Edith Kovács; Tamás Szántai

On the connection between cherry-tree copulas and truncated R-vine copulas

Kybernetika, Vol. 53 (2017), No. 3, 437–460

Persistent URL: http://dml.cz/dmlcz/146936

KYBERNETIKA — VOLUME 53 (2017), NUMBER 3, PAGES 437–460

ON THE CONNECTION BETWEEN CHERRY-TREE COPULAS AND TRUNCATED R-VINE COPULAS

Edith Kovács and Tamás Szántai

Vine copulas are a flexible way for modeling dependences using only pair-copulas as building blocks. However if the number of variables grows the problem gets fastly intractable. For dealing with this problem Brechmann at al. proposed the truncated R-vine copulas. The truncated R-vine copula has the very useful property that it can be constructed by using only pair-copulas and a lower number of conditional pair-copulas. In our earlier papers we introduced the concept of cherry-tree copulas. In this paper we characterize the relation between cherry-tree copulas and truncated R-vine copulas. It turns out that the concept of cherry-tree copula is more general than the concept of truncated R-vine copula. Although both contain in their expressions conditional independences between the variables, the truncated R-vines constructed in greedy way do not exploit the existing conditional independences in the data. We give a necessary and sufficient condition for a cherry-tree copula to be a truncated R-vine copula. We introduce a new method for truncated R-vine modeling. The new idea is that in the first step we construct the top tree by exploiting conditional independences for finding a good-fitting cherry-tree of order k. If this top tree is a tree in an R-vine structure then this will define a truncated R-vine at level k and in the second step we construct a sequence of trees which leads to it. If this top tree is not a tree in an R-vine structure then we can transform it into such a tree at level $k+1$ and then we can again apply the second step. The second step is performed by a backward construction named Backward Algorithm. This way the cherry-tree copulas always can be expressed by pair-copulas and conditional pair-copulas.

Keywords: copula, conditional independences, Regular-vine, truncated vine, cherry-tree copula

Classification: 60C05, 62H05

1. INTRODUCTION

Copulas in general are known to be useful tool for modeling multivariate probability distributions since they make possible to model separately the dependence structure and the univariate marginals. In this paper we show how conditional independences can be utilized in the expression of multivariate copulas. Regarding to this we proved in [19] a theorem which links to a junction tree probability distribution the so called junction tree copula.

DOI: 10.14736/kyb-2017-3-0437

The paper [1] calls the attention to the fact that "conditional independence may reduce the number of the pair-copula decompositions and hence simplify the construction". In this paper the importance of choosing a good factorisation which takes advantage of the conditional independence relations between the random variables is pointed out. In [19] we introduced the concept of cherry-tree copulas which exploits the conditional independences between the variables.

The importance of taking into account the conditional independences between the variables encoded in a Bayesian Network (related to a directed acyclic graph) was explored in the papers written by Kurowicka and Cooke [21], by Hanea et al. [12] and Bauer et al. [3]. Two aspects of this problem were discussed. First, when the Bayesian Network (BN) is known, some of the conditional independences taken from the BN are used to simplify a given expression of the D- or C-vine copulas, which are very special types of vine copulas (see [3]). Second, the problem of reconstruction of the BN from a sample data set was formulated under the assumption that the joint distribution is multivariate normal. For discovering the independences and conditional independences between the variables in Hanea et al. [12] the correlations, the conditional correlations and the determinant of the correlation matrix are used.

The paper [3] is dealing with more general pair-copula constructions related to the non-Gaussian BNs. In their paper the BN is supposed to be known. The formula of probability distribution associated to the given BN is expressed by pair-copulas.

The so-called truncated Regular-vine (R-vine) copula is defined by Kurowicka in [23] and by Brechmann et al. in [6]. In [23] an algorithm was developed for searching the "best truncated vine", which was defined as the one whose nodes of the top trees (trees with most conditioning) correspond to the smallest absolute values of correlations. This restricts the applicability of this method to Gaussian copulas.

In this paper we recall the concept of the cherry-tree copulas. An alternative definition of R-vines using a special hypergraph structure is given in [19]. There we proved that truncated vine copula is a special case of the cherry-tree copula.

In the preliminary section we recall all concepts that we need in the paper. First we will remind some graph theoretical concepts since the conditional independences can be represented on graphs. Then the concepts of copulas and R-vine copulas will be recalled. Finally the multivariate junction tree copula associated to a junction tree probability distribution and an equivalent definition of the R-vine copulas based on the cherry-tree graph structures will also be presented. In the third section we give a necessary and sufficient condition for a cherry-tree copula to be a truncated R-vine copula and an algorithm for obtaining the truncated R-vine structure from a given cherry-tree copula. In the fourth section we give a theorem for transforming a general cherry-tree copula into a truncated R-vine copula. We finish the paper with conclusions and with highlighting the new perspectives given by the paper.

2. PRELIMINARIES

The reader who is familiar with the basic concepts presented in this preliminary section may skip some parts of it.

2.1. Acyclic hypergraph, junction tree, junction tree probability distribution

In this subsection we recall some concepts used in graph theory and probability theory which are needed throughout the paper and present how these can be linked to each other. For a good overview see [24] and [25]. We first present the acyclic hypergraphs and junction trees. Then we introduce the cherry-trees as a special type of junction trees. We finish this subsection with the multivariate joint probability distribution associated to junction trees.

Let $V = \{1, \ldots, d\}$ be a set of vertices and Γ a set of subsets of V called *set of hyperedges*. A *hypergraph* consists of a set V of vertices and a set Γ of hyperedges. We denote a hyperedge by K_i, where K_i is a subset of V.

The *acyclic hypergraph* is a special type of hypergraph which fulfills the following requirements:

- Neither of the hyperedges of Γ is a subset of another hyperedge.
- There exists a numbering of edges for which the *running intersection property* is fulfilled: $\forall j \geq 2 \quad \exists\, i < j : \; K_i \supset K_j \cap (K_1 \cup \ldots \cup K_{j-1})$. (Other formulation is that for all hyperedges K_i and K_j with $i < j-1$, $K_i \cap K_j \subset K_s$ for all $s, i < s < j$.)

Let $S_j = K_j \cap (K_1 \cup \ldots \cup K_{j-1})$, for $j > 1$ and $S_1 = \phi$. Let $R_j = K_j \backslash S_j$. We say that S_j *separates* R_j from $(K_1 \cup \ldots \cup K_{j-1}) \backslash S_j$, and call S_j separator set or shortly separator.

Now we link these concepts to the terminology of junction trees.

The *junction tree* is a special tree stucture which is equivalent to the connected acyclic hypergraphs (see in [24] and [25]). The "nodes" of the tree correspond to the hyperedges of the connected acyclic hypergraph and are called clusters, the "edges" of the junction tree correspond to the separator sets and called separators. The set of all clusters is denoted by Γ, the set of all separators is denoted by $\mathcal{S}$. A junction tree (V, Γ, S) is defined by the set of vertices V, the set of nodes Γ called also set of clusters, and the set of separators S. The junction tree with the largest cluster containing k variables is called *k-width junction tree*.

An important relation between graphs and hypergraphs is given in [24]: A hypergraph is acyclic if and only if it can be considered to be the set of maximal cliques (complete graphs) of a triangulated (*chordal*) graph. We remind here that a clique is a subset of vertices of an undirected graph such that its induced subgraph is complete and a graph is triangulated if every cycle of length greater than 3 has a chord. This means that the vertices in a cluster are all connected with each other.

In the Figure 1 one can see a) a triangulated graph, b) the corresponding acyclic hypergraph and c) the corresponding junction tree.

We consider the random vector $\mathbf{X} = (X_1, \ldots, X_d)^T$. Through the paper we use the following notations: $V = \{1, \ldots, d\}$ the set of indices, $X = \{X_1, \ldots, X_d\}$ the set of the components of the random vector $\mathbf{X}$ and $\mathbf{X}_A$ a random vector with components $X_i, i \in A \subset \{1, \ldots, d\}$.

Major advances in probabilistic inference methods based on graphical representations have been realized by Lauritzen and Spiegelhalter [24] and Lauritzen [25]. However

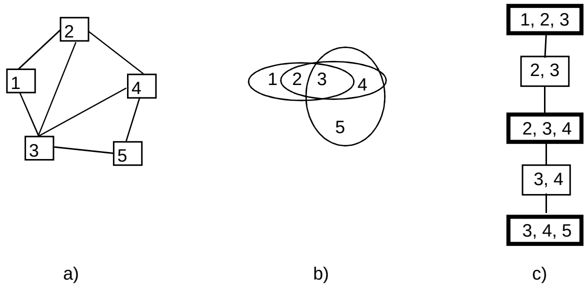

Fig. 1. a) Triangulated graph, b) The corresponding acyclic hypergraph, c) The corresponding junction tree which is a cherry-tree.

probabilistic inference has the inherent disadvantage of being NP-hard. By exploiting the conditional independence relations among the discrete random variables of a probabilistic network the underlying joint probability space maybe decomposed into smaller subspaces corresponding to cliques in a triangulated graph ([24] and [25]).

