Simplified vine copula models: Approximations based on the simplifying assumption[*]

Fabian Spanhel

Department of Statistics, Ludwig-Maximilians-Universität München, Akademiestr. 1, 80799 Munich, Germany.

Center for Quantitative Risk Analysis, Ludwig-Maximilians-Universität München, Akademiestr. 1, 80799 Munich, Germany.
e-mail: spanhel@stat.uni-muenchen.de

Malte S. Kurz

Quantitative Investment Strategy, Scalable Capital GmbH[†], Prinzregentenstr. 48, 80538 Munich, Germany.

Center for Quantitative Risk Analysis, Ludwig-Maximilians-Universität München, Akademiestr. 1, 80799 Munich, Germany.
e-mail: malte.kurz@stat.uni-muenchen.de

Abstract: Vine copulas, or pair-copula constructions, have become an important tool in high-dimensional dependence modeling. Commonly, it is assumed that the data generating copula can be represented by a simplified vine copula (SVC). In this paper, we study the simplifying assumption and investigate the approximation of multivariate copulas by SVCs. We introduce the partial vine copula (PVC) which is a particular SVC where to any edge a j-th order partial copula is assigned. The PVC generalizes the partial correlation matrix and plays a major role in the approximation of copulas by SVCs. We investigate to what extent the PVC describes the dependence structure of the underlying copula. We show that, in general, the PVC does not minimize the Kullback-Leibler divergence from the true copula if the simplifying assumption does not hold. However, under regularity conditions, stepwise estimators of pair-copula constructions converge to the PVC irrespective of whether the simplifying assumption holds or not. Moreover, we elucidate why the PVC is often the best feasible SVC approximation in practice.

Keywords and phrases: Vine copula, pair-copula construction, simplifying assumption, partial vine copula, conditional copula.

Received March 2017.

Contents

1. Introduction

Copulas constitute an important tool to model dependence [28, 17, 25]. While it is easy to construct bivariate copulas, the construction of flexible high-dimensional copulas is a sophisticated problem. The introduction of simplified vine copulas (Joe [16]), or pair-copula constructions (Aas et al. [2]), has been an enormous advance for high-dimensional dependence modeling. Simplified vine copulas (SVCs) are hierarchical structures, constructed upon a sequence of bivariate unconditional copulas, which capture the conditional dependence between pairs of random variables if the data generating process satisfies the simplifying assumption. In this case, all conditional copulas of the data generating vine collapse to unconditional copulas and the true copula can be represented in terms of a SVC. Vine copula methodology and applications have been extensively developed under the simplifying assumption [6, 11, 18, 20, 29], with

studies showing the superiority of SVC models over elliptical copulas and nested Archimedean copulas (Aas and Berg [1], Fischer et al. [8]).

Although some copulas can be expressed as a SVC, the simplifying assumption is not true in general. Hobæk Haff, Aas and Frigessi [14] point out that the simplifying assumption is in general not valid and provide examples of multivariate distributions which do not satisfy the simplifying assumption. Stöber, Joe and Czado [37] show that the Clayton copula is the only Archimedean copula for which the simplifying assumption holds, while the Student-t copula is the only SVC arising from a scale mixture of normal distributions. In fact, it is very unlikely that the unknown data generating process satisfies the simplifying assumption in a strict mathematical sense. As a result, researchers have recently started to investigate new dependence concepts that are related to the simplifying assumption. In particular, studies on the bivariate partial copula, a generalization of the partial correlation coefficient, have (re-)emerged lately [5, 9, 10, 35, 31]. Estimators for non-simplified vine copula models are proposed by Schellhase and Spanhel [33] and Vatter and Nagler [38] using a non-parameteric approach and generalized additive models, respectively.

We introduce the partial vine copula (PVC) which generalizes the partial correlation matrix. The PVC is a particular SVC where to any edge a j-th order partial copula is assigned. We investigate several properties of the PVC and show to what extent the dependence structure of the underlying distribution is captured. The PVC plays a crucial role in terms of approximating a multivariate copula by a SVC. We show that stepwise estimators of SVCs converge to the PVC regardless of whether the simplifying assumption holds. However, we also prove that the PVC may not minimize the Kullback-Leibler divergence from the true copula and thus may not be the best approximation in the space of SVCs. This result is rather surprising, because it implies that it may not be optimal to specify the true copulas in the first tree of a SVC approximation. Moreover, joint and stepwise estimators of SVCs may not converge to the same probability limit any more if the simplifying assumption does not hold. Nevertheless, the PVC is often the best SVC approximation in practice because only a stepwise estimation is feasible. The PVC is used by Nagler and Czado [27] to construct a new non-parametric estimator of a multivariate distribution that can outperform classical non-parametric approaches. Moreover, Kurz and Spanhel [24] apply the PVC to test the simplifying assumption in high-dimensional vine copulas. All in all, these facts highlight the practical importance of the PVC for multivariate dependence modeling.

The rest of this paper is organized as follows. (Simplified) vine copulas, the simplifying assumption, conditional and partial copulas, are discussed in Section 2. The PVC and j-th order partial copulas are introduced in Section 3. Properties of the PVC and examples are presented in Section 4. In Section 5 we analyze the role of the PVC for SVC approximations and explain why the PVC is the best feasible approximation in practical applications. A parametric estimator for the PVC is presented in Section 6 and implications for the stepwise and joint maximum likelihood estimator of SVCs are illustrated. Section 7 contains some concluding remarks.

2. Simplified vine copulas, conditional and partial copulas

This section introduces vine copulas, the simplifying assumption and related concepts. First, notation and assumptions are stated. Then, we discuss (simplified) vine copulas and the simplifying assumption. Thereafter, we introduce the partial copula which can be considered as a generalization of the partial correlation coefficient and as an approximation of a bivariate conditional copula.

2.1. Notation and assumptions

The following notation and assumptions are used throughout the paper. We write $X_{1:d} := (X_1, \ldots, X_d)$, so that $F_{X_{1:d}}(x_{1:d}) := \mathbb{P}(\forall i = 1, \ldots, d \colon X_i \leq x_i)$. Let $\mathrm{d}x_{1:d} := \mathrm{d}x_1 \ldots \mathrm{d}x_d$ denote the variables of integration in $\int f_{X_{1:d}}(x_{1:d})\mathrm{d}x_{1:d}$. $C^\perp$ refers to the independence copula. $X \perp Y$ means that X and Y are stochastically independent. For $1 \leq k \leq d$, the partial derivative of g w.r.t. the k-th argument is denoted by $\partial_k g(x_{1:d})$. We write $\mathbb{1}_{\{A\}} = 1$ if A is true, and $\mathbb{1}_{\{A\}} = 0$ otherwise. For simplicity, we assume that all random variables are real-valued and continuous. Let $d \geq 3$, if not otherwise specified, and $\mathcal{C}_d$ be the space of absolutely continuous d-dimensional copulas with positive density (a.e.). The distribution function of a random vector $U_{1:d}$ with uniform margins is denoted by $F_{1:d} = C_{1:d} \in \mathcal{C}_d$. We set $\mathcal{I}_l^d := \{(i, j) \colon j = l, \ldots, d - 1, i = 1, \ldots, d - j\}$ and $S_{ij} := i + 1 : i + j - 1 := i + 1, \ldots, i + j - 1$. We focus on D-vine copulas, but all results carry over to regular vine copulas (Bedford and Cooke [4], Kurowicka and Joe [23]). An overview of the used notation can be found in Table 1. All proofs are deferred to the Appendix.

2.2. Vine copulas and the simplifying assumption

Definition 2.1. (Simplified D-vine copula or pair-copula construction – Joe [16], Aas et al. [2])
For $(i, j) \in \mathcal{I}_1^d$, let $C_{i,i+j;S_{ij}}^{\mathrm{svc}} \in \mathcal{C}_2$ with density $c_{i,i+j;S_{ij}}^{\mathrm{svc}}$. For $j = 1$ and $i = 1, \ldots, d - j$, we set $C_{i,i+j;S_{ij}}^{\mathrm{svc}} = C_{i,i+1}^{\mathrm{svc}}$ and $u_{k|S_{ij}}^{\mathrm{svc}} = u_k$ for $k = i, i + j$. For $(i, j) \in \mathcal{I}_2^d$, define

$$u_{i|S_{ij}}^{\mathrm{svc}} := F_{i|S_{ij}}^{\mathrm{svc}}(u_i|u_{S_{ij}}) = \partial_2 C_{i,i+j-1;S_{i,j-1}}^{\mathrm{svc}}(u_{i|S_{i,j-1}}^{\mathrm{svc}}, u_{i+j-1|S_{i,j-1}}^{\mathrm{svc}}),$$
$$u_{i+j|S_{ij}}^{\mathrm{svc}} := F_{i+j|S_{ij}}^{\mathrm{svc}}(u_{i+j}|u_{S_{ij}}) = \partial_1 C_{i+1,i+j;S_{i+1,j-1}}^{\mathrm{svc}}(u_{i+1|S_{i+1,j-1}}^{\mathrm{svc}}, u_{i+j|S_{i+1,j-1}}^{\mathrm{svc}}).$$

Then

$$c_{1:d}^{\mathrm{svc}}(u_{1:d}) = \prod_{(i,j)\in\mathcal{I}_1^d} c_{i,i+j;S_{ij}}^{\mathrm{svc}}\left(u_{i|S_{ij}}^{\mathrm{svc}}, u_{i+j|S_{ij}}^{\mathrm{svc}}\right)$$

is the density of a d-dimensional simplified D-vine copula $C_{1:d}^{\mathrm{svc}}$. We denote the space of d-dimensional simplified D-vine copulas by $\mathcal{C}_d^{\mathrm{svc}}$.

TABLE 1
Notation for simplified D-vine copulas. $U_{1:d}$ has standard uniform margins, $d \geq 3$,
$(i,j) \in \mathcal{I}_1^d$, $k = i, i+j$.

Notation	Explanation
$F_{1:d}$ or $C_{1:d}$	cdf and copula of $U_{1:d}$
$\mathcal{C}_d$	space of d-dimensional copulas with positive density
$\mathcal{C}_d^{\mathrm{SVC}}$	space of d-dimensional simplified D-vine copulas with positive density
$\mathcal{I}_1^d$	$\mathcal{I}_1^d := \{(i,j): j = 1, \dots, d-1, i = 1, \dots, d-j\}$, the conditioned set of a D-vine copula density
S_{ij}	$S_{ij} := i+1 : i+j-1 := i+1, \dots, i+j-1$, the conditioning set of an edge in a D-vine
$U_{k\mid S_{ij}}$	$F_{k\mid S_{ij}}(U_k \mid U_{S_{ij}})$, conditional probability integral transform (CPIT) of U_k w.r.t. $U_{S_{ij}}$
$C_{i,i+j;S_{ij}}$	bivariate conditional copula of $F_{i,i+j\mid S_{ij}}$, i.e., $C_{i,i+j;S_{ij}} = F_{U_{i\mid S_{ij}},U_{i+j\mid S_{ij}} \mid U_{S_{ij}}}$
$C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$	arbitrary bivariate (unconditional) copula that is used to model $C_{i,i+j;S_{ij}}$
$C_{i,i+j;S_{ij}}^{\mathrm{P}}$	partial copula of $C_{i,i+j;S_{ij}}$, i.e., $C_{i,i+j;S_{ij}}^{\mathrm{P}} = F_{U_{i\mid S_{ij}},U_{i+j\mid S_{ij}}}$
$C_{i,i+j;S_{ij}}^{\mathrm{PVC}}$	$(j-1)$-th order partial copula of $C_{i,i+j;S_{ij}}$
$U_{k\mid S_{ij}}^{\mathrm{PVC}}$	$F_{k\mid S_{ij}}^{\mathrm{PVC}}(U_k \mid U_{S_{ij}})$, $(j-2)$-th order partial probability integral transform (PPIT) of U_k w.r.t. $U_{S_{ij}}$
$C_{1:d}^{\mathrm{PVC}}$	Partial vine copula (PVC) of $C_{1:d}$, if $d = 3$, then $c_{1:3}^{\mathrm{PVC}}(u_{1:3}) = c_{12}(u_1, u_2)\, c_{23}(u_2, u_3)\, c_{13;2}^{\mathrm{PVC}}(u_{1\mid 2}, u_{3\mid 2})$

Simplified (regular) vine copulas can be considered as an ordered sequence of trees, where j refers to the number of the tree and a bivariate unconditional copula $C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$ is assigned to each of the $d-j$ edges of tree j (Bedford and Cooke [4]). The left hand side of Figure 1 shows the graphical representation of a simplified D-vine copula for $d = 4$, i.e.,

$$c_{1:4}^{\mathrm{SVC}}(u_{1:4}) = \underbrace{c_{12}^{\mathrm{SVC}}(u_1, u_2) c_{23}^{\mathrm{SVC}}(u_2, u_3) c_{34}^{\mathrm{SVC}}(u_3, u_4)}_{\text{first tree}}$$
$$\times \underbrace{c_{13;2}^{\mathrm{SVC}}(u_{1\mid 2}^{\mathrm{SVC}}, u_{3\mid 2}^{\mathrm{SVC}}) c_{24;3}^{\mathrm{SVC}}(u_{2\mid 3}^{\mathrm{SVC}}, u_{4\mid 3}^{\mathrm{SVC}})}_{\text{second tree}} \times \underbrace{c_{14;2:3}^{\mathrm{SVC}}(u_{1\mid 2:3}^{\mathrm{SVC}}, u_{4\mid 2:3}^{\mathrm{SVC}})}_{\text{third tree}}.$$

The bivariate unconditional copulas $C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$ are also called pair-copulas, so that the resulting model is often termed a pair-copula construction. By means of SVC models one can construct a wide variety of flexible multivariate copulas because each of the $d(d-1)/2$ bivariate unconditional copulas $C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$ can be chosen arbitrarily. The resulting model is always a valid d-dimensional copula. Moreover, a pair-copula construction does not suffer from the curse of dimensions because it is build upon a sequence of bivariate unconditional copulas. As
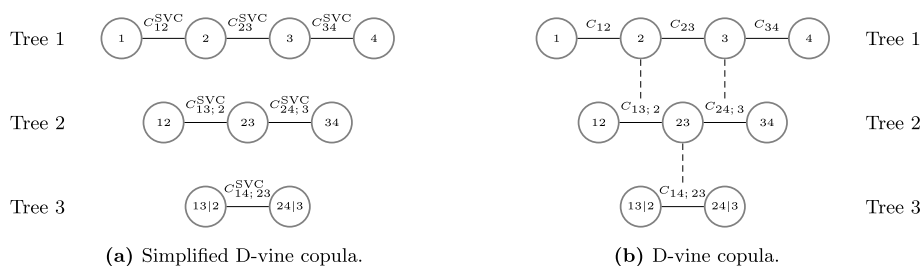
(a) Simplified D-vine copula.   **(b)** D-vine copula.

FIG 1. *(Simplified) D-vine copula representation if $d = 4$. The influence of conditioning variables on the conditional copulas is indicated by dashed lines.*

a result, SVCs are very attractive for high-dimensional applications. Obviously, not every multivariate copula can be represented by a SVC. However, every copula can be represented by the following (non-simplified) D-vine copula.

Definition 2.2. (D-vine copula – Kurowicka and Cooke [22])
Let $U_{1:d}$ be a random vector with cdf $F_{1:d} = C_{1:d} \in \mathcal{C}_d$. For $j = 1$ and $i = 1, \ldots, d-j$, we set $C_{i,i+j;S_{ij}} = C_{i,i+1}$ and $u_{k|S_{ij}} = u_k$ for $k = i, i+j$. For $(i,j) \in \mathcal{I}_2^d$, let $C_{i,i+j;S_{ij}}$ denote the conditional copula of $F_{i,i+j|S_{ij}}$ (Definition 2.5) and let $u_{k|S_{ij}} := F_{k|S_{ij}}(u_k|u_{S_{ij}})$ for $k = i, i+j$. The density of a D-vine copula decomposes the copula density of $U_{1:d}$ into $d(d-1)/2$ bivariate conditional copula densities $c_{i,i+j;S_{ij}}$ according to the following factorization:

$$c_{1:d}(u_{1:d}) = \prod_{(i,j)\in\mathcal{I}_1^d} c_{i,i+j;S_{ij}}(u_{i|S_{ij}}, u_{i+j|S_{ij}}|u_{S_{ij}}).$$

Contrary to a simplified D-vine copula in Definition 2.1, a bivariate conditional copula $C_{i,i+j;S_{ij}}$, which is a function of $j + 1$ variables, is assigned to each edge of a D-vine copula in Definition 2.2. The influence of the conditioning variables on the conditional copulas is illustrated by dashed lines in the right hand side of Figure 1. In applications, the simplifying assumption is typically imposed. That is, it is assumed that all bivariate conditional copulas of the data generating vine copula degenerate to bivariate unconditional copulas.

Definition 2.3. (The simplifying assumption – Hobæk Haff, Aas and Frigessi [14])
The D-vine copula in Definition 2.2 satisfies the simplifying assumption if $c_{i,i+j;S_{ij}}(\cdot, \cdot|u_{S_{ij}})$ does not depend on $u_{S_{ij}}$ for all $(i,j)\in\mathcal{I}_2^d$.

