\}, \ i = 1, 2. \end{aligned}$$

Hence, $[X_3|X_i = x_i] \sim Gamma^{-1}\left(\frac{\nu+1}{2}, \frac{2}{\nu+x_i^2}\right)$.

Inserting this into (16), we obtain

$$\begin{aligned} c_{3-i,3|i}&(F_{3-i|i}(x_{3-i})|x_i), F_{3|i}(x_3|x_i); x_i) \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu+2}{2})} \left(1 + \frac{(x_{3-i} - \rho x_i)^2}{(1-\rho^2)(\nu + x_i^2)} \right)^{\frac{\nu+2}{2}} \sqrt{\frac{\nu + x_i^2}{2x_3}} \\ &\quad \cdot \exp\left\{ -\frac{\nu + x_i^2}{2x_3} \frac{(x_{3-i} - \rho x_i)^2}{(1-\rho^2)(\nu + x_i^2)} \right\}, \ i = 1, 2. \end{aligned}$$

Finally, let

$$u_{3-i|i} = F_{3-i|i}(x_{3-i}|x_i) = t_{\nu+1}\left(\frac{\sqrt{\nu+1}(x_{3-i}-\rho x_i)^2}{\sqrt{(1-\rho^2)(\nu+x_i^2)}}\right)$$
$$u_{3|i} = F_{3|i}(x_3|x_i) = 1 - \mathcal{Q}\left(\frac{\nu+1}{2}, \frac{2}{\nu+x_i^2}\right),\ i=1,2,$$

where t_ν is the cdf of the standard t-distribution with ν degrees of freedom, and $\mathcal{Q}(\alpha,x) = \frac{\Gamma(\alpha,x)}{\Gamma\alpha}$ is the regularised incomplete gamma function, $\Gamma(\alpha,x)$ being the complemented incomplete gamma function [Abramowitz and Stegun, 1972]. Thus,

$$\frac{\sqrt{\nu+1}(x_{3-i}-\rho x_i)^2}{\sqrt{(1-\rho^2)(\nu+x_i^2)}} = t_{\nu+1}^{-1}(u_{3-i|i})$$
$$\frac{\nu+x_i^2}{2x_3} = \mathcal{Q}^{-1}\left(\frac{\nu+1}{2}, 1-u_{3|i}\right),\ i=1,2.$$

We obtain

$$\begin{aligned} c_{3-i,3|i}(u_{3-i|i},&u_{3|i};x_i) \\ &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu+2}{2})}\left(1+\frac{t_{\nu+1}^{-1}(u_{3-i|i})^2}{\nu+1}\right)^{\frac{\nu+2}{2}}\sqrt{\mathcal{Q}^{-1}\left(\frac{\nu+1}{2}, 1-u_{3|i}\right)} \\ &\quad \cdot \exp\left\{-\mathcal{Q}^{-1}\left(\frac{\nu+1}{2}, 1-u_{3|i}\right)\frac{t_{\nu+1}^{-1}(u_{3-i|i})^2}{\nu+1}\right\} \\ &= c_{3-i,3|i}(u_{3-i|i}, u_{3|i}),\ i=1,2, \end{aligned}$$

which is of the simplified form.

A.2. Computations for Example 3.4

All marginals of the multivariate Burr distribution are Burr distributions [Takahasi, 1965]. More specifically,

$$f_{i,\dots,i+j}(x_i,\dots,x_{i+j}) = \frac{\alpha^j\beta^j\prod_{k=0}^j(\theta+k)x_{i+k}^{\beta-1}}{(\alpha\sum_{k=0}^j x_{i+k}^\beta+1)^{\theta+j+1}},\ x_k>0,\ k=0,\dots,j.$$

The copula densities at the second level of the structure (corresponding to lines four to six in (8)) are given by

$$\begin{aligned} c_{i,i+2|i+1}(F_{i|i+1}(x_i|x_{i+1}),&F_{i+2|i+1}(x_{i+2}|x_{i+1});x_{i+1}) \\ &= \frac{f_{i,i+2|i+1}(x_i,x_{i+2}|x_{i+1})}{f_{i|i+1}(x_i|x_{i+1})f_{i+2|i+1}(x_{i+2}|x_{i+1})},\ i=1,2,3. \end{aligned}$$