We say that a probability distribution has the Global Markov (GM) property described by a graph if for any $A, B, C \subset V$ for which C separates A and B in graph theoretical sense $\mathbf{X}_A$ and $\mathbf{X}_B$ are conditionally independent given $\mathbf{X}_C$. This can be formulated in terms of probabilities as

$$P\left(\mathbf{X}_{A\cup B\cup C}\right) = \frac{P\left(\mathbf{X}_{A\cup C}\right) P\left(\mathbf{X}_{B\cup C}\right)}{P\left(\mathbf{X}_C\right)}.$$

The concept of *junction tree probability distribution* is related to the junction tree graph and to the Global Markov property. A junction tree probability distribution is defined as a fraction of some products of marginal probability distributions as follows:

$$P\left(\mathbf{X}\right) = \frac{\prod\limits_{C\in\Gamma} P\left(\mathbf{X}_C\right)}{\prod\limits_{S\in\mathcal{S}} \left[P\left(\mathbf{X}_S\right)\right]^{\nu_S - 1}}, \tag{1}$$

where Γ is the set of clusters of the junction tree, $\mathcal{S}$ is the set of separators, ν_S is the number of those clusters which are linked by the separator S.

Example 2.1. The probability distribution corresponding to Figure 1 is:

$$\begin{aligned} P\left(\mathbf{X}\right) &= \frac{P\left(\mathbf{X}_{\{1,2,3\}}\right) P\left(\mathbf{X}_{\{2,3,4\}}\right) P\left(\mathbf{X}_{\{3,4,5\}}\right)}{P\left(\mathbf{X}_{\{2,3\}}\right) P\left(\mathbf{X}_{\{3,4\}}\right)} \\ &= \frac{P\left(X_1, X_2, X_3\right) P\left(X_2, X_3, X_4\right) P\left(X_3, X_4, X_5\right)}{P\left(X_2, X_3\right) P\left(X_3, X_4\right)}. \end{aligned}$$

In [7] and [8] there were used and named the so called t-cherry-tree graph structures. Since these can be regarded as a special type of junction tree we can give now the following definition. In this paper we will call this structure simply cherry-tree as this does not cause any confusion.

Definition 2.2. We call k order cherry-tree the junction tree with all clusters of size k and all separators of size $k-1$.

Denoting by $\mathcal{C}_{\text{ch}}$ and $\mathcal{S}_{\text{ch}}$ the set of clusters and the set of separators of the cherry junction tree we gave the following definition.

Definition 2.3. (Kovács and Szántai [19]) In the discrete case the probability distribution given by (2) is called *cherry-tree probability distribution*

$$P_{\text{ch}}(\mathbf{X}) = \frac{\prod\limits_{K\in\mathcal{C}_{\text{ch}}} P\left(\mathbf{X}_K\right)}{\prod\limits_{S\in\mathcal{S}_{\text{ch}}} \left(P\left(\mathbf{X}_S\right)\right)^{\nu_s-1}} \tag{2}$$

and in the continuous case the probability distribution given by (3) is called *cherry-tree probability density function*

$$f_{ch}\left(\mathbf{x}\right) = \frac{\prod\limits_{K\in C_{ch}} f_K\left(\mathbf{x}_K\right)}{\prod\limits_{S\in S_{Ch}} \left(f_S\left(\mathbf{x}_S\right)\right)^{\nu_S-1}}, \tag{3}$$

where in both cases ν_S denotes the number of clusters which are linked by the separator S.

The marginal probability distributions and the marginal density functions involved in the above formulae are marginal probability distributions of $P\left(\mathbf{X}\right)$, respectively marginal density functions of $f\left(\mathbf{x}\right)$.

2.2. Copula, R-vine copula

In this subsection we recall some definitions according to copulas and R-vine copulas.

Definition 2.4. A function $C:[0;1]^d \to [0;1]$ is called a d-dimensional copula if it satisfies the following conditions:

1. $C\left(u_1,\ldots,u_d\right)$ is increasing in each component u_i,
2. $C\left(u_1,\ldots,u_{i-1},0,u_{i+1},\ldots,u_d\right)=0$ for all $u_k\in[0;1]$, $k\neq i,\ i=1,\ldots,n$,
3. $C\left(1,\ldots,1,u_i,1,\ldots,1\right)=u_i$ for all $u_i\in[0;1]\,,\ i=1,\ldots,d$,
4. C is d-increasing, i.e for all $\left(u_{1,1},\ldots,u_{1,d}\right)$ and $\left(u_{2,1},\ldots,u_{2,d}\right)$ in $[0;1]^d$ with $u_{1,i}<u_{2,i}$ for all i, we have

$$\sum_{i_1=1}^{2}\cdots\sum_{i_d=1}^{2}\left(-1\right)^{\sum\limits_{j=1}^{d} i_j} C\left(u_{i_1,1},\ldots,u_{i_d,d}\right)\geq 0.$$

Due to Sklar's theorem if $X_1, \ldots, X_d$ are continuous random variables defined on a common probability space, with the univariate marginal cdf's $F_{X_i}(x_i)$ and the joint cdf $F_{X_1,\ldots,X_d}(x_1, \ldots, x_d)$ then there exists a unique copula function $C_{X_1,\ldots,X_d}(u_1, \ldots, u_d) : [0;1]^d \to [0;1]$ such that by the substitution $u_i = F_i(x_i)\,, \; i = 1, \ldots, d$ we get

$$F_{X_1,\ldots,X_d}(x_1, \ldots, x_d) = C_{X_1,\ldots,X_d}(F_1(x_1), \ldots, F_d(x_d))$$

for all $(x_1, \ldots, x_d)^T \in R^d$.

In the following we will use the vectorial notation $F_{\mathbf{X}_D}(\mathbf{x}_D) = C_{\mathbf{X}_D}(\mathbf{u}_D)$, where $\mathbf{u}_D = \left(F_{X_{i_1}}(x_{i_1}), \ldots, F_{X_{i_m}}(x_{i_m})\right)^T$ and $\{i_1, \ldots, i_m\} = D \subseteq V$.

It is well known that

$$f_{X_{i_1},\ldots X_{i_m}}(x_{i_1}, \ldots, x_{i_m}) = c_{X_{i_1},\ldots X_{i_m}}\left(F_{X_{i_1}}(x_{i_1}), \ldots, F_{X_{i_m}}(x_{i_m})\right) \cdot \prod_{k=1}^{m} f_{X_{i_k}}(x_{i_k}).$$

In vectorial terms this can be written as

$$f_{\mathbf{X}_D}(\mathbf{x}_D) = c_{\mathbf{X}_D}(\mathbf{u}_D) \cdot \prod_{i_k \in D} f_{X_{i_k}}(x_{i_k}) \tag{4}$$

and after dividing by the product term

$$c_{\mathbf{X}_D}(\mathbf{u}_D) = \frac{f_{\mathbf{X}_D}(\mathbf{x}_D)}{\prod\limits_{i_k \in D} f_{X_{i_k}}(x_{i_k})}.$$

In many applications occurs that between different pairs of variables there are different dependence structures ([1, 11]) which can not be modeled by a unique multivariate copula function. Therefore a new approach was introduced by Joe in [16]. This is the so called pair-copula construction which is able to encode more types of pair dependences in the same multivariate probability distribution. In this approach a copula is expressed as a product of different types of bivariate copulas and conditional bivariate copulas. A useful tool called R-vine structure was introduced for construction of pair-copulas by Bedford and Cooke in [4, 5] and described in more detail by Kurowicka and Cooke in [22]. This construction uses some special graph sequences.

If it does not cause confusion, instead of $f_{\mathbf{X}_D}$ and $c_{\mathbf{X}_D}$ we will write f_D and c_D. We also introduce the following notations:

- $F_{i,j|D}$ – the conditional probability distribution function of X_i and X_j given $\mathbf{X}_D$;
- $f_{i,j|D}$ – the conditional probability density function of X_i and X_j given $\mathbf{X}_D$,
- $c_{i,j|D}$ – the conditional copula density function corresponding to $f_{i,j|D}$,

where $D \subset V; i, j \in V \backslash D$.

According to [22], the definition of the R-vine graph structure is given as:

Definition 2.5. A *Regular-vine (R-vine) on d variables* consists of a sequence of trees $T_1, T_2, \ldots, T_{d-1}$ with nodes N_i and edges E_i for $i = 1, \ldots, d-1$, which satisfy the following conditions:

- T_1 has nodes $N_1 = \{1, \ldots, d\}$ and edges E_1.
- For $i = 2, \ldots, d-1$ the tree T_i has nodes $N_i = E_{i-1}$.
- For two hyperedges a and b to be joined in T_{i+1} it must hold that $\{a, b\} \in E_i$ and $|a \Delta b| = 2$, where Δ stands for the symmetric difference operator and $|\cdot|$ stands for the cardinality of the set.

We mention here that a and b are subsets of V and $|a| = |b| = i$ in the tree T_i.

The last condition usually is referred to as *proximity condition* that ensures that two nodes in tree T_{i+1} are only connected by an edge if these nodes share a common node in tree T_i.

It is shown in [4] and [22] that the edges in an R-vine tree can be uniquely identified by two nodes (the conditioned nodes), and a set of conditioning nodes, i.e., edges are denoted by $e = j(e), k(e) | D(E)$ where $D(E)$ is the conditioning set and $j(e), k(e) \notin D(E)$. For a good overview see [10].

The next theorem (see [4]), which can be regarded as a central theorem of R-vines, links the probability density function to the copulas assigned to the R-vine structure. In [4] it is shown that there exists a unique probability density assigned to a given R-vine structure. In [5] it is shown that this probability distribution can be expressed as (5).