If the data generating copula satisfies the simplifying assumption it can be represented by a SVC, resulting in fast and simple statistical inference. Several methods for the consistent specification and estimation of pair-copula constructions have been developed under this assumption (Hobæk Haff [13], Dißmann

et al. [6]). However, in view of Definition 2.2 and Definition 2.1 it is evident that it is extremely unlikely that the data generating vine copula strictly satisfies the simplifying assumption in practical applications.

Several questions arise if the data generating process does not satisfy the simplifying assumption and a simplified D-vine copula model (Definition 2.1) is used to approximate a general D-vine copula (Definition 2.2). First of all, what bivariate unconditional copulas $C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$ should be chosen in Definition 2.1 to model the bivariate conditional copulas $C_{i,i+j;S_{ij}}$ in Definition 2.2 so that the best approximation w.r.t. a certain criterion is obtained? If the simplifying assumption does not hold for the data generating vine copula, what is the SVC model that established stepwise procedures (asymptotically) specify and estimate? What are the properties of an optimal approximation? Before we address these questions in Section 5, it is useful to recall the definition of the conditional and partial copula in Section 2.3 and introduce the partial vine copula in Section 3 and Section 4.

2.3. Conditional and partial copulas

Definition 2.4. (Conditional probability integral transform (CPIT))
Let $U_{1:d} \sim F_{1:d} \in \mathcal{C}_d$, $(i,j) \in \mathcal{I}_2^d$ and $k = i, i+j$. We call $U_{k|S_{ij}} := F_{k|S_{ij}}(U_k|U_{S_{ij}})$ the conditional probability integral transform of U_k w.r.t. $U_{S_{ij}}$.

It can be readily verified that, under the assumptions in Definition 2.4, $U_{k|S_{ij}} \sim \mathcal{U}(0,1)$ and $U_{k|S_{ij}} \perp U_{S_{ij}}$. Thus, applying the random transformation $F_{k|S_{ij}}(\cdot|U_{S_{ij}})$ to U_k removes possible dependencies between U_k and $U_{S_{ij}}$. The CPIT $U_{k|S_{ij}}$ can be interpreted as the remaining variation in U_k that can not be explained by $U_{S_{ij}}$. This interpretation is crucial for understanding the conditional and partial copula which are related to the (conditional) joint distribution of CPITs. The conditional copula has been introduced by Patton [30] and we restate its definition here.[1]

Definition 2.5. (Bivariate conditional copula – Patton [30])
Let $U_{1:d} \sim F_{1:d} \in \mathcal{C}_d$ and $(i,j) \in \mathcal{I}_2^d$. The (a.e.) unique conditional copula $C_{i,i+j;S_{ij}}$ of the conditional distribution $F_{i,i+j|S_{ij}}$ is defined by

$$C_{i,i+j;S_{ij}}(a,b|u_{S_{ij}}) := \mathbb{P}(U_{i|S_{ij}} \leq a, U_{i+j|S_{ij}} \leq b|U_{S_{ij}} = u_{S_{ij}})$$
$$= F_{i,i+j|S_{ij}}(F_{i|S_{ij}}^{-1}(a|u_{S_{ij}}), F_{i+j|S_{ij}}^{-1}(b|u_{S_{ij}})|u_{S_{ij}}).$$

Equivalently, we have that

$$F_{i,i+j|S_{ij}}(u_i, u_{i+j}|u_{S_{ij}}) = C_{i,i+j;S_{ij}}(F_{i|S_{ij}}(u_i|u_{S_{ij}}), F_{i+j|S_{ij}}(u_{i+j}|u_{S_{ij}})|u_{S_{ij}}).$$

---

[1] Patton's notation for the conditional copula is given by $C_{i,i+j|S_{ij}}$. Originally, this notation has also been used in the vine copula literature [2, 23, 3]. However, the current notation for a(n) (un)conditional copula that is assigned to an edge of a vine is given by $C_{i,i+j;S_{ij}}$ and $C_{i,i+j|S_{ij}}$ is used to denote $F_{U_i, U_{i+j}|U_{S_{ij}}}$ [18, 37, 21]. In order to avoid possible confusions, we use $C_{i,i+j;S_{ij}}$ to denote a conditional copula and $C_{i,i+j;S_{ij}}^{\mathrm{SVC}}$ to denote an unconditional copula.

Thus, the effect of a change in $u_{S_{ij}}$ on the conditional distribution function $F_{i,i+j|S_{ij}}(u_i, u_{i+j}|u_{S_{ij}})$ can be separated into two effects. First, the values of the CPITs, $(F_{i|S_{ij}}(u_i|u_{S_{ij}}), F_{i+j|S_{ij}}(u_{i+j}|u_{S_{ij}}))$, at which the conditional copula is evaluated, may change. Second, the functional form of the conditional copula $C_{i,i+j;S_{ij}}(\cdot,\cdot|u_{S_{ij}})$ may vary. In comparison to the conditional copula, which is the conditional distribution of two CPITs, the partial copula is the unconditional distribution and copula of two CPITs.

Definition 2.6. (Bivariate partial copula - Bergsma [5])
Let $U_{1:d} \sim F_{1:d} \in \mathcal{C}_d$ and $(i,j) \in \mathcal{I}_2^d$. The partial copula $C_{i,i+j;S_{ij}}^{\mathrm{P}}$ of the distribution $F_{i,i+j|S_{ij}}$ is defined by

$$C_{i,i+j;S_{ij}}^{\mathrm{P}}(a,b) := \mathbb{P}(U_{i|S_{ij}} \leq a, U_{i+j|S_{ij}} \leq b).$$

Since $U_{i|S_{ij}} \perp U_{S_{ij}}$ and $U_{i+j|S_{ij}} \perp U_{S_{ij}}$, the partial copula represents the distribution of random variables which are individually independent of the conditioning vector $U_{S_{ij}}$. This is similar to the partial correlation coefficient, which is the correlation of two random variables from which the linear influence of the conditioning vector has been removed. The partial copula can also be interpreted as the expected conditional copula,

$$C_{i,i+j;S_{ij}}^{\mathrm{P}}(a,b) = \int_{[0,1]^{j-1}} C_{i,i+j;S_{ij}}(a,b|u_{S_{ij}}) \mathrm{d}F_{S_{ij}}(u_{S_{ij}}),$$

and be considered as an approximation of the conditional copula. Indeed, it is easy to show that the partial copula $C_{i,i+j;S_{ij}}^{\mathrm{P}}$ minimizes the Kullback-Leibler divergence from the conditional copula $C_{i,i+j;S_{ij}}$ in the space of absolutely continuous bivariate distribution functions. The partial copula is first mentioned by Bergsma [5] who applies the partial copula to test for conditional independence. Recently, there has been a renewed interest in the partial copula. Spanhel and Kurz [35] investigate properties of the partial copula and mention some explicit examples whereas Gijbels, Omelka and Veraverbeke [9, 10] and Portier and Segers [31] focus on the non-parametric estimation of the partial copula.

3. Higher-order partial copulas and the partial vine copula

A generalization of the partial correlation coefficient that is different from the partial copula is given by a higher-order partial copula. To illustrate this relation, let us recall the common definition of the partial correlation coefficient. Assume that all univariate margins of $Y_{1:d}$ have zero mean and finite variance. For $k = i, i + j$, let $\mathcal{P}(Y_k|Y_{S_{ij}})$ denote the best linear predictor of Y_k w.r.t $Y_{S_{ij}}$ which minimizes the mean squared error so that $\tilde{\mathcal{E}}_{k|S_{ij}} = Y_k - \mathcal{P}(Y_k|Y_{S_{ij}})$ is the corresponding prediction error. The partial correlation coefficient of Y_i and Y_{i+j} given $Y_{S_{ij}}$ is then defined by $\rho_{i,i+j;S_{ij}} = \mathbb{C}\mathrm{orr}[\tilde{\mathcal{E}}_{i|S_{ij}}, \tilde{\mathcal{E}}_{i+j|S_{ij}}]$. An equivalent definition is given as follows. For $i = 1, \ldots, d - 2$, let

$$\mathcal{E}_{i|i+1} := Y_i - \mathcal{P}(Y_i|Y_{i+1}), \quad \text{and} \quad \mathcal{E}_{i+2|i+1} := Y_{i+2} - \mathcal{P}(Y_{i+2}|Y_{i+1}). \qquad (3.1)$$

Moreover, for $j = 3, \ldots, d - 1$, and $i = 1, \ldots, d - j$, define

$$
\begin{aligned}
\mathcal{E}_{i|S_{ij}} &:= \mathcal{E}_{i|S_{i,j-1}} - \mathcal{P}(\mathcal{E}_{i|S_{i,j-1}}|\mathcal{E}_{i+j-1|S_{i,j-1}}), \\
\mathcal{E}_{i+j|S_{ij}} &:= \mathcal{E}_{i+j|S_{i+1,j-1}} - \mathcal{P}(\mathcal{E}_{i+j|S_{i+1,j-1}}|\mathcal{E}_{i+1|S_{i+1,j-1}}).
\end{aligned}
\tag{3.2}
$$

It is easy to show that $\mathcal{E}_{k|S_{ij}} = \tilde{\mathcal{E}}_{k|S_{ij}}$ for all $k = i, i + j$ and $(i, j) \in \mathcal{I}_2^d$. That is, $\mathcal{E}_{k|S_{ij}}$ is the error of the best linear prediction of Y_k in terms of $Y_{S_{ij}}$. Thus, $\rho_{i,i+j;S_{ij}} = \mathbb{C}\mathrm{orr}[\mathcal{E}_{i|S_{ij}}, \mathcal{E}_{i+j|S_{ij}}]$. However, the interpretation of the partial correlation coefficient as a measure of conditional dependence is different depending on whether one considers it as the correlation of $(\tilde{\mathcal{E}}_{i|S_{ij}}, \tilde{\mathcal{E}}_{i+j|S_{ij}})$ or $(\mathcal{E}_{i|S_{ij}}, \mathcal{E}_{i+j|S_{ij}})$. For instance, $\rho_{14;23} = \mathbb{C}\mathrm{orr}[\tilde{\mathcal{E}}_{1|23}, \tilde{\mathcal{E}}_{4|23}]$ can be interpreted as the correlation between Y_1 and Y_4 after each variable has been corrected for the linear influence of $Y_{2:3}$, i.e., $\mathbb{C}\mathrm{orr}[g(\tilde{\mathcal{E}}_{k|23}), h(Y_{2:3})] = 0$ for all linear functions g and h. The idea of the partial copula is to replace the prediction errors $\mathcal{E}_{1|23}$ and $\mathcal{E}_{4|23}$ by the CPITS $U_{1|23}$ and $U_{4|23}$ which are independent of $Y_{2:3}$. On the other side, $\rho_{14;23} = \mathbb{C}\mathrm{orr}\left[\mathcal{E}_{1|23}, \mathcal{E}_{4|23}\right]$ is the correlation of $(\mathcal{E}_{1|2}, \mathcal{E}_{4|3})$ after $\mathcal{E}_{1|2}$ has been corrected for the linear influence of $\mathcal{E}_{3|2}$, and $\mathcal{E}_{4|3}$ has been corrected for the linear influence of $\mathcal{E}_{2|3}$. Consequently, a different generalization of the partial correlation coefficient emerges if we do not only decorrelate the involved random variables in (3.1) and (3.2) but render them independent by replacing each expression of the form $X - \mathcal{P}(X|Z)$ in (3.1) and (3.2) by the corresponding CPIT $F_{X|Z}(X|Z)$. The joint distribution of a resulting pair of random variables is given by the j-th order partial copula. The set of these copulas together with a vine structure constitute the partial vine copula.

Definition 3.1. (Partial vine copula (PVC) and j-th order partial copulas)

Consider the D-vine copula $C_{1:d} \in \mathcal{C}_d$ stated in Definition 2.2. In the first tree, we set for $i = 1, \ldots, d - 1$: $C_{i,i+1}^{\mathrm{PVC}} = C_{i,i+1}$, while in the second tree, we denote for $i = 1, \ldots, d - 2$, $k = i, i + 2$: $C_{i,i+2;i+1}^{\mathrm{PVC}} = C_{i,i+2;i+1}^{\mathrm{P}}$ and $U_{k|i+1}^{\mathrm{PVC}} = U_{k|i+1} = F_{k|i+1}(U_k|U_{i+1})$. In the remaining trees $j = 3, \ldots, d - 1$, for $i = 1, \ldots, d - j$, we define

$$
\begin{aligned}
U_{i|S_{ij}}^{\mathrm{PVC}} &:= F_{i|S_{ij}}^{\mathrm{PVC}}(U_i|U_{S_{ij}}) \\
&:= \partial_2 C_{i,i+j-1;S_{i,j-1}}^{\mathrm{PVC}}(U_{i|S_{i,j-1}}^{\mathrm{PVC}}, U_{i+j-1|S_{i,j-1}}^{\mathrm{PVC}}),
\end{aligned}
\tag{3.3}
$$

$$
\begin{aligned}
U_{i+j|S_{ij}}^{\mathrm{PVC}} &:= F_{i+j|S_{ij}}^{\mathrm{PVC}}(U_{i+j}|U_{S_{ij}}) \\
&:= \partial_1 C_{i+1,i+j;S_{i+1,j-1}}^{\mathrm{PVC}}(U_{i+1|S_{i+1,j-1}}^{\mathrm{PVC}}, U_{i+j|S_{i+1,j-1}}^{\mathrm{PVC}}),
\end{aligned}
\tag{3.4}
$$

and

$$
C_{i,i+j;\,S_{ij}}^{\mathrm{PVC}}(a, b) := \mathbb{P}(U_{i|S_{ij}}^{\mathrm{PVC}} \le a, U_{i+j|S_{ij}}^{\mathrm{PVC}} \le b).
$$

We call the resulting SVC $C_{1:d}^{\mathrm{PVC}}$ the partial vine copula (PVC) of $C_{1:d}$. Its density is given by

$$
c_{1:d}^{\mathrm{PVC}}(u_{1:d}) := \prod_{(i,j)\in\mathcal{I}_1^d} c_{i,i+j;\,S_{ij}}^{\mathrm{PVC}}(u_{i|S_{ij}}^{\mathrm{PVC}}, u_{i+j|S_{ij}}^{\mathrm{PVC}}),
$$

where $u_{k|S_{ij}}^{\text{PVC}}$, $k = i, i+j$ is given by $F_{k|S_{ij}}^{\text{PVC}}$ (defined in (3.3), (3.4)) evaluated at $(u_k, u_{S_{ij}})$. For $k = i, i+j$, we call $U_{k|S_{ij}}^{\text{PVC}}$ the $(j-2)$-th order partial probability integral transform (PPIT) of U_k w.r.t. $U_{S_{ij}}$ and $C_{i,i+j;\,S_{ij}}^{\text{PVC}}$ the $(j-1)$-th order partial copula of $F_{i,i+j|S_{ij}}$ that is induced by $C_{1:d}^{\text{PVC}}$.

Note that the first-order partial copula coincides with the partial copula of a conditional distribution with one conditioning variable. If $j \geq 3$, we call $C_{i,i+j;\,S_{ij}}^{\text{PVC}}$ a higher-order partial copula. It is easy to show that, for all $(i,j) \in \mathcal{I}_2^d$, $U_{i|S_{ij}}^{\text{PVC}}$ is the CPIT of $U_{i|S_{i,j-1}}^{\text{PVC}}$ w.r.t. $U_{i+j-1|S_{i,j-1}}^{\text{PVC}}$ and $U_{i+j|S_{ij}}^{\text{PVC}}$ is the CPIT of $U_{i+j|S_{i+1,j-1}}^{\text{PVC}}$ w.r.t. $U_{i+1|S_{i+1,j-1}}^{\text{PVC}}$. Thus, PPITs are uniformly distributed and higher-order partial copulas are indeed copulas. Since $U_{i|S_{ij}}^{\text{PVC}}$ is the CPIT of $U_{i|S_{i,j-1}}^{\text{PVC}}$ w.r.t. $U_{i+j-1|S_{i,j-1}}^{\text{PVC}}$, it is independent of $U_{i+j-1|S_{i,j-1}}^{\text{PVC}}$. However, in general it is not true that $U_{i|S_{ij}}^{\text{PVC}} \perp U_{S_{ij}}$ as the following lemma clarifies.

Lemma 3.1. (Relation between PPITs and CPITs)
For $(i,j) \in \mathcal{I}_2^d$ and $k = i, i+j$, it holds:

$$U_{k|S_{ij}}^{PVC} \perp U_{S_{ij}} \Leftrightarrow U_{k|S_{ij}}^{PVC} \stackrel{a.s.}{=} U_{k|S_{ij}}.$$

Note that $(U_{i|S_{ij}}^{\text{PVC}}, U_{i+j|S_{ij}}^{\text{PVC}}) \stackrel{d}{=} (U_{i|S_{ij}}, U_{i+j|S_{ij}})$ if and only if $C_{i,i+j;\,S_{ij}}^{\text{PVC}} = C_{i,i+j;\,S_{ij}}^{\text{P}}$. Consequently, if a higher-order partial copula does not coincide with the partial copula, it describes the distribution of a pair of uniformly distributed random variables which are neither jointly nor both individually independent of the conditioning variables of the corresponding conditional copula. If the simplifying assumption holds, then $C_{1:d} = C_{1:d}^{\text{PVC}}$, i.e., higher-order partial copulas, partial copulas and conditional copulas coincide. This insight is used by Kurz and Spanhel [24] to develop tests for the simplifying assumption in high-dimensional vine copulas.