For the numerator, we have

$$f_{i,i+2|i+1}(x_i, x_{i+2}|x_{i+1}) = \frac{f_{i,i+1,i+2}(x_i, x_{i+1}, x_{i+2})}{f_{i+1}(x_{i+1})}$$
$$= \frac{\alpha^2\beta^2(\theta+1)(\theta+2)(\alpha x_{i+1}^\beta + 1)^{\theta+1} x_i^{\beta-1} x_{i+2}^{\beta-1}}{(\alpha(x_i^\beta + x_{i+2}^\beta) + \alpha x_{i+1}^\beta + 1)^{\theta+3}},$$

which is a bivariate Burr distribution in the two scaled variables $\frac{X_k}{(\alpha x_{i+1}^\beta+1)^{\theta+1}}$, $k = i, i+2$. Further,

$$f_{k|i+1}(x_k|x_{i+1}) = \frac{f_{k,i+1}(x_k, x_{i+1})}{f_{i+1}(x_{i+1})}$$
$$= \frac{\alpha\beta(\theta+1)(\alpha x_{i+1}^\beta + 1)^{\theta+1} x_k^{\beta-1}}{(\alpha x_k^\beta + \alpha x_{i+1}^\beta + 1)^{\theta+2}},\ k = i, i+2.$$

Hence,

$$c_{i,i+2|i+1}(F_{i|i+1}(x_i|x_{i+1}), F_{i+2|i+1}(x_{i+2}|x_{i+1}); x_{i+1})$$
$$= \frac{\theta+2}{\theta+1} \frac{(\alpha x_i^\beta + \alpha x_{i+1}^\beta + 1)^{\theta+2}(\alpha x_{i+2}^\beta + \alpha x_{i+1}^\beta + 1)^{\theta+2}(\alpha x_{i+1}^\beta + 1)^{\theta+1}}{(\alpha(x_i^\beta + x_{i+2}^\beta) + \alpha x_{i+1}^\beta + 1)^{\theta+3}}$$
$$= \frac{\theta+2}{\theta+1}\left(\frac{\alpha x_i^\beta + \alpha x_{i+1}^\beta + 1}{\alpha x_{i+1}^\beta + 1}\right)^{\theta+2}\left(\frac{\alpha x_{i+2}^\beta + \alpha x_{i+1}^\beta + 1}{\alpha x_{i+1}^\beta + 1}\right)^{\theta+2}$$
$$\cdot\left(\frac{\alpha x_i^\beta + \alpha x_{i+1}^\beta + 1}{\alpha x_{i+1}^\beta + 1} + \frac{\alpha x_{i+2}^\beta + \alpha x_{i+1}^\beta + 1}{\alpha x_{i+1}^\beta + 1} - 1\right)^{-(\theta+3)}.$$

Letting

$$u_{k|i+1} = F_{k|i+1}(x_k|x_{i+1}) = \int_0^{x_k} f_{k|i+1}(y|x_{i+1})dy$$
$$= 1 - \left(\frac{\alpha x_{i+1}^\beta + 1}{\alpha x_k^\beta + \alpha x_{i+1}^\beta + 1}\right)^{\theta+1},\ k = i, i+j,$$

such that

$$\frac{\alpha x_k^\beta + \alpha x_{i+1}^\beta + 1}{\alpha x_{i+1}^\beta + 1} = (1 - u_{k|i+1})^{-\frac{1}{\theta+1}},\ k = i, i+2,$$

we obtain

$$c_{i,i+2|i+1}(u_{i|i+1}, u_{i+2|i+1}; x_{i+1})$$
$$= \frac{\theta+2}{\theta+1}(1-u_{i|i+1})^{-\frac{\theta+2}{\theta+1}}(1-u_{i+2|i+1})^{-\frac{\theta+2}{\theta+1}}$$
$$\cdot((1-u_{i|i+1})^{-\frac{1}{\theta+1}}(1-u_{i+2|i+1})^{-\frac{1}{\theta+1}} - 1)^{-(\theta+3)}$$
$$= c_{i,i+2|i+1}(u_{i|i+1}, u_{i+2|i+1}),\ i = 1, 2, 3.$$