Theorem 2.6. The joint density of $\mathbf{X} = (X_1, \ldots, X_d)$ is uniquely determined and given by:

$$\begin{aligned} f(x_1, \ldots, x_d) &= \left[\prod_{k=1}^{d} f_k(x_k)\right] \\ &\cdot \prod_{i=2}^{d-1} \prod_{e \in E_i} c_{j(e),k(e)|D(e)} \left(F_{j(e)|D(e)}\left(x_{j(e)}|\mathbf{x}_{D(e)}\right), F_{k(e)|D(e)}\left(x_{k(e)}|\mathbf{x}_{D(e)}\right) |\mathbf{x}_{D(e)}\right), \end{aligned} \tag{5}$$

where $F_{j(e)|D(e)}\left(x_{j(e)}|\mathbf{x}_{D(e)}\right)$ can be calculated as follows:

$$F_{j(e)|D(e)}\left(x_{j(e)}|\mathbf{x}_{D(e)}\right) = \left.\frac{\partial C_{i,j(e)|D(e)\setminus\{i\}}(u_i, u_j)}{\partial u_i}\right|_{\substack{u_i = F_{i|D(e)\setminus\{i\}}\left(x_i|\mathbf{x}_{D(e)\setminus\{i\}}\right) \\ u_j = F_{j(e)|D(e)\setminus\{i\}}\left(x_{j(e)}|\mathbf{x}_{D(e)\setminus\{i\}}\right)}}$$

for $i \in D(e)$.

Thus one can express $F_{j(e)|D(e)}\left(x_{j(e)}|\mathbf{x}_{D(e)}\right)$as a function of two conditional distributions $F_{i|D(e)\setminus\{i\}}\left(x_i|\mathbf{x}_{D(e)\setminus\{i\}}\right)$and $F_{j(e)|D(e)\setminus\{i\}}\left(x_{j(e)}|\mathbf{x}_{D(e)\setminus\{i\}}\right)$, with one less conditioning variable. This formula was given by Joe in [16]. Hence all conditional distribution functions in (5) are nested functions of the univariate marginal distribution functions. In (5) only pair-copulas are involved, therefore these constructions are also called pair-copula constructions.

We emphasize here that in general the parameter of the pair-copulas $c_{j(e),k(e)|D(e)}$ depends on the conditioning set $D(e)$. However in the case of simplified pair copula constructions, the parameter only depends on the conditioning set and not on the values of the corresponding random variables contained in the conditioning set.

In Subsection 2.4 we will give another definition for the R-vine which is related to a sequence of k order cherry-trees.

In [1] the inference of pair-copula decomposition is depicted in three parts:

- The selection of a specific factorization (structure);
- The choice of pair-copula types;
- The estimation of parameters of the chosen pair-copulas.

Our approach deals with finding a good factorization which exploits some of the conditional independences between the random variables.

Many papers are dealing with selecting specific R-vine such as C-vine or D-vine, see for example in [1].

2.3. The multivariate copula associated to a junction tree probability distribution. The cherry-tree copulas.

In this subsection we recall results published in [19] and [20]. In [19] a theorem assuring the existence of a special type of copula density was proved. This copula density was assigned to a junction tree graph structure. Let us consider a random vector $\mathbf{X} = (X_1, X_2, \ldots, X_d)^T$, with the set of indices $V = \{1, 2, \ldots, d\}$. Let (V, Γ, S) be a junction tree defined on the vertex set V, by the cluster set Γ, and the separator set S.

Theorem 2.7. (Kovács and Szántai [19]) The copula density function associated to a junction tree probability distribution

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\prod\limits_{K \in \Gamma} f_{\mathbf{X}_K}(\mathbf{x}_K)}{\prod\limits_{S \in \mathcal{S}} \left[f_{\mathbf{X}_S}(\mathbf{x}_S)\right]^{\nu_S - 1}},$$

is given by

$$c_{\mathbf{X}}(\mathbf{u}_V) = \frac{\prod\limits_{K \in \Gamma} c_{\mathbf{X}_K}(\mathbf{u}_K)}{\prod\limits_{S \in \mathcal{S}} \left[c_{\mathbf{X}_S}(\mathbf{u}_S)\right]^{\nu_S - 1}}, \tag{6}$$

where ν_S is the number of clusters linked by S.

Definition 2.8. (Kovács and Szántai [19]) The copula defined by (6) is called junction tree copula.

Definition 2.9. The junction tree copula associated to a junction tree with the largest cluster containing k elements is called k-width junction tree copula.

We saw that if the conditional independence structure between the random variables makes possible the construction of a junction tree, then the multivariate copula density associated to the joint probability distribution can be expressed as a fraction of some products of lower dimensional copula densities.

As a special case of Theorem 2.7 we state the following theorem for cherry-tree probability distributions.

Theorem 2.10. The copula density function associated to a cherry-tree probability distribution

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\prod\limits_{K \in \mathcal{C}_{ch}} f_{\mathbf{X}_K}(\mathbf{x}_K)}{\prod\limits_{S \in \mathcal{S}_{ch}} [f_{\mathbf{X}_S}(\mathbf{x}_S)]^{v_S - 1}},$$

is given by

$$c_{\mathbf{X}}(\mathbf{u}_V) = \frac{\prod\limits_{K \in \mathcal{C}_{ch}} c_{\mathbf{X}_K}(\mathbf{u}_K)}{\prod\limits_{S \in \mathcal{S}_{ch}} [c_{\mathbf{X}_S}(\mathbf{u}_S)]^{v_S - 1}}, \tag{7}$$

where $(V, \mathcal{C}_{ch}, \mathcal{S}_{ch})$ is a cherry-tree graph structure and v_S is the number of clusters linked by S.

Proof.

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\prod\limits_{K \in \mathcal{C}_{ch}} f_{\mathbf{X}_K}(\mathbf{x}_K)}{\prod\limits_{S \in \mathcal{S}_{ch}} [f_{\mathbf{X}_S}(\mathbf{x}_S)]^{v_S - 1}} = \frac{\prod\limits_{K \in \mathcal{C}_{ch}} c_{\mathbf{X}_K}(\mathbf{u}_K) \cdot \prod\limits_{i_k \in K} f_{X_{i_k}}(x_{i_k})}{\prod\limits_{S \in \mathcal{S}_{ch}} \left[c_{\mathbf{X}_S}(\mathbf{u}_S) \cdot \prod\limits_{i_k \in S} f_{X_{i_k}}(x_{i_k}) \right]^{v_S - 1}}. \tag{8}$$

The question that we have to answer is how many times appears in the nominator respectively in the denominator the probability density function $f_{X_i}(x_i)$ of each X_i random variable.

Since $\bigcup\limits_{K \in \mathcal{C}_{ch}} \mathbf{X}_K = X$ for each random variable $X_i \epsilon X$, $f_{X_i}(x_i)$ appears at least once in the nominator.

Now we prove that in the cherry-tree the number of clusters which contain a variable X_i is greater with 1 than the number of separators which contain the same variable. In a cherry-tree for every index i as a consequence of the running intersection property the following property holds. The clusters which contain a given index i define a sub cherry-tree. If the index i is contained in number of $t(i)$ clusters then i will be contained in $t(i) - 1$ separators as these are edges of the cluster-tree. Therefore for all i $f_{X_i}(x_i)$ will be contained once more times in the nominator than in the denominator.

Applying this result in formula (8) after simplification we obtain

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\prod\limits_{K \in \mathcal{C}_{ch}} c_{\mathbf{X}_K}(\mathbf{u}_K) \prod\limits_{i=1}^{d} f_{X_i}(x_i)}{\prod\limits_{S \in \mathcal{S}_{ch}} [c_{\mathbf{X}_S}(\mathbf{u}_S)]^{v_S - 1}}.$$

Dividing both sides by $\prod\limits_{i=1}^{d} f_{X_i}(x_i)$ we obtain (7). □

Definition 2.11. The copula density function associated to a cherry-tree probability distribution is called cherry-tree copula and its expression is:

$$c_{\mathbf{X}}\left(\mathbf{u}_V\right)=\frac{\prod\limits_{K\in\mathcal{C}_{ch}} c_{\mathbf{X}_K}\left(\mathbf{u}_K\right)}{\prod\limits_{S\in\mathcal{S}_{ch}}\left[c_{\mathbf{X}_S}\left(\mathbf{u}_S\right)\right]^{\nu_S-1}}, \tag{9}$$

where ν_S denotes the number of clusters containing the separator S.

2.4. R-vine structure given by a sequence of cherry-trees. Cherry-vine copula.

In [19] and [20] we gave an alternative definition of R-vines by using the concept of cherry-tree.

Definition 2.12. The cherry-vine graph structure is defined by a sequence of cherry junction trees $T_1, T_2, \dots, T_{d-1}$ as follows

- $T_1$is a regular tree on $V=\{1,\dots,d\}$, the set of edges is $E_1=\left\{e_i^1=(l_i,m_i)\right.$, $\left. i=1,\dots,d-1,\ l_i,m_i\in V\right\}$
- T_2 is the second order cherry junction tree on $V=\{1,\dots,d\}$, with the set of clusters $E_2=\left\{e_i^2, i=1,\dots,d-1 | e_i^2=e_i^1\right\}$, $\left|e_i^1\right|=2$
- T_k is one of the possible k order cherry junction tree on $V=\{1,\dots,d\}$, with the set of clusters $E_k=\left\{e_i^k, i=1,\dots,d-k+1\right\}$, where each e_i^k, $\left|e_i^k\right|=k$ is obtained from the union of two linked clusters in the $(k-1)$ order cherry junction tree T_{k-1}.