Let $k = i, i+j$, and $G_{k|S_{ij}}^{\text{PVC}}(t_k | t_{S_{ij}}) = (F_{k|S_{ij}}^{\text{PVC}})^{-1}(t_k | t_{S_{ij}})$ denote the inverse of $F_{k|S_{ij}}^{\text{PVC}}(\cdot | t_{S_{ij}})$ w.r.t. the first argument. The $(j-1)$-th order partial copula is then given by

$$
\begin{aligned}
C_{i,i+j;\,S_{ij}}^{\text{PVC}}(a,b) &= \mathbb{P}(U_{i|S_{ij}}^{\text{PVC}} \leq a, U_{i+j|S_{ij}}^{\text{PVC}} \leq b) \\
&= \mathbb{E}\big[\mathbb{P}(U_{i|S_{ij}}^{\text{PVC}} \leq a, U_{i+j|S_{ij}}^{\text{PVC}} \leq b | U_{S_{ij}})\big] \\
&= \int_{[0,1]^{j-1}} C_{i,i+j;\,S_{ij}}\Big(F_{i|S_{ij}}\big(G_{i|S_{ij}}^{\text{PVC}}(a|t_{S_{ij}})\big|t_{S_{ij}}\big), \\
&\qquad\qquad F_{i+j|S_{ij}}\big(G_{i+j|S_{ij}}^{\text{PVC}}(b|t_{S_{ij}})\big|t_{S_{ij}}\big)\Big|t_{S_{ij}}\Big)\mathrm{d}F_{S_{ij}}(t_{S_{ij}}).
\end{aligned}
$$

If $j \geq 3$, $C_{i,i+j;\,S_{ij}}^{\text{PVC}}$ depends on $F_{i|S_{ij}}, F_{i+j|S_{ij}}, C_{i,i+j;\,S_{ij}}$, and $F_{S_{ij}}$, i.e., it depends on $C_{i:i+j}$. Moreover, $C_{i,i+j;\,S_{ij}}^{\text{PVC}}$ also depends on $G_{i|S_{ij}}^{\text{PVC}}$ and $G_{i+j|S_{ij}}^{\text{PVC}}$, which are determined by the regular vine structure. Thus, the corresponding PVCs of different regular vines may be different. In particular, if the simplifying assumption does not hold, higher-order partial copulas of different PVCs which refer

to the same conditional distribution may not be identical. This is different from the partial correlation coefficient or the partial copula which do not depend on the structure of the regular vine.

In general, higher-order partial copulas do not share the simple interpretation of the partial copula because they can not be considered as expected conditional copulas. However, higher-order partial copulas can be more attractive from a practical point of view. The estimation of the partial copula of $C_{i,i+j;S_{ij}}$ requires the estimation of the two j-dimensional conditional cdfs $F_{i|S_{ij}}$ and $F_{i+j|S_{ij}}$ to construct pseudo-observations from the CPITs $(U_{i|S_{ij}}, U_{i+j|S_{ij}})$. As a result, a non-parametric estimation of the partial copula is only sensible if j is very small. In contrast, a higher-order partial copula is the distribution of two PPITs $(U_{i|S_{ij}}^{\mathrm{PVC}}, U_{i+j|S_{ij}}^{\mathrm{PVC}})$ which are made up of only two-dimensional functions (Definition 3.1). Thus, the non-parametric estimation of a higher-order partial copula does not suffer from the curse of dimensionality and is also sensible for large j [27]. But also in a parametric framework the specification of the model is much easier for PPITs than for CPITs. This renders higher-order partial copulas very attractive from a modeling point of view. As we show in Section 6, the PVC is also the probability limit of many estimators of pair-copula constructions and thus of great practical importance.

4. Properties of the partial vine copula and examples

In this section, we analyze to what extent the PVC describes the dependence structure of the data generating copula if the simplifying assumption does not hold. We first investigate whether the bivariate margins of $C_{1:d}^{\mathrm{PVC}}$ match the bivariate margins of $C_{1:d}$ and then take a closer look at conditional independence relations. By construction, the bivariate margins $C_{i,i+1}^{\mathrm{PVC}}$ of the PVC given in Definition 3.1 are identical to the corresponding margins $C_{i,i+1}$, for $i = 1, \ldots, d-1$. That is because the PVC explicitly specifies these $d-1$ margins in the first tree of the vine. The other bivariate margins $C_{i,i+j}^{\mathrm{PVC}}$, where $(i,j) \in \mathcal{I}_2^d$, are implicitly specified and given by

$$C_{i,i+j}^{\mathrm{PVC}}(u_i, u_{i+j}) = \int_{[0,1]^{j-1}} C_{i,i+j;S_{ij}}^{\mathrm{PVC}}(F_{i|S_{ij}}^{\mathrm{PVC}}(u_i|u_{S_{ij}}), F_{i+j}^{\mathrm{PVC}}(u_{i+j}|u_{S_{ij}})) \mathrm{d}C_{S_{ij}}^{\mathrm{PVC}}(u_{S_{ij}}).$$

The relation between the implicitly given bivariate margins of the PVC and the underlying copula are summarized in the following lemma.

Lemma 4.1. (Implicitly specified margins of the PVC)
Let $C_{1:d} \in \mathcal{C}_d \backslash \mathcal{C}_d^{svc}$, $(i,j) \in \mathcal{I}_2^d$, and τ_E and ρ_E denote Kendall's τ and Spearman's ρ of the copula $E \in \mathcal{C}_2$. In general, it does not hold that $C_{i,i+j}^{PVC} = C_{i,i+j}, \rho_{C_{i,i+j}^{PVC}} = \rho_{C_{i,i+j}}$, or $\tau_{C_{i,i+j}^{PVC}} = \tau_{C_{i,i+1}}$.

The next example provides a three-dimensional PVC and illustrates the results of Lemma 4.1. Other examples of PVCs in three dimensions are given in Spanhel and Kurz [35].

Example 4.1.

Let $C^{FGM_2}(\theta)$ denote the bivariate FGM copula

$$C^{FGM_2}(u_1, u_2; \theta) = u_1 u_2 [1 + \theta(1 - u_1)(1 - u_2)], \quad |\theta| \leq 1,$$

$C^A(\gamma)$ denote the following asymmetric version of the FGM copula ([28], Example 3.16)

$$C^A(u_1, u_2; \gamma) = u_1 u_2 [1 + \gamma u_1 (1 - u_1)(1 - u_2)], \quad |\gamma| \leq 1, \qquad (4.1)$$

and $C^\perp$ denote the independence copula $C^\perp(u_1, u_2) = u_1 u_2$. Assume that $C_{12} = C^A(\gamma)$, $C_{23} = C^\perp$, $C_{13;2}(\cdot, \cdot; u_2) = C^{FGM_2}(\cdot, \cdot; 1 - 2u_2)$ for all u_2, so that

$$C_{1:3}(u_{1:3}) = \int_0^{u_2} C^{FGM_2}(\partial_2 C^A(u_1, t_2), u_3; 1 - 2t_2) \mathrm{d}t_2.$$

Elementary computations show that the implicit margin is given by

$$C_{13}(u_1, u_3) = u_1 u_3 [\gamma(u_1 - 3u_1^2 + 2u_1^3)(1 - u_3) + 3]/3,$$

which is a copula with quartic sections in u_1 and square sections in u_3 if $\gamma \neq 0$. The corresponding PVC is

$$C_{1:3}^{\mathrm{PVC}}(u_{1:3}) = \int_0^{u_2} C_{13;2}^{\mathrm{PVC}}(F_{1|2}(u_1|t_2), F_{3|2}(u_3|t_2)) \mathrm{d}t_2$$
$$\overset{C_{13;2}^{\mathrm{PVC}}=C^\perp}{=} u_3 \int_0^{u_2} \partial_2 C^A(u_1, t_2) \mathrm{d}t_2$$

and the implicit margin of $C_{1:3}^{\mathrm{PVC}}$ is

$$C_{13}^{\mathrm{PVC}}(u_1, u_3) = C_{1:3}^{\mathrm{PVC}}(u_1, 1, u_3) = u_1 u_3.$$

Moreover, $\rho_{C_{13}} = -\gamma/1080, \tau_{C_{13}} = -\gamma/135$, but $\rho_{C_{13}^{\mathrm{PVC}}} = \tau_{C_{13}^{\mathrm{PVC}}} = 0$.

Higher-order partial copulas can also be used to construct new measures of conditional dependence. For instance, if $X_{1:d}$ is a random vector with copula $C_{1:d} \in \mathcal{C}_d$, higher-order partial Spearman's ρ and Kendall's τ of X_i and X_{i+j} given $X_{S_{ij}}$ are defined by

$$\tau_{C_{i,i+j;S_{ij}}^{\mathrm{PVC}}} = 4 \int_{[0,1]^2} C_{i,i+j;S_{ij}}^{\mathrm{PVC}}(a, b) \mathrm{d}C_{i,i+j;S_{ij}}^{\mathrm{PVC}}(a, b) - 1,$$

$$\rho_{C_{i,i+j;S_{ij}}^{\mathrm{PVC}}} = 12 \int_{[0,1]^2} C_{i,i+j;S_{ij}}^{\mathrm{PVC}}(a, b) \mathrm{d}a\mathrm{d}b - 3.$$

Note that all dependence measures that are derived from a higher-order partial copula are defined w.r.t. a regular vine structure. They coincide with their conditional analogues if the simplifying assumption holds. A partial correlation coefficient of zero is commonly interpreted as an indication of conditional

independence, although this can be quite misleading if the underlying distribution is not close to a Normal distribution (Spanhel and Kurz [35]). Therefore, one might wonder to what extent higher-order partial copulas can be used to check for conditional independencies. If $C_{i,i+j;S_{ij}}^{\mathrm{PVC}}$ equals the independence copula, we say that X_i and X_{i+j} are (j-th order) partially independent given $X_{S_{ij}}$ and write $X_i \overset{\mathrm{PVC}}{\perp} X_{i+j}|X_{S_{ij}}$. The following theorem establishes that there is in general no relation between conditional independence and higher-order partial independence.

Theorem 4.1. (Conditional independence and j-th order partial independence)
Let $d \geq 4$, $(i,j) \in \mathcal{I}_1^d$, and $C_{1:d} \in \mathcal{C}_d \backslash \mathcal{C}_d^{svc}$ be the copula of $X_{1:d}$. It holds that

$$X_i \perp X_{i+2}|X_{i+1} \Rightarrow X_i \overset{\mathrm{PVC}}{\perp} X_{i+2}|X_{i+1},$$

$$\forall j \geq 3 : X_i \perp X_{i+j}|X_{S_{ij}} \not\Rightarrow X_i \overset{\mathrm{PVC}}{\perp} X_{i+j}|X_{S_{ij}},$$

and

$$\forall j \geq 2 : X_i \perp X_{i+j}|X_{S_{ij}} \not\Leftarrow X_i \overset{\mathrm{PVC}}{\perp} X_{i+j}|X_{S_{ij}}.$$

The next five-dimensional example illustrates higher-order partial copulas, PPITs, and the relation between partial independence and conditional independence.

Example 4.2.
Consider the following exchangeable D-vine copula $C_{1:5}$ which does not satisfy the simplifying assumption:

Tree 1: $\qquad\qquad C_{12} = C_{23} = C_{34} = C_{45} = C^{\perp},$ $\qquad\qquad$ (4.2)

Tree 2: $\quad C_{13;2}(a,b|c) = C_{24;3}(a,b|c) = C_{35;4}(a,b|c)$

$$= C^{FGM_2}(a,b;1-2c), \quad \forall(a,b,c) \in [0,1]^3, \quad (4.3)$$

Tree 3: $\qquad\qquad C_{14;2:3} = C_{25;3:4} = C^{\perp},$ $\qquad\qquad$ (4.4)

Tree 4: $\qquad\qquad C_{15;2:4} = C^{\perp},$ $\qquad\qquad$ (4.5)

where $C_{i,i+j;S_{ij}} = C^{\perp}$ means that $C_{i,i+j;S_{ij}}(a,b|u_{S_{ij}}) = ab$ for all $(a,b,u_{S_{ij}}) \in [0,1]^{j+1}$.

The left panel of Figure 2 illustrates the D-vine copula of the data generating process. All conditional copulas of the vine copula in Example 4.2 correspond to the independence copula except for the second tree. For all $i = 1,2,3$, $(U_i, U_{i+1}, U_{i+2}) \sim C^{FGM_3}(1)$, where $C^{FGM_3}(\theta)$ is the three-dimensional FGM copula with $\theta = 1$.[2] We now investigate the PVC of $C_{1:5}$ which is illustrated in the right panel of Figure 2. Since $C_{1:5}$ and $C_{1:5}^{\mathrm{PVC}}$ are exchangeable copulas, we only report the PPITs $U_{1|2}^{\mathrm{PVC}}, U_{1|2:3}^{\mathrm{PVC}}$ and $U_{1|2:4}^{\mathrm{PVC}}$ in the following lemma.

---

[2] The three-dimensional FGM copula is defined as $C^{FGM_3}(u_{1:3};\theta) = \prod_{i=1}^3 u_i + \theta \prod_{i=1}^3 u_i(1-u_i), |\theta| \leq 1$.

(a) Vine copula in Example 4.2.
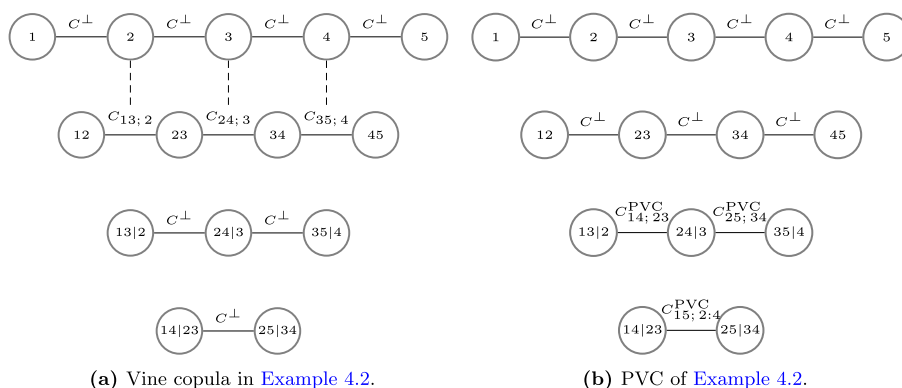
(b) PVC of Example 4.2.

FIG 2. *The non-simplified D-vine copula given in Example 4.2 and its PVC. The influence of conditioning variables on the conditional copulas is indicated by dashed lines.*

Lemma 4.2. (The PVC of Example 4.2)

Let $C_{1:5}$ be defined as in Example 4.2. Then

Tree 1: $\qquad C_{12}^{PVC} = C_{23}^{PVC} = C_{34}^{PVC} = C_{45}^{PVC} = C^{\perp}$,

Tree 2: $\qquad C_{13;2}^{PVC} = C_{24;3}^{PVC} = C_{35;4}^{PVC} = C^{\perp}$,

Tree 3: $\quad C_{14;2:3}^{PVC}(a,b) = C_{25;3:4}^{PVC}(a,b) = C^{FGM_2}(a,b;1/9), \quad \forall (a,b) \in [0,1]^2$,

Tree 4: $\qquad C_{15;2:4}^{PVC} \neq C^{\perp}$,

and

$$U_{1|2}^{PVC} = U_1 = U_{1|2},$$
$$U_{1|2:3}^{PVC} = U_1 \neq U_{1|2:3} = U_1[1 + (1-U_1)(1-2U_2)(1-2U_3)],$$
$$U_{1|2:4}^{PVC} = U_1[1 + \theta(1-U_1)(1-2U_4)] \neq U_{1|2:4} = U_{1|2:3}.$$

Lemma 4.2 demonstrates that j-th order partial copulas may not be independence copulas, although the corresponding conditional copulas are independence copulas. In particular, under the data generating process the edges of the third tree of $C_{1:5}$ are independence copulas. Neglecting the conditional copulas in the second tree and replacing them with first-order partial copulas induces spurious dependencies in the third tree of $C_{1:5}^{\mathrm{PVC}}$. The introduced spurious dependence also carries over to the fourth tree where we have (conditional) independence in fact. Nevertheless, the PVC reproduces the bivariate margins of $C_{1:5}$ pretty well. It can be readily verified that $(C_{13}^{\mathrm{PVC}}, C_{14}^{\mathrm{PVC}}, C_{24}^{\mathrm{PVC}}, C_{25}^{\mathrm{PVC}}, C_{35}^{\mathrm{PVC}}) = (C_{13}, C_{14}, C_{24}, C_{25}, C_{35})$, i.e., except for C_{15}^{PVC}, all bivariate margins of $C_{1:5}^{\mathrm{PVC}}$ match the bivariate margins of $C_{1:5}$ in Example 4.2. Moreover, the mutual information in the third and fourth tree are larger if higher-order partial copulas are used instead of the true conditional copulas. Thus, the spurious dependence

in the third and fourth tree decreases the Kullback-Leibler divergence from $C_{1:5}$ and therefore acts as a countermeasure for the spurious (conditional) independence in the second tree. Lemma 4.2 also reveals that $U_{1|2:4}$ is a function of U_2 and U_3, i.e. the true conditional distribution function $F_{1|2:4}$ depends on u_2 and u_3. In contrast, $F_{1|2:4}^{\mathrm{PVC}}$, the resulting model for $F_{1|2:4}$ which is implied by the PVC, depends only on u_4. That is, the implied conditional distribution function of the PVC depends on the conditioning variable which actually has no effect.