Correspondingly, the third level copula densities (lines seven and eight of (8)) are given by

$$
\begin{aligned}
c_{i,i+3|i+1,i+2}&(F_{i|i+1,i+2}(x_i|x_{i+1},x_{i+2}), F_{i+2|i+1,i+2}(x_{i+3}|x_{i+1},x_{i+2}); x_{i+1}, x_{i+2}) \\
&= \frac{f_{i,i+3|i+1,i+2}(x_i, x_{i+3}|x_{i+1}, x_{i+2})}{f_{i|i+1,i+3}(x_i|x_{i+1}, x_{i+2}) f_{i+3|i+1,i+2}(x_{i+3}|x_{i+1}, x_{i+2})},\ i = 1, 2,
\end{aligned}
$$

with

$$
\begin{aligned}
f_{i,i+3|i+1,i+2}&(x_i, x_{i+3}|x_{i+1}, x_{i+2}) \\
&= \frac{\alpha^2\beta^2(\theta+2)(\theta+3)(\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1)^{\theta+2} x_i^{\beta-1} x_{i+3}^{\beta-1}}{(\alpha(x_i^\beta + x_{i+3}^\beta) + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1)^{\theta+4}},
\end{aligned}
$$

and

$$
f_{k|i+1,i+2}(x_k|x_{i+1}, x_{i+2}) = \frac{\alpha\beta(\theta+2)(\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1)^{\theta+2} x_k^{\beta-1}}{(\alpha x_k^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1)^{\theta+3}},\ k = i, i+3.
$$

We obtain

$$
\begin{aligned}
&c_{i,i+3|i+1,i+2}(F_{i|i+1,i+2}(x_i|x_{i+1}, x_{i+2}), F_{i+3|i+1,i+2}(x_{i+3}|x_{i+1}, x_{i+2}); x_{i+1}, x_{i+2}) \\
&= \frac{\theta+3}{\theta+2} \left(\frac{\alpha x_i^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}{\alpha x_{i+1}^\beta + 1}\right)^{\theta+3} \left(\frac{\alpha x_{i+3}^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}{\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}\right)^{\theta+3} \\
&\quad \cdot \left(\frac{\alpha x_i^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}{\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1} + \frac{\alpha x_{i+3}^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}{\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1} - 1\right)^{-(\theta+4)}.
\end{aligned}
$$

Finally, we substitute with

$$
\begin{aligned}
u_{k|i+1,i+2} &= F_{k|i+1,i+2}(x_k|x_{i+1}, x_{i+2}) \\
&= 1 - \left(\frac{\alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}{\alpha x_k^\beta + \alpha x_{i+1}^\beta + \alpha x_{i+2}^\beta + 1}\right)^{\theta+2},\ k = i, i+3,
\end{aligned}
$$

and obtain

$$
\begin{aligned}
c_{i,i+3|i+1,i+2}&(u_{i|i+1,i+2}, u_{i+3|i+1,i+2}; x_{i+1}, x_{i+2}) \\
&= \frac{\theta+3}{\theta+2}(1 - u_{i|i+1,i+2})^{-\frac{\theta+3}{\theta+2}} (1 - u_{i+3|i+1,i+2})^{-\frac{\theta+3}{\theta+2}} \\
&\quad \cdot ((1 - u_{i|i+1,i+2})^{-\frac{1}{\theta+2}} (1 - u_{i+3|i+1,i+2})^{-\frac{1}{\theta+2}} - 1)^{-(\theta+4)} \\
&= c_{i,i+3|i+1,i+2}(u_{i|i+1,i+2}, u_{i+3|i+1,i+2})\ \ i = 1, 2.
\end{aligned}
$$

Corresponding computations for the top level copula density $c_{15|234}$ results in

$$\begin{aligned}
&c_{15|234}(u_{1|234}, u_{5|234}; x_2, x_3, x_4) \\
&\qquad = \frac{\theta+4}{\theta+3}(1-u_{1|234})^{-\frac{\theta+4}{\theta+3}}(1-u_{5|234})^{-\frac{\theta+4}{\theta+3}} \\
&\qquad\quad \cdot ((1-u_{1|234})^{-\frac{1}{\theta+3}}(1-u_{5|234})^{-\frac{1}{\theta+3}}-1)^{-(\theta+4)} \\
&\qquad = c_{15|234}(u_{1|234}, u_{5|234}).
\end{aligned}$$

Hence, the D-vine (8) is of the simplified form.