It is straightforward to see that Definition 2.12 is equivalent with Definition 2.5. Next we define the pair-copulas assigned to the cherry-vine structure given in Definition 2.12

The copula densities $c_{l_i,m_i}\left(F_{l_i}\left(x_{li}\right),F_{m_i}\left(x_{m_i}\right)\right)$ are assigned to the edges of the tree T_1.

The copula densities

$$c_{a_{ij}^l,b_{ij}^l|S_{ij}^l}\left(F_{a_{ij}^l|S_{ij}^l}\left(x_{a_{ij}^l}|\mathbf{x}_{S_{ij}^l}\right),F_{b_{ij}^l|S_{ij}^l}\left(x_{b_{ij}^l}|\mathbf{x}_{S_{ij}^l}\right)\,\middle|\,\mathbf{x}_{S_{ij}^l}\right)$$

are assigned to each pair of clusters e_i^l and e_j^l, which are linked in the junction tree T_l, where:

$$\begin{aligned} S^l &= e_i^l\cap e_j^l, \\ a_{ij}^l &= e_i^l-S_{ij}^l \\ b_{ij}^l &= e_i^l-S_{ij}^l, \end{aligned} \tag{10}$$

for $l=2,\dots,d-1$. It is easy to see that a_{ij}^l and b_{ij}^l, $l=2,\dots,d-1$ contain a single element only.

The following theorem is a consequence of Theorem 2.6.

Theorem 2.13. The probability distribution associated to the cherry-vine structure given in Definition 2.12 can be expressed as:

$$f\left(x_1,\ldots,x_d\right)=\left[\prod_{i=1}^{d} f_i\left(x_i\right)\right]\left[\prod_{\left(l_i m_i\right)\in E_1} c_{l_i m_i}\left(F_{l_i}\left(x_{l_i}\right),F_{m_i}\left(x_{m_i}\right)\right)\right]\cdot$$
$$\prod_{l=2}^{d-1}\prod_{e_i^l,e_j^l\in N\left(T_l\right)} c_{a_{i,j}^l,b_{i,j}^l|S_{ij}^l}\left(F_{a_{i,j}^l|S_{ij}^l}\left(x_{a_{i,j}^l}|\mathbf{x}_{S_{ij}^l}\right),F_{b_{i,j}^l|S_{ij}^l}\left(x_{b_{i,j}^l}|\mathbf{x}_{S_{ij}^l}\right)|\mathbf{x}_{S_{ij}^l}\right) \qquad (11)$$

where $e_i^l,e_j^l\in N\left(T_l\right)$ denotes that e_i^l,e_j^l are linked in the cherry tree T_l, and $S_{ij}^l,a_{i,j}^l,b_{i,j}^l$, are defined by (10) and $F_{a_{i,j}^l|S_{ij}^l}$ is defined in similar way as in Theorem 2.6.

The Formula (11) can be applied for simplified and also not simplified pair-copula constructions.

We illustrate these concepts on the following example in the case of simplified pair-copula constructions.

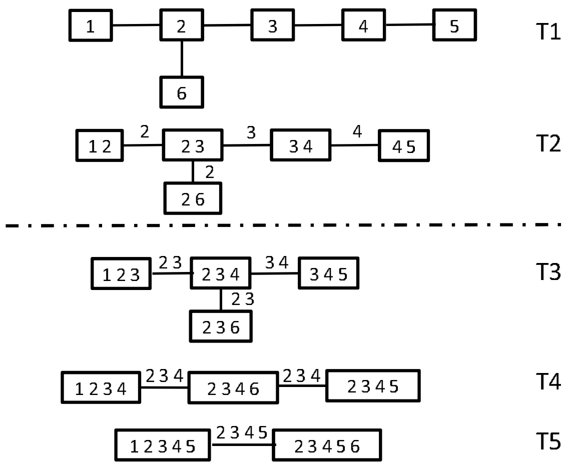

Fig. 2. Example for an R-vine structure on 6 variables using Definition 2.12.

Example 2.14. The edge set of the first tree and the sequence of the cherry-trees (in Figure 2) together with the copula densities determined by Definition 2.12 are the following:

$$T_1:\quad E_1=\left\{e_1^1=(1,2),e_2^1=(2,3),e_3^1=(2,6),e_4^1=(3,4),e_5^1=(4,5)\right\},$$
$$c_{e_1^1}=c_{1,2},c_{e_2^1}=c_{2,3},c_{e_3^1}=c_{2,6},c_{e_4^1}=c_{3,4},c_{e_5^1}=c_{4,5};$$
$$T_2:\quad E_2=\left\{e_1^2=(1,2),e_2^2=(2,3),e_3^2=(2,6),e_4^2=(3,4),e_5^2=(4,5)\right\}$$
$$S_{1,2}^2=e_1^2\cap e_2^2=\{2\},$$
$$a_{1,2}^2=e_1^2-S_{1,2}^2=\{1\},b_{1,2}^2=e_2^2-S_{1,2}^2=\{3\},c_{a_{1,2}^2,b_{1,2}^2|S_{1,2}^2}=c_{1,3|2};$$
$$S_{2,3}^2=e_2^2\cap e_3^2=\{2\},$$
$$a_{2,3}^2=e_2^2-S_{2,3}^2=\{3\},b_{2,3}^2=e_2^2-S_{2,3}^2=\{6\},c_{a_{2,3}^2,b_{2,3}^2|S_{2,3}^2}=c_{3,6|2};$$

$$S^2_{2,4} = e^2_2 \cap e^2_4 = \{3\},$$
$$a^2_{2,4} = e^2_2 - S^2_{2,4} = \{2\}, b^2_{2,4} = e^2_4 - S^2_{2,4} = \{4\}, c_{a^2_{2,4}, b^2_{2,4} | S^2_{2,4}} = c_{2,4|3};$$
$$S^2_{4,5} = e^2_4 \cap e^2_5 = \{4\},$$
$$a^2_{4,5} = e^2_4 - S^2_{4,5} = \{3\}, b^2_{4,5} = e^2_5 - S^2_{4,5} = \{5\}, c_{a^2_{4,5}, b^2_{4,5} | S^2_{4,5}} = c_{3,5|4};$$
$$T_3: \quad E_3 = \left\{ e^3_1 = (1,2,3), e^3_2 = (2,3,4), e^3_3 = (2,3,6), e^3_4 = (3,4,5) \right\}$$
$$S^3_{1,2} = e^3_1 \cap e^3_2 = \{2,3\},$$
$$a^3_{1,2} = e^3_1 - S^3_{1,2} = \{1\}, b^3_{1,2} = e^3_2 - S^3_{1,2} = \{4\}, c_{a^3_{1,2}, b^3_{1,2} | S^3_{1,2}} = c_{1,4|2,3};$$
$$S^3_{2,3} = e^3_2 \cap e^3_3 = \{2,3\},$$
$$a^3_{2,3} = e^3_2 - S^3_{2,3} = \{4\}, b^3_{2,3} = e^3_3 - S^3_{2,3} = \{6\}, c_{a^3_{2,3}, b^3_{2,3} | S^3_{2,3}} = c_{4,6|2,3};$$
$$S^3_{2,4} = e^3_2 \cap e^3_4 = \{3,4\},$$
$$a^3_{2,4} = e^3_2 - S^3_{2,4} = \{2\}, b^3_{2,4} = e^3_4 - S^3_{2,4} = \{5\}, c_{a^3_{2,4}, b^3_{2,4} | S^3_{2,4}} = c_{2,5|3,4};$$
$$T_4: \quad E_4 = \left\{ e^4_1 = (1,2,3,4), e^4_2 = (2,3,4,5), e^4_3 = (2,3,4,6) \right\}$$
$$S^4_{1,2} = e^4_1 \cap e^4_2 = \{2,3,4\},$$
$$a^4_{1,2} = e^4_1 - S^4_{1,2} = \{1\}, b^4_{1,2} = e^4_2 - S^4_{1,2} = \{5\}, c_{a^4_{1,2}, b^4_{1,2} | S^4_{1,2}} = c_{1,5|2,3,4};$$
$$S^3_{2,3} = e^4_2 \cap e^4_3 = \{2,3,4\},$$
$$a^4_{2,3} = e^4_2 - S^4_{2,3} = \{5\}, b^4_{2,3} = e^4_3 - S^4_{2,3} = \{6\}, c_{a^4_{2,3}, b^4_{2,3} | S^4_{2,3}} = c_{5,6|2,3,4};$$
$$T_5: \quad E_5 = \left\{ e^5_1 = (1,2,3,4,5), e^5_2 = (2,3,4,5,6) \right\}$$
$$S^5_{1,2} = e^5_1 \cap e^5_2 = \{2,3,4,5\},$$
$$a^5_{1,2} = e^5_1 - S^5_{1,2} = \{1\}, b^5_{1,2} = e^5_2 - S^5_{1,2} = \{6\}, c_{a^5_{1,2}, b^5_{1,2} | S^5_{1,2}} = c_{1,6|2,3,4,5}.$$

We draw the attention to the fact that the copulas assigned to the first tree T_1 are not conditional copulas in accordance with the formula (11).