5. Approximations based on the partial vine copula

The specification and estimation of SVCs is commonly based on procedures that asymptotically minimize the Kullback-Leibler divergence (KLD) in a stepwise fashion. For instance, if a parametric vine copula model is used, the step-by-step ML estimator (Hobæk Haff [12, 13]) is often employed in order to select and estimate the parametric pair-copula families of the vine. In this case, one estimates tree after tree and sequentially minimizes the estimated KLD conditional on the estimates from the previous trees. But also the non-parametric methods of Kauermann and Schellhase [20] and Nagler and Czado [27] proceed in a stepwise manner and asymptotically minimize the KLD of each pair-copula separately under appropriate conditions. In this section, we investigate the role of the PVC when it comes to approximating non-simplified vine copulas.

5.1. Tree-by-tree KLD minimization

Let $C_{1:d} \in \mathcal{C}_d$ and $C_{1:d}^{\mathrm{SVC}} \in \mathcal{C}_d^{\mathrm{SVC}}$. The KLD of $C_{1:d}^{\mathrm{SVC}}$ from the true copula $C_{1:d}$ is given by

$$D_{KL}(C_{1:d}||C_{1:d}^{\mathrm{SVC}}) = \mathbb{E}\left[\log \frac{c_{1:d}(U_{1:d})}{c_{1:d}^{\mathrm{SVC}}(U_{1:d})}\right],$$

where the expectation is taken w.r.t. the true distribution $C_{1:d}$. We now decompose the KLD into the Kullback-Leibler divergences related to each of the $d-1$ trees. For this purpose, let $j = 1, \ldots, d-1$ and define

$$\mathbf{T}_j := \left\{ (C_{i,i+j;S_{ij}}^{\mathrm{SVC}})_{i=1,\ldots,d-j} \colon C_{i,i+j;S_{ij}}^{\mathrm{SVC}} \in \mathcal{C}_2 \text{ for } 1 \leq i \leq d-j \right\},$$

so that $\mathbf{T}_{1:j} = \times_{k=1}^{j} \mathbf{T}_k$ represents all possible SVCs up to and including the j-th tree. Let $\mathcal{T}_j \in \mathbf{T}_j, \mathcal{T}_{1:j-1} \in \mathbf{T}_{1:j-1}$. The KLD of the SVC associated with $\mathcal{T}_{1:d-1}$ is given by

$$D_{KL}(C_{1:d}||\mathcal{T}_{1:d-1}) = \sum_{j=1}^{d-1} D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1})), \tag{5.1}$$

where

$$D_{KL}^{(1)}(\mathcal{T}_1(\mathcal{T}_{1:0}))) := D_{KL}^{(1)}(\mathcal{T}_1) := \sum_{i=1}^{d-1} \mathbb{E}\left[\log \frac{c_{i,i+1}(U_i, U_{i+1})}{c_{i,i+1}^{\mathrm{SVC}}(U_i, U_{i+1})}\right]$$

denotes the KLD related to the first tree, and for the remaining trees $j = 2, \ldots, d-1$, the related KLD is

$$D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1})) := \sum_{i=1}^{d-j} \mathbb{E}\left[\log \frac{c_{i,i+j;S_{ij}}(U_{i|S_{ij}}, U_{i+j|S_{ij}}|U_{S_{ij}})}{c_{i,i+j;S_{ij}}^{\mathrm{SVC}}(U_{i|S_{ij}}^{\mathrm{SVC}}, U_{i+j|S_{ij}}^{\mathrm{SVC}})}\right].$$

For instance, if $d = 3$, the KLD can be decomposed into the KLD related to the first tree $D_{KL}^{(1)}$ and to the second tree $D_{KL}^{(2)}$ as follows

$$D_{KL}(C_{1:3}||\mathcal{T}_{1:2}) = D_{KL}(C_{1:3}||(\mathcal{T}_1, \mathcal{T}_2)) = D_{KL}^{(1)}(\mathcal{T}_1) + D_{KL}^{(2)}(\mathcal{T}_2(\mathcal{T}_1))$$

$$= \mathbb{E}\left[\log \frac{c_{12}(U_{1:2})c_{23}(U_{2:3})}{c_{12}^{\mathrm{SVC}}(U_{1:2})c_{23}^{\mathrm{SVC}}(U_{2:3})}\right] + \mathbb{E}\left[\log \frac{c_{13;2}\big(\partial_2 C_{12}(U_{1:2}), \partial_1 C_{23}(U_{2:3})|U_2\big)}{c_{13;2}^{\mathrm{SVC}}\big(\partial_2 C_{12}^{\mathrm{SVC}}(U_{1:2}), \partial_1 C_{23}^{\mathrm{SVC}}(U_{2:3})\big)}\right].$$

Note that the KLD related to tree j depends on the specified copulas in the lower trees because they determine at which values the copulas in tree j are evaluated. The following theorem shows that, if one sequentially minimizes the KLD related to each tree, then the optimal SVC is the PVC.

Theorem 5.1. (Tree-by-tree KLD minimization using the PVC)
Let $C_{1:d} \in \mathcal{C}_d$ be the data generating copula and $\mathcal{T}_j^{PVC} := (C_{i,i+j;S_{ij}}^{PVC})_{i=1,\ldots,d-j}$, so that $\mathcal{T}_{1:j}^{PVC} := \times_{k=1}^j \mathcal{T}_k^{PVC}$ collects all copulas of the PVC up to and including the j-th tree. It holds that

$$\forall j = 1, \ldots, d-1: \quad \underset{\mathcal{T}_j \in \mathbf{T}_j}{\arg\min} \; D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1}^{PVC})) = \mathcal{T}_j^{PVC}. \tag{5.2}$$

According to Theorem 5.1, if the true copulas are specified in the first tree, one should choose the first-order partial copulas in the second tree. Moreover, the second-order partial copulas should then be specified in the third tree and so on to minimize the KLD tree-by-tree. Theorem 5.1 also remains true if we replace $\mathcal{C}_2$ in the definition of $\mathbf{T}_j$ by the space of absolutely continuous bivariate cdfs. The PVC ensures that random variables in higher trees are uniformly distributed since the resulting random variables in higher trees are PPITs. If one uses a different approximation, such as the one used by Hobæk Haff, Aas and Frigessi [14] and Stöber, Joe and Czado [37], then the random variables in higher trees are not necessarily uniformly distributed and pseudo-copulas (Fermanian and Wegkamp [7]) can be used to further minimize the KLD.

5.2. Global KLD minimization

The previous sequential minimization neglects that the KLD related to a tree depends on the copulas that are specified in the former trees. For instance, if $d = 3$, the KLD of the first tree $D_{KL}^{(1)}(\mathcal{T}_1)$ is minimized over the copulas $(C_{12}^{\mathrm{SVC}}, C_{23}^{\mathrm{SVC}})$ in the first tree $\mathcal{T}_1$. However, the effect of the chosen copulas in the first tree $\mathcal{T}_1$ on the KLD related to the second tree $D_{KL}^{(2)}(\mathcal{T}_2(\mathcal{T}_1))$ is not taken into

account. Therefore, we now analyze whether the PVC also globally minimizes the KLD. Note that specifying the wrong margins in the first tree $\mathcal{T}_1$, e.g., $(C_{12}^{\mathrm{svc}}, C_{23}^{\mathrm{svc}}) \neq (C_{12}, C_{23})$, increases $D_{KL}^{(1)}(\mathcal{T}_1)$ in any case. Thus, without any further investigation, it is absolutely indeterminate whether the definite increase in $D_{KL}^{(1)}(\mathcal{T}_1)$ can be overcompensated by a possible decrease in $D_{KL}^{(2)}(\mathcal{T}_2(\mathcal{T}_1))$ if another approximation is chosen. The next theorem shows that the PVC is in general not the global minimizer of the KLD.

Theorem 5.2. (Global KLD minimization if $C_{1:d} \in \mathcal{C}_d^{\mathrm{svc}}$ or $C_{1:d} \in \mathcal{C}_d \backslash \mathcal{C}_d^{\mathrm{svc}}$)
If $C_{1:d} \in \mathcal{C}_d^{SVC}$, i.e., the simplifying assumption holds for $C_{1:d}$, then

$$\underset{C_{1:d}^{SVC} \in \mathcal{C}_d^{SVC}}{\arg\min} \ D_{KL}(C_{1:d} || C_{1:d}^{SVC}) = C_{1:d}^{PVC}. \tag{5.3}$$

If the simplifying assumption does not hold for $C_{1:d}$, then $C_{1:d}^{PVC}$ might not be a global minimum. That is, $\exists C_{1:d} \in \mathcal{C}_d \backslash \mathcal{C}_d^{SVC}$ such that

$$\underset{C_{1:d}^{SVC} \in \mathcal{C}_d^{SVC}}{\arg\min} \ D_{KL}(C_{1:d} || C_{1:d}^{SVC}) \neq C_{1:d}^{PVC}, \tag{5.4}$$

and $\forall (\mathcal{T}_2, \ldots, \mathcal{T}_{d-1}) \in \times_{k=2}^{d-1} \mathbf{T}_k$

$$\underset{\tilde{\mathcal{T}}_{1:d-1} \in \mathbf{T}_{1:d-1}}{\arg\min} \ D_{KL}(C_{1:d} || \tilde{\mathcal{T}}_{1:d-1}) \neq (\mathcal{T}_1^{PVC}, \mathcal{T}_2, \ldots, \mathcal{T}_{d-1}). \tag{5.5}$$

Theorem 5.2 states that, if the simplifying assumption does not hold, the KLD may not be minimized by choosing the true copulas in the first tree, first-order partial copulas in the second tree and higher-order partial copulas in the remaining trees (see (5.4)). It follows that, if the objective is the minimization of the KLD, it may not be optimal to specify the true copulas in the first tree, no matter what bivariate copulas are specified in the other trees (see (5.5)). This rather puzzling result can be explained by the fact that, if the simplifying assumption does not hold, then the approximation error of the implicitly modeled bivariate margins is not minimized (see Lemma 4.1). For instance, if $d = 3$, a departure from the true copulas (C_{12}, C_{23}) in the first tree increases the KLD related to the first tree, but it can decrease the KLD of the implicitly modeled margin C_{13}^{svc} from C_{13}. As a result, the increase in $D_{KL}^{(1)}$ can be overcompensated by a larger decrease in $D_{KL}^{(2)}$, so that the KLD can be decreased.

Theorem 5.2 does not imply that the PVC never minimizes the KLD from the true copula. For instance, if $d = 3$ and if $C_{13;2}^{\mathrm{PVC}} = C^{\perp}$, then $D_{KL}(C_{1:3} || C_{1:3}^{\mathrm{PVC}})$ is an extremum, which directly follows from equation (5.2) since

$$\underset{\mathcal{T}_1 \in \mathbf{T}_1}{\arg\min} D_{KL}(C_{1:3} || (\mathcal{T}_1, (C^{\perp}))) = \underset{\mathcal{T}_1 \in \mathbf{T}_1}{\arg\min} D_{KL}^{(1)}(\mathcal{T}_1).$$

It is an open problem whether and when the PVC can be the global minimizer of the KLD. Unfortunately, the SVC approximation that globally minimizes the

KLD is not tractable. However, if the SVC approximation that minimizes the KLD does not specify the true copulas in the first tree, the random variables in higher trees are not CPITs. Thus, it is not guaranteed that these random variables are uniformly distributed and we could further decrease the KLD by assigning pseudo-copulas (Fermanian and Wegkamp [7]) to the edges in the higher trees. It can be easily shown that the resulting best approximation is then a pseudo-copula. Consequently, the best approximation satisfying the simplifying assumption is in general not a SVC but a simplified vine pseudo-copula if one considers the space of regular vines where each edge corresponds to a bivariate cdf.

While the PVC may not be the best approximation in the space of SVCs, it is often the best feasible SVC approximation in practical applications. That is because the stepwise specification and estimation of a SVC is feasible for (very) large dimensions which is not true for a joint specification and estimation. For instance, assume all pair-copula families of a parametric vine copula are chosen simultaneously and the selection is done by means of information criteria. In this case, we have to estimate $K^{d(d-1)/2}$ different models, where d is the dimension and K the number of possible pair-copula families that can be assigned to each edge. On the contrary, a stepwise procedure only requires the estimation of $Kd(d-1)/2$ models. To illustrate the computational burden, consider the R-package VineCopula [34] where $K = 40$. For this number of pair-copula families, a joint specification requires the estimation of 64,000 ($d = 3$) or more than four billion ($d = 4$) models whereas only 120 ($d = 3$) or 240 ($d = 4$) models are needed for a stepwise specification. For many non-parametric estimation approaches (kernels [27], empirical distributions [15]), only the sequential estimation of a SVC is possible. The only exception is the spline-based approach of Kauermann and Schellhase [20]. However, due to the large number of parameters and the resulting computational burden, a joint estimation is only feasible for $d \leq 5$ [19].

6. Convergence to the partial vine copula

If the data generating process satisfies the simplifying assumption, consistent stepwise procedures for the specification and estimation of parametric and non-parametric SVC models asymptotically minimize the KLD from the true copula. Theorem 5.1 implies that this is not true in general if the data generating process does not satisfy the simplifying assumption. An implication of this result for the application of SVCs is pointed out in the next corollary.

6.1. Step-by-step and joint ML estimates: Theory

Corollary 6.1.
Denote the sample size by N. Let $C_{1:d} \in \mathcal{C}_d$ be the data generating copula and $C_{1:d}^{svc}(\theta) \in \mathcal{C}_d^{svc}$, $\theta \in \Theta$, be a parametric SVC so that $\exists_1 \theta_{PVC} \in \Theta : C_{1:d}^{svc}(\theta_{PVC}) = C_{1:d}^{PVC}$. The pseudo-true parameters which minimize the KLD from the true distribution are assumed to exist (see White [39] for sufficient conditions) and

denoted by

$$\theta^\star = \underset{\theta \in \Theta}{\arg\min}\, D_{KL}(C_{1:d}||C_{1:d}^{svc}(\theta)).$$

Let $\hat{\theta}_S$ denote the (semi-parametric) step-by-step ML estimator and $\hat{\theta}_J$ denote the (semi-parametric) joint ML estimator defined in Hobæk Haff [12, 13]. Under regularity conditions (e.g., Condition 1 and Condition 2 in [36]) and for $N \to \infty$, it holds that:

(i) $\hat{\theta}^S \overset{p}{\to} \theta^{PVC}$.

(ii) $\hat{\theta}^J \overset{p}{\to} \theta^\star$.

(iii) $\exists C_{1:d} \in \mathcal{C}_d \backslash \mathcal{C}_d^{svc}$ such that $\hat{\theta}^S \overset{p}{\nrightarrow} \theta^\star$.

Corollary 6.1 shows that the step-by-step and joint ML estimator may not converge to the same limit (in probability) if the simplifying assumption does not hold for the data generating vine copula. For this reason, we investigate in the following the difference between the step-by-step and joint ML estimator in finite samples. Note that the convergence of kernel-density estimators to the PVC has recently been established by Nagler and Czado [27]. However, in this case, only a sequential estimation of a SVC is possible and thus the best feasible approximation in the space of SVCs is given by the PVC.

6.2. Step-by-step and joint ML estimates: Simulation study

We compare the step-by-step and the joint ML estimator under the assumption that the pair-copula families of the PVC are specified for the parametric vine copula model. For this purpose, we simulate data from two three-dimensional copulas $C_{1:3}$ with sample sizes $N = 500, 2500, 25000$, perform a step-by-step and joint ML estimation, and repeat this 1000 times. For ease of exposition and because the qualitative results are not different, we consider copulas where $C_{12} = C_{23}$ and only present the estimates for $(\theta_{12}, \theta_{13;2})$.

Example 6.1. (PVC of the Frank copula)
Let $C^{\mathrm{Fr}}(\theta)$ denote the bivariate Frank copula with dependence parameter θ and $C^{\mathrm{P\text{-}Fr}}(\theta)$ be the partial Frank copula [35] with dependence parameter θ. Let $C_{1:3}$ be the true copula with $(C_{12}, C_{23}, C_{13;2}) = (C^{\mathrm{Fr}}(5.74), C^{\mathrm{Fr}}(5.74), C^{\mathrm{P\text{-}Fr}}(5.74))$, i.e., $C_{1:3} = C_{1:3}^{\mathrm{PVC}}$, and $C_{1:3}^{\mathrm{svc}}(\theta) = (C^{\mathrm{Fr}}(\theta_{12}), C^{\mathrm{Fr}}(\theta_{23}), C^{\mathrm{P\text{-}Fr}}(\theta_{13;2}))$ be the parametric SVC that is fitted to data generated from $C_{1:3}$.

Example 6.1 presents a data generating process which satisfies the simplifying assumption, implying $\theta^{\mathrm{PVC}} = \theta^\star$. It is the PVC of the three-dimensional Frank copula with Kendall's τ approximately equal to 0.5. Figure 3 shows the corresponding box plots of joint and step-by-step ML estimates and their difference. The left panel confirms the results of Hobæk Haff [12, 13]. Although the joint ML estimator is more efficient, the loss in efficiency for the step-by-step ML estimator is negligible and both estimators converge to the true parameter
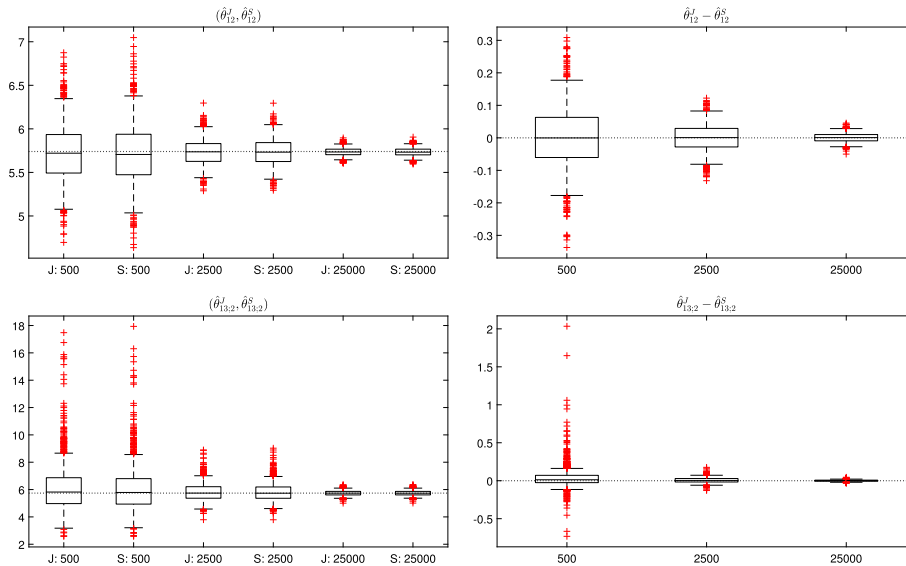
FIG 3. *Box plots of joint (J) and sequential (S) ML estimates and their difference for sample sizes $N = 500, 2500, 25000$, if the data is generated from $C_{1:3}$ in Example 6.1 and the pair-copula families of the SVC are given by the corresponding PVC. The dotted line indicates the pseudo-true parameter and zero, respectively. The end of the whiskers is 0.953 times the inter-quartile range, corresponding to approximately 95% coverage if the data is generated by a normal distribution.*

value. Moreover, the right panel of Figure 3 shows that the difference between joint and step-by-step ML estimates is never statistically significant at a 5% level. Since the computational time for a step-by-step ML estimation is much lower than for a joint ML estimation [12], the step-by-step ML estimator is very attractive for estimating high-dimensional vine copulas that satisfy the simplifying assumption. Moreover, the step-by-step ML estimator is then inherently suited for selecting the pair-copula families in a stepwise manner. However, if the simplifying assumption does not hold for the data generating vine copula, the step-by-step and joint ML estimator can converge to different limits (Corollary 6.1), as the next example demonstrates.

Example 6.2. (Frank copula)
Let $C_{1:3}$ be the Frank copula with dependence parameter $\theta = 5.74$, i.e., $C_{1:3} \neq C_{1:3}^{\text{PVC}}$, and $C_{1:3}^{\text{SVC}}(\theta) = (C^{\text{Fr}}(\theta_{12}), C^{\text{Fr}}(\theta_{23}), C^{\text{P-Fr}}(\theta_{13;2}))$ be the parametric SVC that is fitted to data generated from $C_{1:3}$.

Example 6.2 is identical to Example 6.1, with the only difference that the conditional copula is varying in such a way that the resulting three-dimensional copula is a Frank copula. Although the Frank copula does not satisfy the simplifying assumption, it is pretty close to a copula for which the simplifying assumption holds, because the variation in the conditional copula is strongly limited for many Archimedean copulas (Mesfioui and Quessy [26]). Nevertheless, the
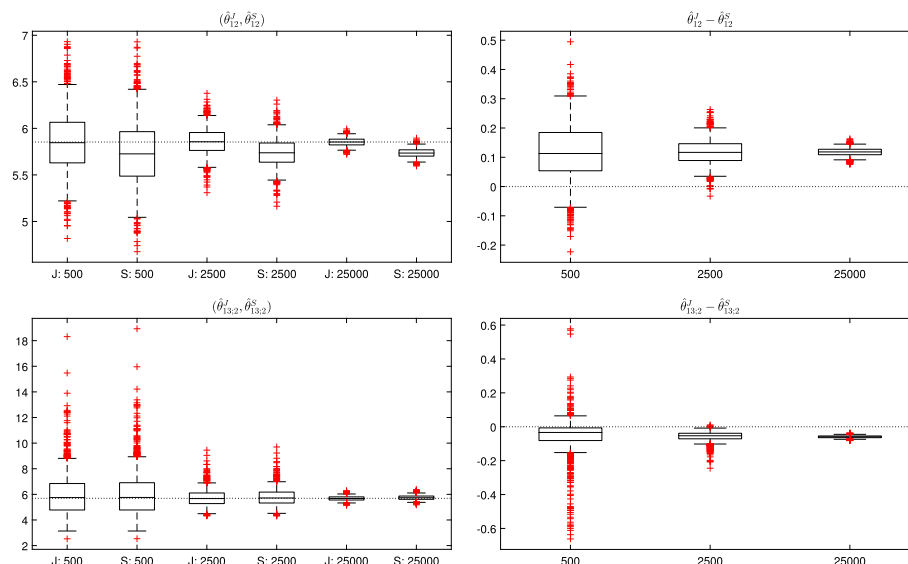
FIG 4. *Box plots of joint (J) and sequential (S) ML estimates and their difference for sample sizes $N = 500, 2500, 25000$, if the data is generated from $C_{1:3}$ in Example 6.2 and the pair-copula families of the SVC are given by the corresponding PVC. The dotted line indicates the pseudo-true parameter and zero, respectively. The end of the whiskers is 0.953 times the inter-quartile range, corresponding to approximately 95% coverage if the data is generated by a normal distribution.*

right panel of Figure 4 shows that the step-by-step and joint ML estimates for θ_{12} are significantly different at the 5% level if the sample size is 2500 observations. The difference between step-by-step and joint ML estimates for $\theta_{13;2}$ is less pronounced, but also highly significant for sample sizes with 2500 observations or more. In Example 6.2 the step-by-step ML estimator is not a consistent estimator of the SVC model that minimizes the KLD from the underlying copula. In contrast, the joint ML estimator is still a consistent minimizer. A third example where the distance between the data generating copula and the PVC and thus the difference between the step-by-step and joint ML estimates is more pronounced is given in Appendix A.9.

7. Conclusion

We introduced the partial vine copula (PVC) which is a particular simplified vine copula (SVC) that coincides with the data generating copula if the simplifying assumption holds. The PVC can be regarded as a generalization of the partial correlation matrix where partial correlations are replaced by j-th order partial copulas. While a higher-order partial copula of the PVC is related to the partial copula, it does not suffer from the curse of dimensionality and can be estimated for high-dimensional data [27]. We analyzed to what extent the

dependence structure of the underlying distribution is reproduced by the PVC. In particular, we showed that a pair of random variables may be considered as conditionally (in)dependent according to the PVC although this is not the case for the data generating process.

We also revealed the importance of the PVC for the modeling of high-dimensional distributions by means of SVCs. Up to now, the estimation of vine copulas has almost always been based on the assumption that the data generating process satisfies the simplifying assumption. Moreover, the implications that follow if the simplifying assumption is not true have not been investigated. We showed that the PVC is the SVC approximation that minimizes the Kullback-Leibler divergence in a stepwise fashion. Since almost all estimators of SVCs proceed sequentially, it follows that many estimators of SVCs converge to the PVC regardless of whether the simplifying assumption holds. However, we also proved that the PVC may not minimize the Kullback-Leibler divergence from the true copula and thus may not be the best SVC approximation in theory. Nevertheless, due to the prohibitive computational burden or simply because only a stepwise model specification and estimation is possible, the PVC is often the best feasible SVC approximation in practice.

The analysis in this paper showed the relative optimality of the PVC when it comes to approximating multivariate copulas by SVCs. Obviously, it is easy to construct (theoretical) examples where the PVC does not provide a good approximation in absolute terms. But such examples do not provide any information about the appropriateness of the simplifying assumption in practice. To investigate whether the PVC is a good approximation in applications, one can use Lemma 3.1 to develop tests for the simplifying assumption, see Kurz and Spanhel [24]. Moreover, even in cases where the simplifying assumption is strongly violated, an estimator of the PVC can yield an approximation that is superior to competing approaches. Recently, it has been demonstrated in Nagler and Czado [27] that the PVC can be used to obtain a constrained kernel-density estimator that outperforms unconstrained kernel-density estimators.

Appendix

A.1. Proof of Lemma 3.1

$U_{k|S_{ij}}^{\text{PVC}} \overset{a.s.}{=} U_{k|S_{ij}} \Rightarrow U_{k|S_{ij}}^{\text{PVC}} \perp U_{S_{ij}}$ is true because $U_{k|S_{ij}}$ is a CPIT. For the converse, let $A := \times_{k=i+1}^{i+j-1}[0, u_k]$ and consider

$$\mathbb{P}(U_{k|S_{ij}}^{\text{PVC}} \leq a, U_{S_{ij}} \leq u_{S_{ij}}) = \int_A F_{k|S_{ij}}\big((F_{k|S_{ij}}^{\text{PVC}})^{-1}(a|t_{S_{ij}})|t_{S_{ij}}\big)\mathrm{d}C_{S_{ij}}(t_{S_{ij}}).$$

(A.1)

Since $U_{k|S_{ij}}^{\text{PVC}} \sim U(0,1)$ it follows that if $U_{k|S_{ij}}^{\text{PVC}} \perp U_{S_{ij}}$ then $\mathbb{P}(U_{k|S_{ij}}^{\text{PVC}} \leq a, U_{S_{ij}} \leq u_{S_{ij}}) = aC_{S_{ij}}(u_{S_{ij}})$ for all $(a, u_{S_{ij}}) \in [0,1]^j$. This implies that

$$\mathbb{P}(U_{k|S_{ij}}^{\text{PVC}} \leq a, U_{S_{ij}} \leq u_{S_{ij}}) = \int_A a \, \mathrm{d}C_{S_{ij}}(t_{S_{ij}})$$

equals the right hand side of (A.1) for all $(a, u_{S_{ij}}) \in [0, 1]^j$. It follows that the integrands must be identical (a.e.) as well and $F_{k|S_{ij}}((F_{k|S_{ij}}^{\mathrm{PVC}})^{-1}(a|t_{S_{ij}})|t_{S_{ij}}) = a$ for all $a \in [0, 1]$ and almost every $t_{S_{ij}} \in [0, 1]^{j-1}$. Thus $F_{k|S_{ij}} = F_{k|S_{ij}}^{\mathrm{PVC}}$ (a.e.). Since the CPIT $U_{k|S_{ij}} = F_{k|S_{ij}}(U_k|U_{S_{ij}})$ and the PPIT $U_{k|S_{ij}}^{\mathrm{PVC}} = F_{k|S_{ij}}^{\mathrm{PVC}}(U_k|U_{S_{ij}})$ are defined as (a.e.) identical transformations of the random variables $(U_k, U_{S_{ij}})$ it follows that $U_{k|S_{ij}}^{\mathrm{PVC}} \overset{a.s}{=} U_{k|S_{ij}}$.

A.2. Proof of Lemma 4.1

Let $C_{1:3}^\star \in \mathcal{C}_3^{\mathrm{SVC}}$ be the SVC given in Example 4.1. We define $C_{1:d}$ as follows. Let $C_{1,d-1;2:d-2} = C_{12}^\star, C_{2,d;3:d-1} = C_{23}^\star, C_{1,d;2:d-1} = D_{1,3;2}^\star$, where $D_{1,3;2}^\star$ is the corresponding conditional copula in Example 4.1 and $C_{i,i+j;S_{ij}} = E_{k,l} \in \mathcal{C}_2, (k, l) \in \mathcal{I}_1^d$ means that $C_{i,i+j;S_{ij}}(a, b|u_{S_{ij}}) = E_{k,l}(a, b)$ for all $(a, b, u_{S_{ij}}) \in [0, 1]^{j+1}$. Moreover, let $C_{i,i+j;S_{ij}} = C^\perp$ for $(i, j) \in \mathcal{I}_1^d \setminus \{(1, d-2), (2, d-2), (1, d-1)\}$. The conclusion now follows from Example 4.1.

A.3. Proof of Theorem 4.1

W.l.o.g. assume that the margins of $X_{1:d}$ are uniform. Let $C^{FGM_3}(u_{1:3}; \theta) = \prod_{i=1}^3 u_i + \theta \prod_{i=1}^3 u_i(1 - u_i), |\theta| \leq 1$, be the three-dimensional FGM copula, $d \geq 4$, and $(i, j) \in \mathcal{I}_1^d$. It is obvious that $C_{i,i+2;i+1} = C^\perp \Rightarrow C_{i,i+2;i+1}^{\mathrm{PVC}} = C^\perp$ is true. Let $J \in \{2, \ldots, d-2\}$ be fixed. Assume that $C_{1:d}$ has the following D-vine copula representation of the non-simplified form

$$C_{1,1+J;2:J} = \partial_3 C^{FGM_3}(u_1, u_{1+J}, u_2; 1)$$
$$C_{2,2+J;3:J+1} = \partial_3 C^{FGM_3}(u_2, u_{2+J}, u_{1+J}; 1)$$

and $C_{i,i+j;S_{i,j}} = C^\perp$ for all other $(i, j) \in \mathcal{I}_1^d$. Using the same arguments as in the proof of Lemma 4.2 we obtain

$$C_{i,i+J;i+1:J-1}^{\mathrm{PVC}} = C^\perp, \quad i = 1, 2,$$
$$C_{1,2+J;2:J+1}^{\mathrm{PVC}} = C^{FGM_2}(1/9).$$

This proves that $C_{i,i+2;i+1} = C^\perp \Leftarrow C_{i,i+2;i+1}^{\mathrm{PVC}} = C^\perp$ is not true in general and that, for $j \geq 3$, neither the statement $C_{i,i+j;S_{ij}} = C^\perp \Rightarrow C_{i,i+j;S_{ij}}^{\mathrm{PVC}} = C^\perp$ nor the statement $C_{i,i+j;S_{ij}} = C^\perp \Leftarrow C_{i,i+j;S_{ij}}^{\mathrm{PVC}} = C^\perp$ is true in general.

A.4. Proof of Lemma 4.2

We show a more general result and set

$$C_{i,i+2;i+1}(u_i, u_{i+2}|u_{i+1}) = C^{FGM_2}(u_i, u_{i+2}; g(u_{i+1}))$$

in (4.3) where $g\colon [0,1] \to [-1,1]$ is a non-constant measurable function such that

$$\forall u \in [0.5, 1]\colon g(0.5 + u) = -g(0.5 - u). \tag{A.2}$$

For $i = 1, 2, 3$, the copula in the second tree of the PVC is given by

$$
\begin{aligned}
C^{\mathrm{PVC}}_{i,i+2;i+1}(a,b) &= \mathbb{P}(U_{i|i+1} \leq a, U_{i+2|i+1} \leq b) \\
&= \int_{[0,1]} C_{i,i+2;i+1}(a,b|u_{i+1})\mathrm{d}u_{i+1} \\
&\overset{(4.3)}{=} ab\big(1 + (1-a)(1-b)\int_{[0,1]} g(u_{i+1})\mathrm{d}u_{i+1}\big) \overset{(\mathrm{A.2})}{=} ab, \quad (\mathrm{A.3})
\end{aligned}
$$

which is the independence copula. For $i = 1, 2, k = i, i+3$, the true CPIT of U_k w.r.t. $U_{i+1:i+2}$ is a function of $U_{i+1:i+2}$ because

$$U_{i|i+1:i+2} = U_i[1 + g(U_{i+1})(1 - U_i)(1 - 2U_{i+2})], \tag{A.4}$$

$$U_{i+3|i+1:i+2} = U_{i+3}[1 + g(U_{i+2})(1 - U_{i+3})(1 - 2U_{i+1})]. \tag{A.5}$$

However, for $i = 1, 2, k = i, i+3$, the PPIT of U_k w.r.t. $U_{i+1:i+2}$ is not a function of $U_{i+1:i+2}$ because

$$
\begin{aligned}
U^{\mathrm{PVC}}_{i|i+1:i+2} &= F^{\mathrm{PVC}}_{i|i+1:i+2}(U_i|U_{i+1:i+2}) = F_{U_{i|i+1}|U_{i+2|i+1}}(U_{i|i+1}|U_{i+2|i+1}) \\
&= \partial_2 C^{\mathrm{PVC}}_{i,i+2;i+1}(U_{i|i+1}, U_{i+2|i+1}) \overset{(\mathrm{A.3})}{=} U_{i|i+1} \overset{(4.2)}{=} U_i, \quad (\mathrm{A.6})
\end{aligned}
$$

and, by symmetry,

$$U^{\mathrm{PVC}}_{i+3|i+1:i+2} = U_{i+3}. \tag{A.7}$$

For $i = 1, 2$, the joint distribution of these first-order PPITs is a copula in the third tree of the PVC which is given by

$$
\begin{aligned}
C^{\mathrm{PVC}}_{i,i+3;i+1:i+2}(a,b) &= \mathbb{P}(U^{\mathrm{PVC}}_{i|i+1:i+2} \leq a, U^{\mathrm{PVC}}_{i+3|i+1:i+2} \leq b) \\
&\overset{(\mathrm{A.6}),(\mathrm{A.7})}{=} \mathbb{P}(U_i \leq a, U_{i+3} \leq b) = C_{i,i+3}(a,b) \quad (\mathrm{A.8}) \\
&\overset{(4.4)}{=} \int_{[0,1]^2} F_{i|i+1:i+2}(a|u_{i+1:i+2})F_{i+3|i+1:i+2}(b|u_{i+1:i+2})\mathrm{d}u_{i+1:i+2} \\
&\overset{(\mathrm{A.4}),(\mathrm{A.5})}{=} ab\bigg[1 + (1-a)(1-b)\int_{[0,1]} g(u_{i+1})(1 - 2u_{i+1})\mathrm{d}u_{i+1} \\
&\qquad \times \int_{[0,1]} g(u_{i+2})(1 - 2u_{i+2})\mathrm{d}u_{i+2}\bigg] \\
&= ab[1 + \theta(1-a)(1-b)] = C^{FGM_2}(\theta),
\end{aligned}
$$

where $\theta := 4(\int_{[0,1]} u g(u) \mathrm{d}u)^2 > 0$, by the properties of g. Thus, a copula in the third tree of the PVC is a bivariate FGM copula whereas the true conditional copula is the independence copula.