The joint probability density function of $\mathbf{X} = (X_1, \ldots, X_6)$ can be expressed by Theorem 2.13 as follows:

$$
\begin{aligned}
& f(x_1, x_2, x_3, x_4, x_5, x_6) \\
& = \left(\prod_{i=1}^{6} f(x_i) \right) c_{1,2}(F_1(x_1), F_2(x_2)) \cdot c_{2,3}(F_2(x_2), F_3(x_3)) \cdot c_{2,6}(F_2(x_2), F_6(x_6)) \\
& \cdot c_{3,4}(F_3(x_3), F_4(x_4)) \\
& \cdot c_{4,5}(F_4(x_4), F_5(x_5)) \\
& \cdot c_{1,3|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \\
& \cdot c_{3,6|2}\left(F_{3|2}(x_3|x_2), F_{6|2}(x_6|x_2)\right) \\
& \cdot c_{2,4|3}\left(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)\right) \\
& \cdot c_{3,5|4}\left(F_{3|4}(x_3|x_4), F_{5|4}(x_5|x_4)\right) \\
& \cdot c_{1,4|2,3}\left(F_{1|2,3}(x_1|x_2, x_3), F_{4|2,3}(x_4|x_2, x_3)\right) \\
& \cdot c_{4,6|2,3}\left(F_{4|2,3}(x_4|x_2, x_3), F_{6|2,3}(x_6|x_2, x_3)\right) \\
& \cdot c_{2,5|3,4}\left(F_{2|3,4}(x_2|x_3, x_4), F_{5|3,4}(x_5|x_3, x_4)\right) \\
& \cdot c_{1,5|2,3,4}\left(F_{1|2,3,4}(x_1|x_2, x_3, x_4), F_{5|2,3,4}(x_5|x_2, x_3, x_4)\right) \\
& \cdot c_{5,6|2,3,4}\left(F_{5|2,3,4}(x_1|x_2, x_3, x_4), F_{6|2,3,4}(x_6|x_2, x_3, x_4)\right) \\
& \cdot c_{1,6|2,3,4,5}\left(F_{1|2,3,4,5}(x_1|x_2, x_3, x_4, x_5), F_{6|2,3,4,5}(x_6|x_2, x_3, x_4, x_5)\right)
\end{aligned}
$$

In this example we expressed the probability density function in a simplified form as in general each conditional pair copula depends on the conditioning variables (see (11)).

Here we call the attention that our R-vine representation by a sequence of cherry-trees was also used in a recent paper by Hobaeck–Haff et al. [15], Section 3. However in their paper it was not clearly declared that this representation was introduced in our paper [20].

In the following sections we give some theorems dealing with the relation between truncated R-vines and cherry-tree copulas.

3. TRUNCATED R-VINE AS A SPECIAL CASE OF CHERRY-TREE COPULA

In this section we give theorems highlighting the relation between the truncated R-vine and cherry-tree copulas.

As the number of variables grows, the number of conditional pair-copulas grows rapidly. For example in [11] in the case of 16 variables the number of pair-copulas involved, which have to be modeled and fitted is $120 = 15 + 14 + \cdots + 2 + 1$. To keep such structure tractable for inference and model selection, the simplifying assumption that copulas of conditional distributions do not depend on the variables which they are conditioned on is popular. Although this assumption leads in many cases to misspecifications as it was pointed out in [2] and [14]. In [13] there are presented classes of distributions where simplification is applicable. An idea to overcome the fitting of a large number of pair-copulas with large conditioning set is to exploit the conditional independences between the random variables. This idea was already discussed for Gaussian copulas in [22], based on the idea inspired by Whittaker [27]. However our approach is more general.

Aas et al. in [1] gave the relation between conditional independences and conditional pair-copulas:

Remark 3.1. X_i and X_j are conditionally independent given the set of variables $\mathbf{X}_A, A \subset V \setminus \{i,j\}$ if and only if

$$c_{ij|A}\left(F_{i|A}\left(x_i|\mathbf{x}_A\right), F_{j|A}\left(x_j|\mathbf{x}_A\right) |\ \mathbf{x}_A\right) = 1.$$

The following theorem is an important consequence of Theorem 2.6.

Theorem 3.2. If in an R-vine the conditional copula densities corresponding to the trees $T_k, T_{k+1}, \ldots, T_{d-1}$ are all equal to 1 then there exists a joint probability distribution which can be expressed only with the conditional copula densities assigned to $T_1, \ldots, T_{k-1}$:

$$f\left(x_1, \ldots, x_d\right) = \left[\prod_{i=1}^{d} f_i\left(x_i\right)\right] \left[\prod_{(l_i m_i) \in E_1} c_{l_i m_i}\left(F_{l_i}\left(x_{l_i}\right), F_{m_i}\left(x_{m_i}\right)\right)\right] \cdot$$
$$\prod_{l=2}^{k-1} \prod_{e_i^l, e_j^l \in N(T_l)} c_{a_{i,j}^l, b_{i,j}^l | S_{ij}^l}\left(F_{a_{i,j}^l | S_{ij}^l}\left(x_{a_{i,j}^l} | \mathbf{x}_{S_{ij}^l}\right), F_{b_{i,j}^l | S_{ij}^l}\left(x_{b_{i,j}^l} | \mathbf{x}_{S_{ij}^l}\right) | \mathbf{x}_{S_{ij}^l}\right)$$

where $e_i^l, e_j^l \in N\left(T_l\right)$ denotes that e_i^l, e_j^l are linked in the cherry tree T_l, and $S_{ij}^l, a_{i,j}^l, b_{i,j}^l$, are defined by (10) and $F_{a_{i,j}^l | S_{ij}^l}$ is defined in similar way as in Theorem 2.6.

The following definition of *truncated vine at level k* is given in [6].

Definition 3.3. A *pair-wisely truncated R-vine copula at level k* (or truncated R-vine copula at level k) is a special R-vine copula with the property that all pair-copulas with conditioning set equal to, or larger than k, are set to bivariate independence copulas.

We call the attention that Brechmann denotes the first tree T_0. To be consistent with our earlier notations we will denote the first tree by T_1.

In their approach Brechmann et al. [6] construct the truncated vine copulas by choosing the pairs of variables with the strongest Kendall's correlation in the first tree. In the last trees the pair-copulas were set to one. We claim that the strong dependences in the lower trees do not imply necessarily conditional independences in the last trees. In Figure 3 a conditional independence structure at level 3 is given which can not be achieved by a sequence of trees as a truncated R-vine.

Another approach, which is much closer to ours, is given by Kurowicka in [23]. Her idea was to build trees with lowest dependence (conditional independences) in the top trees, starting with the last tree (node). Her method uses partial correlations which in case of Gaussian copula are theoretically well grounded.

There arise the following questions. What special properties has the probability distribution, obtained by setting all the conditional copula densities associated to the trees $T_k, \ldots, T_{d-1}$ to 1 and what conditional independences are encoded in the obtained copula?

If the conditional copulas associated to the tree T_3 of Figure 3:

$$\begin{aligned} &c_{1,4|2,3}\left(F_{1|23}\left(x_1|x_2,x_3\right), F_{4|2,3}\left(x_4|x_2,x_3\right)\right), \\ &c_{4,6|2,3}\left(F_{1|23}\left(x_1|x_2,x_3\right), F_{4|2,3}\left(x_4|x_2,x_3\right)\right), \\ &c_{2,5|3,4}\left(F_{2,5|3,4}\left(x_2|x_3,x_4\right)\right), F_{5|3,4}\left(x_5|x_3,x_4\right)\right), \end{aligned} \tag{12}$$

are equal to 1, these imply the following conditional independences between the variables:

$$X_1 \perp X_4|X_2,X_3; \quad X_4 \perp X_6|X_2,X_3; \quad X_2 \perp X_5|X_3,X_4. \tag{13}$$

In this case the junction tree copula associated to T_3 in Figure 3 gives the expression of the multivariate copula as a cherry-tree copula.

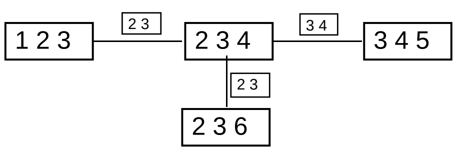

Fig. 3. 3-rd order cherry junction tree.

The cherry-tree copula density assigned to the truncated R-vine structure in Figure 2 is:

$$
\begin{aligned}
& f\left(x_1, x_2, x_3, x_4, x_5, x_6\right) \\
& =\left(\prod_{i=1}^{6} f\left(x_i\right)\right) \cdot c_{1,2}\left(F_1\left(x_1\right), F_2\left(x_2\right)\right) \cdot c_{2,3}\left(F_2\left(x_2\right), F_3\left(x_3\right)\right) \\
& \cdot c_{2,6}\left(F_2\left(x_2\right), F_6\left(x_6\right)\right) \cdot c_{3,4}\left(F_3\left(x_3\right), F_4\left(x_4\right)\right) \cdot c_{4,5}\left(F_4\left(x_4\right), F_5\left(x_5\right)\right) \\
& \cdot c_{1,3|2}\left(F_{1|2}\left(x_1|x_2\right), F_{3|2}\left(x_3|x_2\right)\right) \cdot c_{3,6|2}\left(F_{3|2}\left(x_3|x_2\right), F_{6|2}\left(x_6|x_2\right)\right) \\
& \cdot c_{2,4|3}\left(F_{2|3}\left(x_2|x_3\right), F_{4|3}\left(x_4|x_3\right)\right) \cdot c_{3,5|4}\left(F_{3|4}\left(x_3|x_4\right), F_{5|4}\left(x_5|x_4\right)\right).
\end{aligned}
$$

Let us recall the following results.