The CPITs of U_1 or U_5 w.r.t. $U_{2:4}$ are given by

$$U_{1|2:4} = \quad F_{1|2:4}(U_1|U_{2:4}) = \partial_2 C_{14;2:3}(U_{1|2:3}, U_{4|2:3}|U_{2:3}) \overset{(4.4)}{=} U_{1|2:3}$$

$$\overset{(A.4)}{=} U_1[1 + g(U_2)(1 - U_1)(1 - 2U_3)], \tag{A.9}$$

$$U_{5|2:4} = \quad U_5[1 + g(U_4)(1 - U_5)(1 - 2U_3)], \tag{A.10}$$

whereas the corresponding second-order PPITs are given by

$$U_{1|2:4}^{\mathrm{PVC}} = F_{1|2:4}^{\mathrm{PVC}}(U_1|U_{2:4}) = F_{U_{1|2:3}^{\mathrm{PVC}}|U_{4|2:3}^{\mathrm{PVC}}}(U_{1|2:3}^{\mathrm{PVC}}|U_{4|2:3}^{\mathrm{PVC}}) \overset{(A.6),(A.7)}{=} F_{U_1|U_4}(U_1|U_4)$$

$$= U_{1|4} = \partial_2 C_{14}(U_1, U_4) \overset{(A.8)}{=} U_1[1 + \theta(1 - U_1)(1 - 2U_4)], \tag{A.11}$$

$$U_{5|2:4}^{\mathrm{PVC}} = U_5[1 + \theta(1 - U_5)(1 - 2U_2)]. \tag{A.12}$$

For the copula in the fourth tree of the PVC it holds

$$C_{15;2:4}^{\mathrm{PVC}}(a, b) = \mathbb{P}(U_{1|2:4}^{\mathrm{PVC}} \le a, U_{5|2:4}^{\mathrm{PVC}} \le b) \overset{(A.11),(A.12)}{=} \mathbb{P}(U_{1|4} \le a, U_{5|2} \le b)$$

$$= \mathbb{P}(U_1 \le F_{1|4}^{-1}(a|U_4), U_5 \le F_{5|2}^{-1}(b|U_2))$$

$$= \int_{[0,1]^3} F_{15|2:4}(F_{1|4}^{-1}(a|u_4), F_{5|2}^{-1}(b|u_2)|u_{2:4}) c_{2:4}(u_{2:4}) \mathrm{d}u_{2:4}$$

$$= \int_{[0,1]^3} C_{15;2:4}(F_{1|2:4}(F_{1|4}^{-1}(a|u_4)|u_{2:4}), F_{5|2:4}(F_{5|2}^{-1}(b|u_2)|u_{2:4})|u_{2:4})$$

$$\times c_{2:4}(u_{2:4}) \mathrm{d}u_{2:4}$$

$$\overset{(4.5)}{=} \int_{[0,1]^3} F_{1|2:4}(F_{1|4}^{-1}(a|u_4)|u_{2:4}) F_{5|2:4}(F_{5|2}^{-1}(b|u_2)|u_{2:4}) c_{2:4}(u_{2:4}) \mathrm{d}u_{2:4}$$

$$\overset{(A.9),(A.10)}{=} \int_{[0,1]^3} F_{1|2:3}(F_{1|4}^{-1}(a|u_4)|u_{2:3}) F_{5|3:4}(F_{5|2}^{-1}(b|u_2)|u_{3:4}) c_{2:4}(u_{2:4}) \mathrm{d}u_{2:4}$$

$$\overset{(A.4),(A.5)}{=} \int_{[0,1]^3} F_{1|4}^{-1}(a|u_4)[1 + g(u_2)(1 - F_{1|4}^{-1}(a|u_4))(1 - 2u_3)] F_{5|2}^{-1}(b|u_2)$$

$$\times [1 + g(u_4)(1 - F_{5|2}^{-1}(b|u_2))(1 - 2u_3)][1 + g(u_3)(1 - 2u_2)(1 - 2u_4)] \mathrm{d}u_{2:4}$$

$$= \int_{[0,1]^2} F_{1|4}^{-1}(a|u_4) F_{5|2}^{-1}(b|u_2)$$

$$\times \left[1 + \int_{[0,1]} (1 - 2u_3)^2 \mathrm{d}u_3 (1 - F_{1|4}^{-1}(a|u_4))(1 - F_{5|2}^{-1}(b|u_2)) g(u_4) g(u_2) \right.$$

$$+ \int_{[0,1]} (1 - 2u_3)g(u_3)\mathrm{d}u_3(1 - F_{5|2}^{-1}(b|u_2))(1 - 2u_2)(1 - 2u_4)g(u_4)$$

$$+ \int_{[0,1]} (1 - 2u_3)g(u_3)\mathrm{d}u_3(1 - F_{1|4}^{-1}(a|u_4))(1 - 2u_4)(1 - 2u_2)g(u_2) \bigg] \mathrm{d}u_2\mathrm{d}u_4,$$

where we used that $\int_{[0,1]}(1 - 2u_3)\mathrm{d}u_3 = 0$, $\int_{[0,1]} g(u_3)\mathrm{d}u_3 \overset{(A.2)}{=} 0$ and $\int_{[0,1]}(1 - 2u_3)^2 g(u_3)\mathrm{d}u_3 \overset{(A.2)}{=} 0$. By setting $\gamma := -2\int_{[0,1]} ug(u)\mathrm{d}u$ we can write the copula function as

$$C_{15;\,2:4}^{\mathrm{PVC}}(a,b) = \int_{[0,1]^2} F_{1|4}^{-1}(a|u_4)F_{5|2}^{-1}(b|u_2)$$

$$\times \left[1 + \frac{1}{3}(1 - F_{1|4}^{-1}(a|u_4))(1 - F_{5|2}^{-1}(b|u_2))g(u_4)g(u_2) \right.$$

$$+ \gamma(1 - F_{5|2}^{-1}(b|u_2))(1 - 2u_2)(1 - 2u_4)g(u_4)$$

$$\left. + \gamma(1 - F_{1|4}^{-1}(a|u_4))(1 - 2u_4)(1 - 2u_2)g(u_2) \right] \mathrm{d}u_2\mathrm{d}u_4$$

$$= \left(\int_0^1 F_{1|4}^{-1}(a|u_4)\mathrm{d}u_4 \right) \left(\int_0^1 F_{5|2}^{-1}(b|u_2)\mathrm{d}u_2 \right)$$

$$+ \gamma \left(\int_0^1 (1 - 2u_2)(F_{5|2}^{-1}(b|u_2) - (F_{5|2}^{-1}(b|u_2))^2)\mathrm{d}u_2 \right)$$

$$\times \left(\int_0^1 (1 - 2u_4)g(u_4)F_{1|4}^{-1}(a|u_4)\mathrm{d}u_4 \right)$$

$$+ \gamma \left(\int_0^1 (1 - 2u_4)(F_{1|4}^{-1}(a|u_4) - (F_{1|4}^{-1}(a|u_4))^2)\mathrm{d}u_4 \right)$$

$$\times \left(\int_0^1 (1 - 2u_2)g(u_2)F_{5|2}^{-1}(b|u_2)\mathrm{d}u_2 \right)$$

$$+ \frac{1}{3} \left(\int_0^1 g(u_4)(F_{1|4}^{-1}(a|u_4) - (F_{1|4}^{-1}(a|u_4))^2)\mathrm{d}u_4 \right)$$

$$\times \left(\int_0^1 g(u_2)(F_{5|2}^{-1}(b|u_2) - (F_{5|2}^{-1}(b|u_2))^2)\mathrm{d}u_2 \right)$$

If $(U, V) \sim C^{FGM_2}(\theta)$, the quantile function is given by (cf. Remillard [32])

$$F_{U|V}^{-1}(u|v) = \frac{1 + h(v) - \sqrt{(1 + h(v))^2 - 4h(v)u}}{2h(v)},$$

with $h(v) := \theta(1 - 2v)$, which implies

$$\frac{\partial}{\partial u} F_{U|V}^{-1}(u|v) = \frac{1}{\sqrt{(1 + h(v))^2 - 4h(v)u}} =: G(u, v), \qquad \text{(A.13)}$$

and

$$\frac{\partial}{\partial u} (F_{U|V}^{-1}(u|v))^2 = \frac{1}{h(v)} [(1 + h(v))G(u, v) - 1]. \qquad \text{(A.14)}$$

For the density of the copula in the fourth tree of the PVC it follows

$$c_{15;\,2:4}^{\mathrm{PVC}}(a, b) = \frac{\partial^2}{\partial a \partial b} C_{15;\,2:4}^{\mathrm{PVC}}(a, b)$$

$$\stackrel{\text{(A.13),(A.14)}}{=} \left(\int_0^1 G(a, u_4) \mathrm{d}u_4 \right) \left(\int_0^1 G(b, u_2) \mathrm{d}u_2 \right)$$

$$+ \frac{1}{\gamma} \left(1 - \int_0^1 G(b, u_2) \mathrm{d}u_2 \right) \left(\int_0^1 (1 - 2u_4) g(u_4) G(a, u_4) \mathrm{d}u_4 \right)$$

$$+ \frac{1}{\gamma} \left(1 - \int_0^1 G(a, u_4) \mathrm{d}u_4 \right) \left(\int_0^1 (1 - 2u_2) g(u_2) G(b, u_2) \mathrm{d}u_2 \right)$$

$$+ \frac{1}{3} \left(\int_0^1 \frac{g(u_4)}{h(u_4)} [1 - G(a, u_4)] \, \mathrm{d}u_4 \right) \left(\int_0^1 \frac{g(u_2)}{h(u_2)} [1 - G(b, u_2)] \, \mathrm{d}u_2 \right)$$

$$= \frac{1}{4\theta^2} \log(\sigma(a)) \log(\sigma(b))$$

$$+ \frac{1}{\gamma} \left(1 - \frac{1}{2\theta} \log(\sigma(b)) \right) \left(\int_0^1 (1 - 2u_4) g(u_4) G(a, u_4) \mathrm{d}u_4 \right)$$

$$+ \frac{1}{\gamma} \left(1 - \frac{1}{2\theta} \log(\sigma(a)) \right) \left(\int_0^1 (1 - 2u_2) g(u_2) G(b, u_2) \mathrm{d}u_2 \right)$$

$$+ \frac{1}{3} \left(\int_0^1 \frac{g(u_4)}{h(u_4)} [1 - G(a, u_4)] \, \mathrm{d}u_4 \right) \left(\int_0^1 \frac{g(u_2)}{h(u_2)} [1 - G(b, u_2)] \, \mathrm{d}u_2 \right),$$

where

$$\sigma(i) = \left(\frac{\sqrt{(1 + \theta)^2 - 4\theta i} + 1 - 2i + \theta}{\sqrt{(1 - \theta)^2 + 4\theta i} + 1 - 2i - \theta} \right) \qquad \text{for } i \in \{a, b\}.$$

If we set $g(u) := 1 - 2u$, then $\theta = 1/9$ and $\gamma = 1/3$, and we get

$$c_{15;\,2:4}^{\mathrm{PVC}}(a, b) = \frac{81}{4} \prod_{i=a,b} \log(s(i)) + 27 \prod_{i=a,b} \left(1 - \frac{81}{4} \log(s(i)) \right)$$

$$+ \frac{2187}{2} \sum_{(i,j) \in I_{a,b}} \left(1 - \frac{9}{2} \log(s(i)) \right) \left[(6j^2 - 6j + 1) \log(s(j)) \right.$$

$$+ \frac{1}{9}\left((6j - \frac{26}{9})\sqrt{25 - 9j} - (6j - \frac{28}{9})\sqrt{16 + 9j}\right)\Bigg]$$

where

$$s(i) = \frac{\sqrt{25 - 9i} + 5 - 9i}{\sqrt{16 + 9i} + 4 - 9i} \quad \text{for } i \in \{a, b\} \text{ and } I_{a,b} := \{(a, b), (b, a)\}.$$

Evaluating the density shows that $C_{15;2:4}^{\mathrm{PVC}}$ is not the independence copula.

A.5. Proof of Theorem 5.1

The KLD related to tree j, $D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1}))$, is minimized when the negative cross entropy related to tree j is maximized. The negative cross entropy related to tree j is given by

$$H^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1})) := \sum_{i=1}^{d-j} \mathbb{E}\left[\log c_{i,i+j;S_{ij}}^{\mathrm{SVC}}(F_{i|S_{ij}}^{\mathrm{SVC}}(U_i|U_{S_{ij}}), F_{i+j|S_{ij}}^{\mathrm{SVC}}(U_{i+j}|U_{S_{ij}}))\right]$$

$$=: \sum_{i=1}^{d-j} H_i^{(j)}(c_{i,i+j;S_{ij}}^{\mathrm{SVC}}, F_{i|S_{ij}}^{\mathrm{SVC}}, F_{i+j|S_{ij}}^{\mathrm{SVC}}),$$

where for $j = 1$ we set $F_{k|S_{ij}}^{\mathrm{SVC}}(u_k|u_{S_{ij}}) = u_k$, $k = i, i+j$. Obviously, to maximize $H^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1}))$ w.r.t. $\mathcal{T}_j$ we can maximize each $H_i^{(j)}(c_{i,i+j;S_{ij}}^{\mathrm{SVC}}, F_{i|S_{ij}}^{\mathrm{SVC}}, F_{i+j|S_{ij}}^{\mathrm{SVC}})$ individually for all $i = 1, \ldots, d - j$. If $j = 1$, then

$$H_i^{(j)}(c_{i,i+j;S_{ij}}^{\mathrm{SVC}}, F_{i|S_{ij}}^{\mathrm{SVC}}, F_{i+j|S_{ij}}^{\mathrm{SVC}}) = \mathbb{E}\left[\log \frac{c_{i,i+1}(U_i, U_{i+1})}{c_{i,i+1}^{\mathrm{SVC}}(U_i, U_{i+1})}\right]$$

which is maximized for $C_{i,i+1}^{\mathrm{SVC}} = C_{i,i+1}$ by Gibbs' inequality. Thus, if $j = 1$, then

$$\operatorname*{arg\,min}_{\mathcal{T}_j \in \mathbf{T}_j} \; D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1}^{\mathrm{PVC}})) = \mathcal{T}_j^{\mathrm{PVC}}. \tag{A.15}$$

To show that (A.15) holds for $j \geq 2$ we use induction. Assume that

$$\operatorname*{arg\,min}_{\mathcal{T}_j \in \mathbf{T}_j} \; D_{KL}^{(j)}(\mathcal{T}_j(\mathcal{T}_{1:j-1}^{\mathrm{PVC}})) = \mathcal{T}_j^{\mathrm{PVC}}$$

holds for $1 \leq j \leq d-2$. To minimize the KLD related to tree $j+1 =: n$ w.r.t. $\mathcal{T}_n$, conditional on $\mathcal{T}_{1:n-1} = \mathcal{T}_{1:n-1}^{\mathrm{PVC}}$, we have to maximize the negative cross entropy which is maximized if

$$H_i^{(n)}(c_{i,i+n;S_{i,n}}^{\mathrm{SVC}}, F_{i|S_{i,n}}^{\mathrm{PVC}}, F_{i+n|S_{i,n}}^{\mathrm{PVC}})$$

$$= \mathbb{E}\left[\log c_{i,i+n;S_{i,n}}^{\mathrm{SVC}}\left(F_{i|S_{i,n}}^{\mathrm{PVC}}(U_i|U_{S_{i,n}}), F_{i+n|S_{i,n}}^{\mathrm{PVC}}(U_{i+n}|U_{S_{i,n}})\right)\right]$$

is maximized for all $i = 1, \ldots, d - n$. Using the substitutions

$$u_i = (F^{\mathrm{PVC}}_{i|S_{i,n}})^{-1}(t_i|u_{S_{i,n}}) = G_{i|S_{i,n}}(t_i|u_{S_{i,n}})$$

and

$$u_{i+n} = (F^{\mathrm{PVC}}_{i+n|S_{i,n}})^{-1}(t_{i+n}|u_{S_{i,n}}) = G_{i+n|S_{i,n}}(t_{i+n}|u_{S_{i,n}}),$$

we obtain

$$H_i^{(n)}(c^{\mathrm{SVC}}_{i,i+n;S_{i,n}}, F^{\mathrm{PVC}}_{i|S_{i,n}}, F^{\mathrm{PVC}}_{i+n|S_{i,n}}) = \int_{[0,1]^{n+1}} \log c^{\mathrm{SVC}}_{i,i+n;S_{i,n}}(t_i, t_{i+n})$$

$$\times c_{i,i+n;S_{i,n}} \Big(F_{i|S_{i,n}}\big(G_{i|S_{i,n}}(t_i|u_{S_{i,n}})\big|u_{S_{i,n}}\big),$$

$$F_{i+n|S_{i,n}}\big(G_{i+n|S_{i,n}}(t_{i+n}|u_{S_{i,n}})\big|u_{S_{i,n}}\big)\Big|u_{S_{i,n}}\Big)$$

$$\times \frac{\prod_{k=i,i+n} f_{k|S_{i,n}}\big(G_{k|S_{i,n}}(t_k|u_{S_{i,n}})\big|u_{S_{i,n}}\big)}{\prod_{k=i,i+n} f^{\mathrm{PVC}}_{k|S_{i,n}}\big(G_{k|S_{i,n}}(t_k|u_{S_{i,n}})\big|u_{S_{i,n}}\big)} c_{S_{i,n}}(u_{S_{i,n}}) \mathrm{d}u_{S_{i,n}} \mathrm{d}t_i \mathrm{d}t_{i+n}$$

$$= \int_{[0,1]^2} \log c^{\mathrm{SVC}}_{i,i+n;S_{i,n}}(t_i, t_{i+n})$$

$$\times \Big(\int_{[0,1]^{n-1}} c_{i,i+n;S_{i,n}} \Big(F_{i|S_{i,n}}\big(G_{i|S_{i,n}}(t_i|u_{S_{i,n}})\big|u_{S_{i,n}}\big),$$

$$F_{i+n|S_{i,n}}\big(G_{i+n|S_{i,n}}(t_{i+n}|u_{S_{i,n}})\big|u_{S_{i,n}}\big)\Big|u_{S_{i,n}}\Big)$$

$$\times \frac{\prod_{k=i,i+n} f_{k|S_{i,n}}\big(G_{k|S_{i,n}}(t_k|u_{S_{i,n}})\big|u_{S_{i,n}}\big)}{\prod_{k=i,i+n} f^{\mathrm{PVC}}_{k|S_{i,n}}\big(G_{k|S_{i,n}}(t_k|u_{S_{i,n}})\big|u_{S_{i,n}}\big)} c_{S_{i,n}}(u_{S_{i,n}}) \mathrm{d}u_{S_{i,n}} \Big) \mathrm{d}t_i \mathrm{d}t_{i+n}$$

$$= \int_{[0,1]^2} \log c^{\mathrm{SVC}}_{i,i+n;S_{i,n}}(t_i, t_{i+n}) c^{\mathrm{PVC}}_{i,i+n;S_{i,n}}(t_i, t_{i+n}) \mathrm{d}t_i \mathrm{d}t_{i+n},$$

which is maximized for $c^{\mathrm{SVC}}_{i,i+n;S_{i,n}} = c^{\mathrm{PVC}}_{i,i+n;S_{i,n}} = c^{\mathrm{PVC}}_{i,i+(j+1);S_{i,j+1}}$ by Gibbs' inequality.

A.6. Proof of Theorem 5.2

Equation (5.3) is obvious, since $C^{\mathrm{PVC}}_{1:d}$ is the data generating process. Equation (5.5) immediately follows from the equations (5.1) and (5.4). Using the same arguments as in Appendix A.2, the validity of (5.4) for $d = 3$ implies the validity of (5.4) for $d \geq 3$. However, even for $d = 3$, the KLD is a triple integral and does not exhibit an analytical expression if the data generating process is a non-simplified vine copula. Thus, the hard part is to show that there exists a data generating copula which does not satisfy the simplifying assumption and for which the PVC does not minimize the KLD. We prove equation (5.4) for $d = 3$ by means of the following example.

Example A.1.
Let $g\colon [0,1] \to [-1,1]$ be a measurable function. Consider the data generating process

$$C_{1:3}(u_{1:3}) = \int_0^{u_2} C^{FGM_2}\big(u_1, u_3; g(z)\big)\mathrm{d}z,$$

i.e., the two unconditional bivariate margins (C_{12}, C_{23}) are independence copulas and the conditional copula is a FGM copula with varying parameter $g(u_2)$. The first-order partial copula is also a FGM copula given by

$$C_{13;2}^{\text{SVC}}(u_1, u_3; \theta_{13;2}^{\text{PVC}}) = u_1 u_3[1 + \theta_{13;2}^{\text{PVC}}(1 - u_1)(1 - u_3)], \quad \theta_{13;2}^{\text{PVC}} := \int_0^1 g(u_2)\mathrm{d}u_2.$$

We set $C_{23}^{\text{SVC}} = C_{23}, C_{13;2}^{\text{SVC}} = C_{13;2}^{\text{PVC}}$, and specify a parametric copula $C_{12}^{\text{SVC}}(\theta_{12})$, $\theta_{12} \in \Theta_{12} \subset \mathbb{R}$, with conditional cdf $F_{1|2}^{\text{SVC}}(u_1|u_2; \theta_{12})$ and such that $C_{12}^{\text{SVC}}(0)$ corresponds to the independence copula. Thus, $(C_{12}^{\text{SVC}}(0), C_{23}, C_{13;2}^{\text{PVC}}) = (C_{12}, C_{23}, C_{13;2}^{\text{PVC}})$. We also assume that $c_{12}^{\text{SVC}}(u_1, u_2; \theta_{12})$ and $\partial_{\theta_{12}} c_{12}^{\text{SVC}}(u_1, u_2; \theta_{12})$ are both continuous on $(u_1, u_2, \theta_{12}) \in (0,1)^2 \times \Theta_{12}$.

We now derive necessary and sufficient conditions such that

$$D_{KL}(C_{1:3}||C_{12}^{\text{SVC}}(\theta_{12}), C_{23}, C_{13;2}^{\text{PVC}}) := D_{KL}(C_{1:3}||((C_{12}^{\text{SVC}}(\theta_{12}), C_{23}), (C_{13;2}^{\text{PVC}})))$$

attains an extremum at $\theta_{12} = 0$.

Lemma A.1. (Extremum of the KLD in Example A.1)
Let $C_{1:3}$ be given as in Example A.1. For $u_1 \in (0,1)$, we define

$$h(u_1; g) := \int_0^1 \partial_{\theta_{12}} F_{1|2}^{svc}(u_1|u_2; \theta_{12})\Big|_{\theta_{12}=0} g(u_2)\,du_2,$$

$$K(u_1; \theta_{13;2}^{PVC}) := \frac{1}{h(u_1; g)} \int_0^1 \int_0^1 \partial_{\theta_{12}} \log c_{13;2}^{svc}\big(F_{1|2}^{svc}(u_1|u_2; \theta_{12}), u_3; \theta_{13;2}^{PVC}\big)\Big|_{\theta_{12}=0}$$
$$\times c_{1:3}(u_{1:3})\,du_2\,du_3.$$

Then, $\forall u_1 \in (0, 0.5)\colon K(0.5 + u_1; \theta_{13;2}^{PVC}) > 0 \Leftrightarrow \theta_{13;2}^{PVC} > 0$, and $D_{KL}(C_{1:3}||C_{12}^{svc}(\theta_{12}), C_{23}, C_{13;2}^{PVC})$ has an extremum at $\theta_{12} = 0$ if and only if

$$\partial_{\theta_{12}} D_{KL}(C_{1:3}||C_{12}^{svc}(\theta_{12}), C_{23}, C_{13;2}^{PVC})\Big|_{\theta_{12}=0}$$
$$= \int_0^{0.5} K(0.5 + u_1; \theta_{13;2}^{PVC})[h(0.5 + u_1; g) - h(0.5 - u_1; g)]\,du_1 = 0. \quad \text{(A.16)}$$

Proof. See Appendix A.7. $\qquad\square$

It depends on the data generating process whether the condition in Lemma A.1 is satisfied and $D_{KL}(C_{1:3}||C_{12}^{\text{SVC}}(0), C_{23}, C_{13;2}^{\text{PVC}})$ is an extremum or not as we illustrate in the following. If $\theta_{13;2}^{\text{PVC}} = 0$, then $K(u_1; \theta_{13;2}^{\text{PVC}}) = 0$ for all

$u_1 \in (0, 1)$, or if g does not depend on u_2, then $h(u_1; g) = 0$ for all $u_1 \in (0, 1)$. Thus, the integrand in (A.16) is zero and we have an extremum if one of these conditions is true. Assuming $\theta_{13;2}^{\mathrm{PVC}} \neq 0$ and that g depends on u_2, we see from (A.16) that g and C_{12}^{SVC} determine whether we have an extremum at $\theta_{12} = 0$. Depending on the copula family that is chosen for C_{12}^{SVC}, it may be possible that the copula family alone determines whether $D_{KL}(C_{1:3}||C_{12}^{\mathrm{SVC}}(0), C_{23}, C_{13;2}^{\mathrm{PVC}})$ is an extremum. For instance, if C_{12}^{SVC} is a FGM copula we obtain

$$h(u_1; g) = u_1(1 - u_1) \int_0^1 (1 - 2u_2)g(u_2)\mathrm{d}u_2$$

so that

$$h(0.5 + u_1; g) = h(0.5 - u_1; g), \quad \forall u_1 \in (0, 0.5).$$

This symmetry of h across 0.5 implies that (A.16) is satisfied for all functions g.

If we do not impose any constraints on the bivariate copulas in the first tree of the SVC approximation, then $D_{KL}(C_{1:3}||C_{12}^{\mathrm{SVC}}(0), C_{23}, C_{13;2}^{\mathrm{PVC}})$ may not even be a local minimizer of the KLD. For instance, if C_{12}^{SVC} is the asymmetric FGM copula given in (4.1), we find that

$$h(u_1; g) = u_1^2(1 - u_1) \int_0^1 (1 - 2u_2)g(u_2)\mathrm{d}u_2.$$

If $\Lambda := \int_0^1 (1 - 2u_2)g(u_2)\mathrm{d}u_2 \neq 0$, e.g., g is a non-negative function which is increasing, say $g(u_2) = u_2$, then, depending on the sign of Λ, either

$$h(0.5 + u_1; g) > h(0.5 - u_1; g), \quad \forall u_1 \in (0, 0.5),$$

or

$$h(0.5 + u_1; g) < h(0.5 - u_1; g), \quad \forall u_1 \in (0, 0.5),$$

so that the integrand in (A.16) is either strictly positive or negative and thus $D_{KL}(C_{1:3}||C_{12}, C_{23}, C_{13;2}^{\mathrm{PVC}})$ can not be an extremum. Since $\theta_{12} \in [-1, 1]$, it follows that $D_{KL}(C_{1:3}||C_{12}^{\mathrm{SVC}}(0), C_{23}, C_{13;2}^{\mathrm{PVC}})$ is not a local minimum. As a result, we can, relating to the PVC, further decrease the KLD from the true copula if we adequately specify "wrong" copulas in the first tree and choose the first-order partial copula in the second tree of the SVC approximation.

A.7. Proof of Lemma A.1

The KLD attains an extremum if and only if the negative cross entropy attains an extremum. The negative cross entropy is given by

$$\begin{aligned}
H_{1:3}(C_{1:3}||C_{1:3}^{\mathrm{SVC}}(\theta_{12}; \theta_{13;2}^{\mathrm{PVC}})) &= \mathbb{E}[\log c_{1:3}^{\mathrm{SVC}}(U_{1:3}; \theta_{12}, \theta_{13;2}^{\mathrm{PVC}})] \\
&= \mathbb{E}[\log[c_{12}^{\mathrm{SVC}}(U_{1:2}; \theta_{12})c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(U_1|U_2; \theta_{12}), U_3; \theta_{13;2}^{\mathrm{PVC}}\big)]] \\
&= \mathbb{E}[\log c_{12}^{\mathrm{SVC}}(U_{1:2}; \theta_{12})] + \mathbb{E}[\log c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(U_1|U_2; \theta_{12}), U_3; \theta_{13;2}^{\mathrm{PVC}}\big)].
\end{aligned}$$

If the negative cross entropy attains an extremum then the derivative of $\mathbb{E}[\log c_{1:3}^{\mathrm{SVC}}(U_{1:3};\theta_{12})]$ w.r.t. θ_{12} is zero. Since $c_{12}^{\mathrm{SVC}}(u_1,u_2;\theta_{12})$ and $\partial_{\theta_{12}} c_{12}^{\mathrm{SVC}}(u_1,u_2;\theta_{12})$ are both continuous on $(u_1,u_2,\theta_{12}) \in (0,1)^2 \times \Theta_{12}$, we can apply Leibniz's rule for differentiation under the integral sign to conclude that

$$\partial_{\theta_{12}}\mathbb{E}[\log c_{12}^{\mathrm{SVC}}(U_{1:2};\theta_{12})]\big|_{\theta_{12}=0} = \mathbb{E}[\partial_{\theta_{12}}\log c_{12}^{\mathrm{SVC}}(U_{1:2};\theta_{12})\big|_{\theta_{12}=0}] = 0$$

because $C_{12}^{\mathrm{SVC}}(0)$ is the true copula of $U_{1:2}$. Thus, the derivative evaluated at $\theta_{12} = 0$ becomes

$$\partial_{\theta_{12}}\mathbb{E}[\log c_{1:3}^{\mathrm{SVC}}(U_{1:3};\theta_{12},\theta_{13;2}^{\mathrm{PVC}})]\Big|_{\theta_{12}=0}$$

$$= \partial_{\theta_{12}}\mathbb{E}[\log c_{12}^{\mathrm{SVC}}(U_{1:2};\theta_{12})]\Big|_{\theta_{12}=0}$$

$$\quad + \partial_{\theta_{12}}\mathbb{E}[\log c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(U_1|U_2;\theta_{12}),U_3;\theta_{13;2}^{\mathrm{PVC}}\big)]\Big|_{\theta_{12}=0}$$

$$= \partial_{\theta_{12}}\mathbb{E}[\log c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(U_1|U_2;\theta_{12}),U_3;\theta_{13;2}^{\mathrm{PVC}}\big)]\Big|_{\theta_{12}=0}$$

$$= \partial_{\theta_{12}}\int_{[0,1]^3} \log c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12}),u_3;\theta_{13;2}^{\mathrm{PVC}}\big)c_{1:3}(u_{1:3})\mathrm{d}u_{1:3}\Big|_{\theta_{12}=0}$$

$$= \int_{[0,1]^3} \frac{\partial_1 c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12}),u_3;\theta_{13;2}^{\mathrm{PVC}}\big)}{c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12}),u_3;\theta_{13;2}^{\mathrm{PVC}}\big)} \partial_{\theta_{12}}F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12})\Big|_{\theta_{12}=0}$$

$$\quad \times c_{1:3}(u_{1:3})\mathrm{d}u_{1:3},$$

where $\partial_1 c_{13;2}^{\mathrm{SVC}}(u,v;\theta_{13;2}^{\mathrm{PVC}})$ is the partial derivative w.r.t. u and we have used Leibniz's integral rule to perform the differentiation under the integral sign for the second last equality which is valid since the integrand and its partial derivative w.r.t. θ_{12} are both continuous in $u_{1:3}$ and θ_{12} on $(0,1)^3 \times (-1,1)$.