Theorem 3.4. (Kovács and Szántai [19]) A general k-width junction tree copula (see Definition 2.9) can be expressed as a k order cherry-tree copula.

This theorem shows why the k order cherry-tree copulas are so powerful in multivariate copula modeling.

Another important result is given by the following theorem.

Theorem 3.5. (Kovács and Szántai [19]) A k order cherry-tree copula can be expressed as a $(k+1)$ order cherry-tree copula.

As a consequence of Theorem 3.5 we have the following theorem.

Theorem 3.6. (Kovács and Szántai [19]) Any copula having a structure of truncated vine at level k is a k order cherry-tree copula.

Remark 3.7. Since truncated R-vine copula is a cherry-tree copula it can be defined by formula (9), where the set of clusters and separators are defined only by the first tree after truncation which we call top tree.

The top tree can be achieved in multiple ways, i. e. there are many sequences of cherry-trees leading to it, therefore in our opinion the good sequence of cherry trees is not necessarily the sequence which greedy way maximizes associations in the lower trees. The construction of the sequence of cherry trees in a greedy way may not result a good truncated vine at level k since the conditional independences in the higher trees will not necessarily be achieved. Rather we claim that it is more useful to choose that sequence which uses those pair-copulas which can be good fitted to the empirical data.

Theorem 3.8. A k order cherry-tree copula is a truncated R-vine copula if and only if its separators define a $(k-1)$ order cherry-tree.

Proof. The first implication is that if the separators of the tree T_k form a $(k-1)$ order cherry-tree, then the k order cherry-tree can be expressed as a truncated R-vine. For this statement we give a constructive proof by the following algorithm.

We will show that there exists a sequence of cherry-trees which leads to the given kth order cherry-tree. This means that the kth order cherry-tree is a truncated R-vine at level k.

Algorithm 3.9. Backward Algorithm.

Algorithm for obtaining a truncated R-vine structure from a cherry-tree structure.

Input: A k order cherry-tree graph structure, i. e. a set of clusters of size k and the set of separators of size k-1 enhanced with the property that the separators define a $k-1$ order cherry-tree.

Output: An R-vine truncated at level k.

We obtain recursively an $(m-1)$ width cherry-tree from an m-width cherry-tree, for $m = k, \dots, 1$ by the following two steps:

- Step 1. The separators of the m-width cherry-tree will be the clusters in the $(m-1)$-width cherry-tree, which will be linked if between them is one cluster in the m-width cherry-tree, and they are different.
- Step 2. The leaf clusters (those clusters which contain a simplicial node i. e. a node which is not contained by any other cluster) are transformed into $(m-1)$-width clusters, by deleting one node which is not simplicial. We emphasize here that it is essential to delete the same node from all leaf clusters which are connected to the same cluster. This guaranties that the $m-1$ order cherry-tree structure obtained is enhanced with the property that its separators define an $m-2$ order cherry-tree. The $m-1$-width cluster obtained in this way will be connected to one of the clusters obtained in Step 1, which was the $m-1$-width separator linked to it in the m-width cherry-tree.

An application of this algorithm can be seen in Figure 4.

Now we prove the other implication: If the k order cherry-tree copula can be expressed by an R-vine truncated at level k then the separators define a $(k-1)$ order cherry-tree. We prove this by proving an equivalent statement. If the separators do not define a $(k-1)$ order cherry-tree, then it cannot be expressed as an R-vine truncated at level k. We prove this on the example in Figure 5.

Let T_4 be the 4 order cherry-tree in Figure 5. Its separators do not define a 3-rd order cherry-tree. We will prove, that there does not exist any 3-rd order cherry-tree T_3 with the property that T_4 can be obtained from it by Defintion 2.12, which means that there does not exist a truncated R-vine structure which leads to it.

We will show that there does not exist a T_3 cherry-tree with clusters in E_3, such that the clusters in $E_4 = \{(1,2,3,5), (1,3,4,6), (1,2,3,4), (1,2,4,7)\}$, could be obtained by the union of two linked clusters belonging to E_3.

There are two possibilities:

1) The clusters $(1,2,3), (1,3,4), (1,2,4)$ are clusters of E_3. This cannot be the case because the running intersection property could not be fulfilled.

2) At least one of these clusters is not in E_3. Without loss of generality let us suppose that $(1,2,3)$ is not a cluster in E_3. This means that one of the pairs $(1,2), (2,3)$ and $(1,3)$ are not connected in T_3.

   Without loss of generality let us suppose that $(1,2)$ are not linked. By Definition 2.12 this means that $(1,2,3,5)$ in T_4 can be obtained from the union of $(1,3,5)$ and $(2,3,5)$ which are linked in T_3.

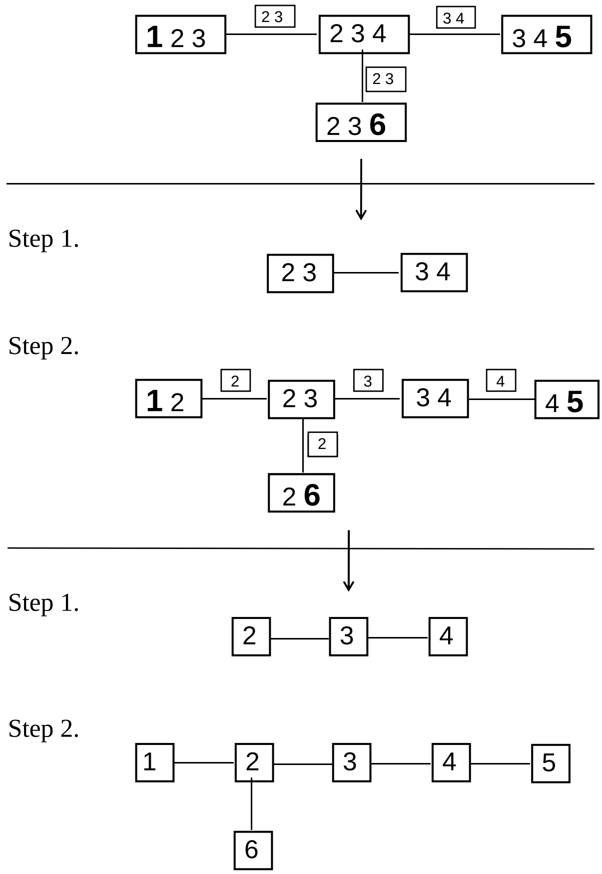

Fig. 4. Application of Algorithm 1 to a given 3-rd order cherry-tree in order to obtain a truncated R-vine at level 3 which leads to it.

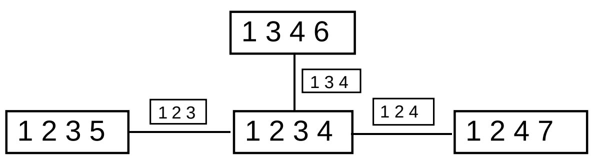

Fig. 5. A 4 order cherry-tree copula which cannot be achieved as a truncated R-vine.

Now there are two sub-cases again.

2a) $(1, 3, 5)$ is a leaf cluster (only one cluster is connected directly to it, in this case $(2, 3, 5)$). This leads to contradiction because 1 appears in at least one of other third order clusters, contained for example in $(1,2,4,7)$.

2b) $(1,3,5)$ is not a leaf cluster. In this case it is linked to another cluster by $(1,3),(1,5)$ or $(3,5)$. The other cluster has the form $(1,3,k),(1,5,k)$ or $(3,5,k)$, with $k \in \{4,6,7\}$. By Defintion 2.12 the clusters of T_4 are obtained by the union of the linked clusters in T_3. So by taking the union of $(1,3,5)$ with any of the clusters $(1,3,k),(1,5,k)$ or $(3,5,k)$ we obtain a 4 order cluster $(1,3,5,k)$, which also leads to contradiction because only one cluster of T_4 contains 5 but in this case we would have two clusters $(1,2,3,5)$ and $(1,3,5,k)$ both of them containing 5.

□

Definition 3.10. The truncated R-vine obtained by the Algorithm 3.9 (Backward Algorithm) started from a given cherry-tree as the top tree is called cherry-vine structure.

Remark 3.11. Algorithm 3.9 can produce more cherry-vine structures as in Step 2 we may proceed in different directions.

Remark 3.12. As it can be seen from the proof of Theorem 3.8 not every cherry-tree copula is a truncated vine copula.

Lemma 3.13. A necessary and sufficient condition for a cherry-tree copula to be a truncated R-vine copula is that each cluster has to be connected to its neighbors with at most two different separators.

Proof. First the necessity. If a cluster is connected to its neighbors by more than two different separators then the cherry-tree copula is not a truncated R-vine, see Figure 5.

Now the sufficiency. We want to prove that if a cherry-tree copula is a truncated R-vine copula then each cluster has at most two different separators. This is equivalent to the following. If a cluster of a cherry-tree has more than two different separators then it is not a truncated R-vine.

Let $\{i_1,\dots,i_k\}$ be an arbitrary cluster of a k order cherry-tree. Let us suppose that it is connected to its neighbors by three different separators. Without loss of generality let us denote these separators as follows: $S_{\setminus i_1}=\{i_2,\dots,i_k\}$, $S_{\setminus i_2}=\{i_1,i_3\dots,i_k\}$, $S_{\setminus i_3}=\{i_1,i_2,i_4\dots,i_k\}$. Any two of them will define a $(k-1)$ order cherry-tree but all three will not, since for any permutation of the three separators, there will be an element which do not fulfill the running intersection property. For example if the following connection is proposed

$$S_{\setminus i_1}-S_{\setminus i_2}-S_{\setminus i_3}$$

then i_2 occurs in the first and last set, but not in the set on the path between them. □

This relationship was also discussed in a recent paper by Hobaek–Haff et al. [15].

Lemma 3.13 can be used for checking whether a cherry-tree copula is or is not a truncated R-vine copula.

At this point we can conclude that the truncated vine at level k is a k order cherry-tree copula, but not every k order cherry-tree copula can be obtained as a truncated vine at level k.

4. METHODS FOR CONSTRUCTING CHERRY-TREE COPULAS AND TRUNCATED R-VINE COPULAS

Regarding Kurowicka's approach ([23]): "We start building the vine from the top node, and progress to the lower trees, ensuring that regularity condition is satisfied and partial correlations corresponding to these nodes are the smallest. If we assume that we can assign the independent copula to nodes of the vine with small absolute values of partial correlations, then this algorithm will be useful in finding an optimal truncation of a vine structure", we claim, that there are copulas which have conditional independences in the top trees ($m \geq k$), however they have not a truncated R-vine structure at level k.

We may have the following decomposition from the last node backward, which leads to the cherry-tree which is not truncated R-vine in Figure 6.

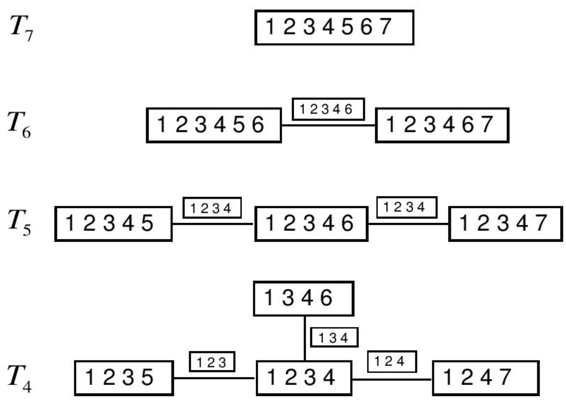

Fig. 6. The backward decomposition which leads to a 4th order cherry-tree, but not an R-vine truncated at level 4.

At this end it may arise the following question. How can we express a cherry-tree copula by two dimensional copulas? We have two possibilities:

First, if the cherry-tree copula is a truncated R-vine copula (the separators form a tree as we have seen in Theorem 3.8), then use Algorithm 3.9 to achieve a truncated vine structure, to which the pair-copulas will be assigned. In this case its formula is the following:

$$\frac{\prod\limits_{K\in\mathcal{C}_{ch}} c_K\left(F_{\mathbf{X}_K}(\mathbf{x}_K)\right)}{\prod\limits_{S\in\mathcal{S}_{ch}} \left[c_S\left(F_{\mathbf{X}_S}(\mathbf{x}_S)\right)\right]^{\nu_S-1}} = \left[\prod_{(l_i m_i)\in E_1} c_{l_i m_i}\left(F_{l_i}\left(x_{l_i}\right), F_{m_i}\left(x_{m_i}\right)\right)\right] \cdot$$
$$\prod_{l=2}^{k-1}\prod_{e_i^l,e_j^l\in N(T_l)} c_{a^l_{i,j},b^l_{i,j}|S^l_{ij}}\left(F_{a^l_{i,j}|S^l_{ij}}\left(x_{a^l_{i,j}}|\mathbf{x}_{S^l_{ij}}\right), F_{b^l_{i,j}|S^l_{ij}}\left(x_{b^l_{i,j}}|\mathbf{x}_{S^l_{ij}}\right)|\mathbf{x}_{S^l_{ij}}\right).$$

where $e_i^l, e_j^l \in N(T_l)$ denotes that e_i^l, e_j^l are linked in the cherry tree T_l, and $S^l_{ij}, a^l_{i,j}, b^l_{i,j}$, are defined by (10) and $F_{a^l_{i,j}|S^l_{ij}}$ is defined in similar way as in Theorem 2.6.

Second, if the cherry-tree copula is not a truncated R-vine copula the following theorem will be powerful for solving this problem.

Before stating the theorem it is important to call the attention on the following. Because the same cherry-tree can be represented graphically in multiple ways (when ν_S is greater than 1), it is important to start with a cherry-tree where the same separator links a cluster to other clusters. This means that from all clusters linked by the same separator we choose one and all the others will be linked to it. In this way all the other clusters will be neighbors of the chosen one.

Theorem 4.1. Starting from any k order cherry-tree copula the $(k+1)$ cherry-tree copula obtained by joining the neighbor clusters via Definition 2.12 will be a truncated R-vine copula.

Proof. Since via Lemma 3.13 we are interested only in the number of the different separators, we may suppose without loss of generality that all separators have multiplicity one.

We have two cases.

In the first case the k order cherry-tree copula is already a truncated R-vine copula, then by Definition 2.12 the obtained $k+1$ order cherry-tree copula is also a truncated R-vine.

In the second case we suppose that the k order cherry-tree copula is not a truncated R-vine copula. This means by Theorem 4.1 that the set of separators do not define a cherry-tree. Lemma 3.13 implies that there exists at least one cluster $C_k^* = \{i_1, i_2, \ldots, i_k\}$ which is connected to its neighbors by more than two different separators.

By joining the neighbor clusters in the k order cherry-tree using Definition 2.12 we obtain a $(k+1)$ order cherry-tree in which the separators correspond to the clusters in the k order tree. Let us consider the k-width clusters denoted by $C_1^k, \ldots, C_k^k$ such that $C_2^k, \ldots, C_k^k$ are all connected to C_1^k by different separators. (This will be a cherry-tree which is not truncated vine.) Let us form the clusters $C_i^{k+1} = C_1^k \cup C_i^k, i = 2, \ldots, k$. This way we obtain $k-1$ clusters of width $k+1$. Without loss of generality we fix one of them $C^\star = C_2^{k+1}$. $C^\star$ is connected to the other $C_i^{k+1}, i = 3, \ldots, k$ by C_1^k. On the other hand if in the original k-width tree C_2^k was connected to other clusters then $C^\star$ will be connected to them by C_2^k. At this end it turns out that $C^\star$ will be connected by at most 2 separators C_1^k and possibly C_2^k to the other clusters in the $k+1$-width tree.

For a better understanding let us illustrate this on an example. Let be $C_k^* = \{1, 2, 3, 4\}$ see picture a) of Figure 7. Then in picture b) the process of joining the neighbor clusters is presented. In picture c) of Figure 7 we have $C_{k+1}^* = \{1, 2, 3, 4, 5\}$ which has two different separators connected to it, one of them is $C_k^* = \{1, 2, 3, 4\}$ the other is $\{2, 3, 4, 5\}$.

We emphasize here, that by joining any two k order clusters in the k order cherry-tree we obtain a $(k+1)$ order cluster which will have at most two neighbor clusters in the $(k+1)$ order tree, connected to it by different separators. By Lemma 3.13 the $(k+1)$ tree obtained from the k order tree by Definition 2.12, will have a truncated R-vine structure. □