To compute the integral we observe that

$$\partial_1 c_{13;2}^{\mathrm{SVC}}(u,v;\theta_{13;2}^{\mathrm{PVC}}) = -2\theta_{13;2}^{\mathrm{PVC}}(1-2v),$$

$$\partial_1 c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;0),u_3;\theta_{13;2}^{\mathrm{PVC}}\big) = -2\theta_{13;2}^{\mathrm{PVC}}(1-2u_3),$$

$$\frac{\partial_1 c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12}),u_3;\theta_{13;2}^{\mathrm{PVC}}\big)}{c_{13;2}^{\mathrm{SVC}}\big(F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12}),u_3;\theta_{13;2}^{\mathrm{PVC}}\big)}\bigg|_{\theta_{12}=0} = \frac{-2\theta_{13;2}^{\mathrm{PVC}}(1-2u_3)}{1+\theta_{13;2}^{\mathrm{PVC}}(1-2u_1)(1-2u_3)}.$$

Note that $m(u_1,u_3;\theta_{13;2}^{\mathrm{PVC}}) := \frac{-2\theta_{13;2}^{\mathrm{PVC}}(1-2u_3)}{1+\theta_{13;2}^{\mathrm{PVC}}(1-2u_1)(1-2u_3)}$ does not depend on u_2. Moreover, with $c_{1:3}(u_{1:3}) = 1 + g(u_2)(1-2u_1)(1-2u_3)$,

$$\int_0^1 \partial_{\theta_{12}}F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12},\theta_{13;2}^{\mathrm{PVC}})\Big|_{\theta_{12}=0} c_{1:3}(u_{1:3})\mathrm{d}u_2$$

$$= \int_0^1 \partial_{\theta_{12}}F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12})\Big|_{\theta_{12}=0}\mathrm{d}u_2 + (1-2u_1)(1-2u_3)$$

$$\quad \times \int_0^1 \partial_{\theta_{12}}F_{1|2}^{\mathrm{SVC}}(u_1|u_2;\theta_{12})\Big|_{\theta_{12}=0}g(u_2)\mathrm{d}u_2$$

$$= (1 - 2u_1)(1 - 2u_3)h(u_1; g),$$

where the second equality follows because

$$\int_0^1 \partial_{\theta_{12}} F_{1|2}^{\mathrm{SVC}}(u_1|u_2; \theta_{12})\mathrm{d}u_2 = \partial_{\theta_{12}} \int_0^1 F_{1|2}^{\mathrm{SVC}}(u_1|u_2; \theta_{12})\mathrm{d}u_2 = \partial_{\theta_{12}} u_1 = 0.$$

Thus, integrating out u_2, we obtain

$$\partial_{\theta_{12}} \mathbb{E}[\log c_{1:3}^{\mathrm{SVC}}(U_{1:3}; \theta_{12})]\Big|_{\theta_{12}=0}$$

$$= \int_{[0,1]^2} m(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})(1 - 2u_1)(1 - 2u_3)h(u_1; g)\mathrm{d}u_1\mathrm{d}u_3$$

$$= \int_{[0,1]^2} f(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})h(u_1; g)\mathrm{d}u_1\mathrm{d}u_3, \tag{A.17}$$

where $f(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}}) := m(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})(1 - 2u_1)(1 - 2u_3)$. We note that $\forall u_1 \in (0, 0.5)$, $u_3 \in (0, 1)$:

$$f(0.5 + u_1, u_3; \theta_{13;2}^{\mathrm{PVC}}) > 0 \Leftrightarrow \theta_{13;2}^{\mathrm{PVC}} > 0,$$

$$f(0.5 - u_1, u_3; \theta_{13;2}^{\mathrm{PVC}}) = -f(0.5 + u_1, 1 - u_3; \theta_{13;2}^{\mathrm{PVC}}).$$

So, if $u_1 \in (0, 0.5)$ then

$$\int_0^1 f(0.5 - u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3 = \int_0^1 f(0.5 - u_1, 1 - u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3$$

$$= -\int_0^1 f(0.5 + u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3.$$

Thus, if we define $K(u_1; \theta_{13;2}^{\mathrm{PVC}}) := \int_0^1 f(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3$ we have that $\forall u_1 \in (0, 0.5)$:

$$K(0.5 + u_1; \theta_{13;2}^{\mathrm{PVC}}) > 0 \Leftrightarrow \theta_{13;2}^{\mathrm{PVC}} > 0,$$

$$K(0.5 - u_1; \theta_{13;2}^{\mathrm{PVC}}) = -K(0.5 + u_1; \theta_{13;2}^{\mathrm{PVC}}). \tag{A.18}$$

Plugging this into our integral (A.17) yields

$$\partial_{\theta_{12}} \mathbb{E}[\log c_{1:3}^{\mathrm{SVC}}(U_{1:3}; \theta_{12}, \theta_{13;2}^{\mathrm{PVC}})]\Big|_{\theta_{12}=0}$$

$$= \int_0^1 \int_0^1 f(u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})h(u_1; g)\mathrm{d}u_1\mathrm{d}u_3$$

$$= \int_0^{0.5} h(0.5 - u_1; g) \left(\int_0^1 f(0.5 - u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3 \right) \mathrm{d}u_1$$

$$+ \int_0^{0.5} h(0.5 + u_1; g) \left(\int_0^1 f(0.5 + u_1, u_3; \theta_{13;2}^{\mathrm{PVC}})\mathrm{d}u_3 \right) \mathrm{d}u_1$$

$$= \int_0^{0.5} h(0.5 - u_1; g) K(0.5 - u_1; \theta_{13;2}^{\mathrm{PVC}}) \mathrm{d}u_1$$

$$+ \int_0^{0.5} h(0.5 + u_1; g) K(0.5 + u_1; \theta_{13;2}^{\mathrm{PVC}}) \mathrm{d}u_1$$

$$\overset{(A.18)}{=} \int_0^{0.5} K(0.5 + u_1; \theta_{13;2}^{\mathrm{PVC}})[h(0.5 + u_1; g) - h(0.5 - u_1; g)] \mathrm{d}u_1.$$

Note that if $\theta_{13;2}^{\mathrm{PVC}} = 0$, then $K(u_1; \theta_{13;2}^{\mathrm{PVC}}) = 0$ for all $u_1 \in (0, 1)$, or if g does not depend on u_2, then $h(u_1; g) = 0$ for all $u_1 \in (0, 1)$, so in both cases the integrand is zero and we have an extremum.

A.8. Proof of Corollary 6.1

Corollary 6.1 (i) and (ii) follow directly from Theorem 1 in Spanhel and Kurz [36], which states the asymptotic distribution of approximate rank Z-estimators if the data generating process is not nested in the parametric model family. Corollary 6.1 (iii) follows then from Theorem 5.2 and Theorem 5.1.

A.9. An example where the difference between $\hat{\theta}^S$ and $\hat{\theta}^J$ is more pronounced

Example A.2.
Let $C^{\mathrm{BB1}}(\theta, \delta)$ denote the BB1 copula with dependence parameter (θ, δ) and $C^{\mathrm{Sar}}(\alpha)$ be the Sarmanov copula with cdf

$$C(u, v; \alpha) = uv\big(1 + (3\alpha + 5\alpha^2 \prod_{i=u,v}(1 - 2i)) \prod_{i=u,v}(1 - i)\big),$$

for $|\alpha| \leq \sqrt{7}/5$. The partial Sarmanov copula is given by $C^{\mathrm{P\text{-}Sar}}(u, v; a, b) = uv\big(1 + (3a + 5b \prod_{i=u,v}(1 - 2i)) \prod_{i=u,v}(1 - i)\big)$, where $|a| \leq \sqrt{7}/5$ and $a^2 \leq b \leq (\sqrt{1 - 3a^2} + 1)/5$. Define $S(u_2) = (1 + \exp(u_2))^{-1}$ and $f(u_2) = 1 - 2S(10u_2 - 0.5)) + 2(1 - 2u_2)S(-5)$ so that $g(u_2) = 0.1\left(\sqrt{7} + 1\right)\left(1 - f(u_2)\right) - 0.2$. Let $C_{1:3}$ be the true copula with $(C_{12}, C_{23}, C_{13;2}) = (C^{\mathrm{BB1}}(2, 2), C^{\mathrm{BB1}}(2, 2), C^{\mathrm{Sar}}(g(u_2)))$ and $C_{1:3}^{\mathrm{SVC}}(\theta, \delta, a, b) = (C^{\mathrm{BB1}}(\theta_{12}, \delta_{12}), C^{\mathrm{BB1}}(\theta_{23}, \delta_{23}), C^{\mathrm{P\text{-}Sar}}(a, b))$ be the parametric SVC that is fitted to data generated from $C_{1:3}$.

Note that g is a sigmoid function, with $(g(0), g(1)) = (-0.2, \sqrt{7}/5)$, so that Spearman's rho of the conditional copula $C^{\mathrm{Sar}}(g(u_2))$ varies in the interval $(g(0), g(1)) = (-0.2, \sqrt{7}/5)$ because $\rho_{C^{\mathrm{Sar}}} = \alpha$. Figure 5 shows that the difference between step-by-step and joint ML estimates for the two parameters of the first copula in the first tree is already (individually) significant at the 5% level if the sample size is 500 observations. Thus, the difference between step-by-step and joint ML estimates can be relevant for moderate sample sizes if the variation in the conditional copula is strong enough. Once again, the difference between step-by-step and joint ML estimates is less pronounced for the parameters of $C_{13;2}^{\mathrm{SVC}}$ but it also becomes highly significant with sufficient sample size.
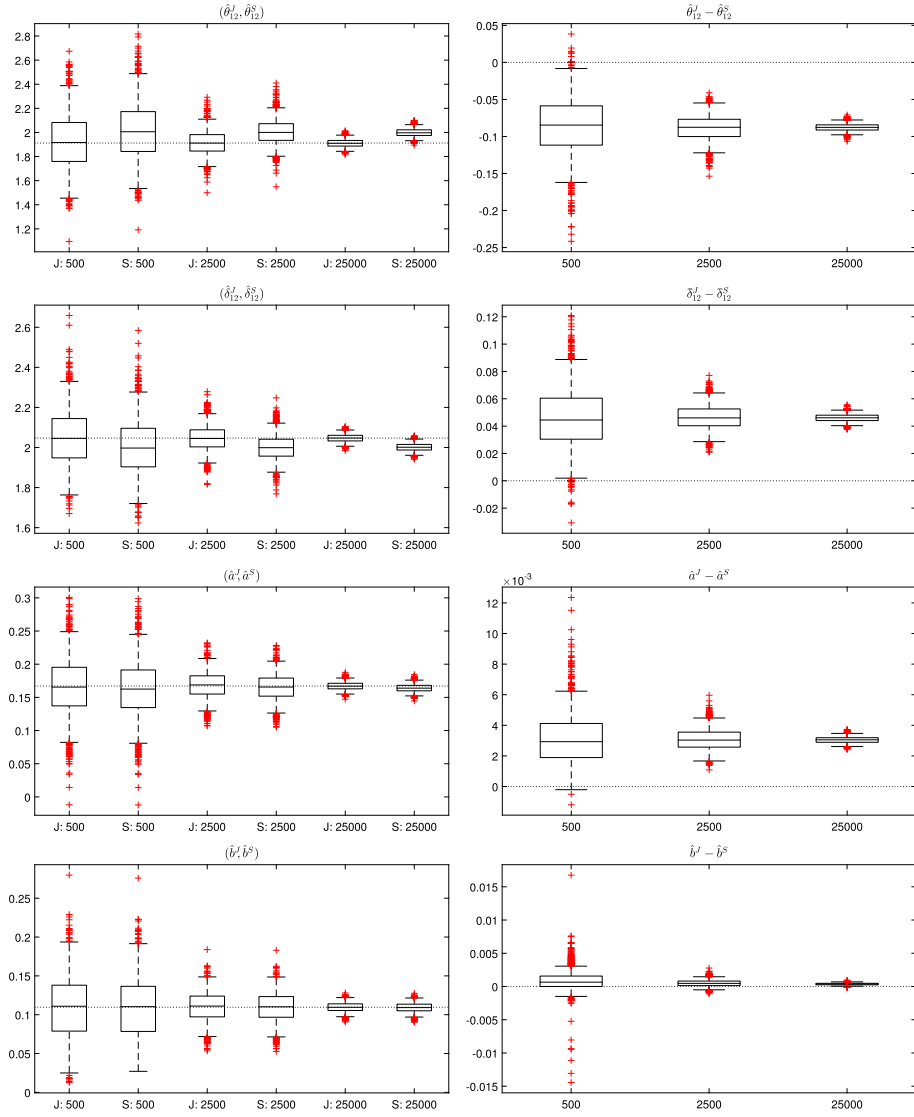
FIG 5. *Box plots of joint (J) and sequential (S) ML estimates and their difference for sample sizes $N = 500, 2500, 25000$, if the data is generated from $C_{1:3}$ in Example A.2 and the pair-copula families of the SVC are given by the corresponding PVC. The dotted line indicates the pseudo-true parameter and zero, respectively. The end of the whiskers is 0.953 times the inter-quartile range, corresponding to approximately 95% coverage if the data is generated by a normal distribution.*

Acknowledgements

comments which helped to improve this paper. We also would like to thank Roger Cooke and Irène Gijbels for interesting discussions on the simplifying assumption.

References

[1] AAS, K. and BERG, D. (2009). Models for construction of multivariate dependence – a comparison study. *The European Journal of Finance* **15** 639–659.

[2] AAS, K., CZADO, C., FRIGESSI, A. and BAKKEN, H. (2009). Pair-Copula Constructions of Multiple Dependence. *Insurance: Mathematics and Economics* **44** 182–198. MR2517884

[3] ACAR, E. F., GENEST, C. and NEŠLEHOVÁ, J. (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis* **110** 74–90. MR2927510

[4] BEDFORD, T. and COOKE, R. M. (2002). Vines – A New Graphical Model for Dependent Random Variables. *The Annals of Statistics* **30** 1031–1068. MR1926167

[5] BERGSMA, W. P. (2004). Testing conditional independence for continuous random variables. URL: https://www.eurandom.tue.nl/reports/2004/048-report.pdf.

[6] DISSMANN, J., BRECHMANN, E. C., CZADO, C. and KUROWICKA, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis* **59** 52–69. MR3000041

[7] FERMANIAN, J.-D. and WEGKAMP, M. H. (2012). Time-dependent copulas. *Journal of Multivariate Analysis* **110** 19–29. MR2927507

[8] FISCHER, M., KÖCK, C., SCHLÜTER, S. and WEIGERT, F. (2009). An empirical analysis of multivariate copula models. *Quantitative Finance* **9** 839–854. MR2588311

[9] GIJBELS, I., OMELKA, M. and VERAVERBEKE, N. (2015a). Estimation of a Copula when a Covariate Affects only Marginal Distributions. *Scandinavian Journal of Statistics* **42** 1109–1126. MR3426313

[10] GIJBELS, I., OMELKA, M. and VERAVERBEKE, N. (2015b). Partial and average copulas and association measures. *Electronic Journal of Statistics* **9** 2420–2474. MR3425363

[11] GROTHE, O. and NICKLAS, S. (2013). Vine constructions of Lévy copulas. *Journal of Multivariate Analysis* **119** 1–15. MR3061411

[12] HOBÆK HAFF, I. (2012). Comparison of estimators for pair-copula constructions. *Journal of Multivariate Analysis* **110** 91–105. MR2927511

[13] HOBÆK HAFF, I. (2013). Parameter estimation for pair-copula constructions. *Bernoulli* **19** 462–491. MR3037161

[14] 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. MR2595309

[15] Hobæk Haff, I. and Segers, J. (2015). Nonparametric estimation of pair-copula constructions with the empirical pair-copula. *Computational Statistics & Data Analysis* **84** 1–13. MR3292793

[16] Joe, H. (1996). Families of m-Variate Distributions with Given Margins and m(m-1)/2 Bivariate Dependence Parameters. *Lecture Notes-Monograph Series* **28** 120–141. MR1485527

[17] Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall, London. MR1462613

[18] Joe, H., Li, H. and Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis* **101** 252–270. MR2557632

[19] Kauermann, G., Schellhase, C. and Ruppert, D. (2013). Flexible Copula Density Estimation with Penalized Hierarchical B-splines. *Scandinavian Journal of Statistics* **40** 685–705. MR3145112

[20] Kauermann, G. and Schellhase, C. (2014). Flexible pair-copula estimation in D-vines using bivariate penalized splines. *Statistics and Computing* **24** 1081–1100. MR3253855

[21] Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis* **120** 85–101. MR3072719

[22] Kurowicka, D. and Cooke, R. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling.* Wiley, Chichester. MR2216540

[23] Kurowicka, D. and Joe, H., eds. (2011). *Dependence Modeling.* World Scientific, Singapore. MR2856966

[24] Kurz, M. S. and Spanhel, F. (2017). Testing the simplifying assumption in high-dimensional vine copulas. ArXiv e-prints `arXiv:1706.02338`.

[25] McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management. Princeton series in finance.* Princeton Univ. Press, Princeton NJ. MR2175089

[26] Mesfioui, M. and Quessy, J.-F. (2008). Dependence structure of conditional Archimedean copulas. *Journal of Multivariate Analysis* **99** 372–385. MR2396969

[27] Nagler, T. and Czado, C. (2016). Evading the curse of dimensionality in nonparametric density estimation with simplified vine copulas. *Journal of Multivariate Analysis* **151** 69–89. MR3545278

[28] Nelsen, R. B. (2006). *An Introduction to Copulas. Springer Series in Statistics.* Springer, New York. MR2197664

[29] Nikoloulopoulos, A. K., Joe, H. and Li, H. (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis* **56** 3659–3673. MR2943919

[30] Patton, A. J. (2006). Modelling Asymmetric Exchange Rate Dependence. *International Economic Review* **47** 527–556. MR2216591

[31] Portier, F. and Segers, J. (2018). On the weak convergence of the empirical conditional copula under a simplifying assumption. *Journal of Multivariate Analysis* **166** 160–181. MR3799641

[32] Remillard, B. (2013). *Statistical Methods for Financial Engineering.* CRC Press, Boca Raton, FL. MR3098545

[33] Schellhase, C. and Spanhel, F. (2018). Estimating non-simplified vine copulas using penalized splines. *Statistics and Computing* **28** 387–409. MR3747570

[34] Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T. and Erhardt, T. (2016). VineCopula: Statistical Inference of Vine Copulas URL: https://CRAN.R-project.org/package=VineCopula, R package version 2.0.5.

[35] Spanhel, F. and Kurz, M. S. (2016a). The partial copula: Properties and associated dependence measures. *Statistics & Probability Letters* **119** 76–83. MR3555271

[36] Spanhel, F. and Kurz, M. S. (2016b). Estimating parametric simplified vine copulas under general conditions. unpublished working paper.

[37] Stöber, J., Joe, H. and Czado, C. (2013). Simplified pair copula constructions—Limitations and extensions. *Journal of Multivariate Analysis* **119** 101–118. MR3061418

[38] Vatter, T. and Nagler, T. (2018). Generalized Additive Models for Pair-Copula Constructions. *Journal of Computational and Graphical Statistics* **27** 715–727. MR3890864

[39] White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica* **50** 1–25. MR0640163