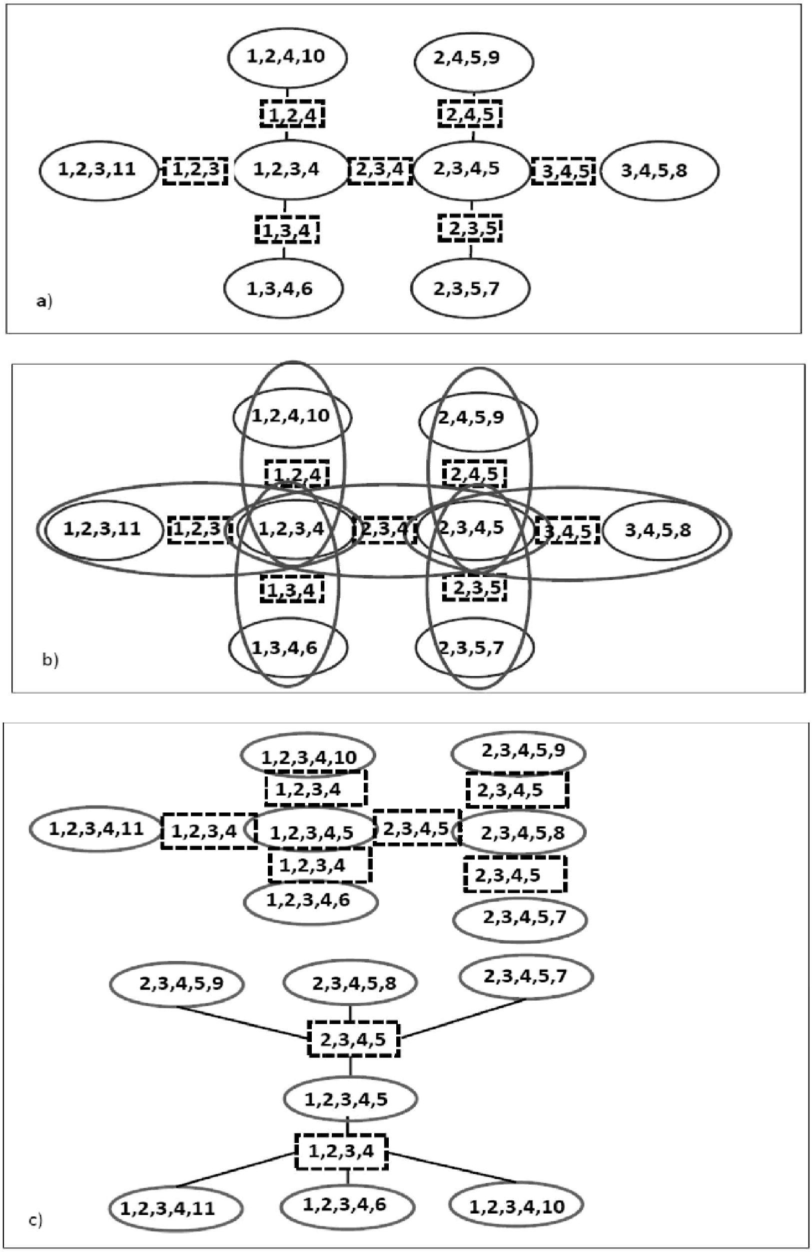

Fig. 7. a) A cherry-tree copula which is not truncated R-vine. b) Joining the neighboring clusters via Definition 2.12. c) The obtained 5th order cherry-tree which is a truncated R-vine in two representations.

We conclude this section with the following. It is easy to see how restrictive is to search for truncated R-vines only by building from bottom up the first trees in a greedy way. A truncated R-vine is defined by its top tree, the lower trees can be chosen by fulfilling the condition given in Definition 2.12. It is important to remark here that the number of possible sequences of trees is reduced by fixing the top tree. Therefore it may be combined with greedy methodologies in order to find a good fitting truncated R-vine copula.

The new idea is searching good fitting cherry-tree copulas and then to express it by a truncated R-vine copula. We proved in Theorem 4.1 that any cherry-tree copula can be transformed into a truncated R-vine which can be reached by using the Backward Algorithm.

5. CONCLUSIONS

In modeling multivariate probability distribution, an important task is to exploit some conditional independences existing between the random variables. We introduced in [26] and [17] the discrete cherry-tree probability distributions, then in [19] the cherry-tree copulas. The results of the present paper link the cherry-tree copula to the truncated R-vine which makes possible the use of cherry-tree structures in modeling continuous probability distributions, too.

If the number of variables grows the general R-vine copula modeling gets untractable. A method to overcome this problem is exploiting the conditional independences between the variables. The cherry-tree copulas are able to exploit these conditional independences. Another model containing conditional independences is the truncated R-vine. In the literature it was mainly fitted in greedy way from bottom to up. Truncated R-vines contain conditional independences but they do not use them in the model selection. This paper gives a possibility to overcome this drawback.

In this paper we clarify the relation between the cherry-tree copula and the truncated R-vine copula. The cherry-tree copula is more general than the truncated R-vine copula, but the truncated R-vine copula has the powerful property that it can be expressed by pair-wise copulas and pair-wise conditional copulas. We proved that a k order cherry-vine copula can be either expressed as a truncated R-vine copula at level k (by using the Backward Algorithm) or transformed into a $k+1$ order cherry-tree copula which can be expressed by a truncated vine copula at level $k+1$ (Theorem 4.1). In this way the cherry-tree copula gets also this powerful property.

In [19] we proved that any general k-width junction tree copula can be embedded in a k order cherry-tree copula. This shows the power of cherry-tree copulas related to general junction tree copulas.

Since the conditional independence structure of the truncated R-vine copula is completely characterized by the top tree (at a given level), in our opinion finding good truncated R-vines should be started by finding a good top tree (cherry-tree). A possible method for this, starting from a data set, is presented in [18]. Then one can construct the sequence of the cherry-trees which leads to the top cherry-tree. This is the so called cherry-vine structure.

We believe our approach may open a new perspective in modeling continuous multivariate probability distributions by exploiting the conditional independences between the components of the random vector. We challenge the vine copula community to search for good models from this perspective.

ACKNOWLEDGEMENT

We would like to thank the anonymous reviewers for their constructive comments and thoughtful questions. These have helped us greatly to improve our paper.

(Received July 22, 2016)

REFERENCES

[1] K. Aas, C. Czado, A. Frigessi, and H. Bakken: Pair-copula constructions of multiple dependence. Insur. Math. Econom. *44* (2009), 182–198. DOI:10.1016/j.insmatheco.2007.02.001

[2] E. F. Acar, C. Genest, and J. Nešlehová: Beyond simplified pair-copula constructions. J. Multivariate Anal. *110* (2012), 74–90. DOI:10.1016/j.jmva.2012.02.001

[3] A. Bauer, C. Czado, and T. Klein: Pair-copula construction for non-Gaussian DAG models. Canad. J. Stat. *40* (2012), 1, 86–109. DOI:10.1002/cjs.10131

[4] T. Bedford and R. Cooke: Probability density decomposition for conditionally dependent random variables modeled by vines. Ann. Math. Artif. Intell. *32* (2001), 245–268. DOI:10.1023/a:1016725902970

[5] T. Bedford and R. Cooke: Vines – a new graphical model for dependent random variables. Ann. Statist. *30* (2002), 4, 1031–1068. DOI:10.1214/aos/1031689016

[6] E. C. Brechmann, C. Czado, and K. Aas: Truncated regular vines in high dimensions with applications to financial data. Canad. J. Statist. *40* (2012), 1, 68–85. DOI:10.1002/cjs.10141

[7] J. Bukszár and A. Prékopa: Probability bounds with cherry trees. Math. Oper. Res. *26* (2001), 174–192. DOI:10.1287/moor.26.1.174.10596

[8] J. Bukszár and T. Szántai: Probability bounds given by hypercherry trees. Optim. Methods Software *17* (2002), 409–422. DOI:10.1080/1055678021000033955

[9] T. M. Cover and J. A. Thomas: Elements of Information Theory. Wiley Interscience, New York 1991. DOI:10.1002/0471200611

[10] C. Czado: Pair-copula constructions of multivariate copulas. In: Copula Theory and Its Applications (P. Jaworski, F. Durante, W. Härdle, and T. Rychlik, eds.), Springer, Berlin 2010. DOI:10.1007/978-3-642-12465-5_4

[11] J. Dissman, E. C. Brechmann, C. Czado, and D. Kurowicka: Selecting and estimating regular vine copulae and application to financial returns. Comput. Statist. Data Anal. *59* (2013), 52–69. DOI:10.1016/j.csda.2012.08.010

[12] A. Hanea, D. Kurowicka, and R. Cooke: Hybrid method for quantifying and analyzing Bayesian belief networks. Qual. Reliab. Engrg. *22* (2006), 708–729. DOI:10.1002/qre.808

[13] I. Hobaek Haff, K. Aas, and A. Frigessi: On the simplified pair-copula construction – simply useful or too simplistic? J. Multivariate Anal. *101* (2010), 5, 1296–1310. DOI:10.1016/j.jmva.2009.12.001

[14] I. Hobaek Haff and J. Segers: Nonparametric estimation of pair-copula constructions with the empirical pair-copula. 2010. http://arxiv.org/abs/1201.5133

[15] I. Hobaek-Haff, K. Aas, A. Frigessi, and V. Lacal: Structure learning in Bayesian Networks using regular vines. Computat. Statist. Data Anal. *101* (2016), 186–208. DOI:10.1016/j.csda.2016.03.003

[16] H. Joe: Multivariate Models and Dependence Concepts. Chapman and Hall, London 1997. DOI:10.1201/b13150

[17] E. Kovács and T. Szántai: On the approximation of discrete multivariate probability distribution using the new concept of t-cherry junction tree. Lect. Notes Economics Math. Systems *633*, Proc. IFIP/IIASA/GAMM Workshop on Coping with Uncertainty, Robust Solutions, 2008, IIASA, Laxenburg 2010, pp. 39–56. DOI:10.1007/978-3-642-03735-1_3

[18] E. Kovács and T. Szántai: Multivariate copula expressed by lower dimensional copulas. 2010.arXiv:1009.2898

[19] E. Kovács and T. Szántai: Hypergraphs in the characterization of regular-vine copula structures. In: Proc. 13th International Conference on Mathematics and its Applications, Timisoara 2012(a), pp. 335–344.

[20] E. Kovács and T. Szántai: Vine copulas as a mean for the construction of high dimensional probability distribution associated to a Markov network. 2012(b).arXiv:1105.1697

[21] D. Kurowicka and R. Cooke: The vine copula method for representing high dimensional dependent distributions: Application to continuous belief nets. In: Proc. 2002 Winter Simulation Conference 2002, pp. 270–278. DOI:10.1109/wsc.2002.1172895

[22] D. Kurowicka and R. M. Cooke: Uncertainty Analysis with High Dimensional Dependence Modelling. John Wiley, Chichester 2006. DOI:10.1002/0470863072

[23] D. Kurowicka: Optimal truncation of vines. In: Dependence-Modeling – Handbook on Vine Copulas (D. Kurowicka and H. Joe, eds.), Word Scientific Publishing, Singapore 2011.

[24] S. L. Lauritzen and D. J. Spiegelhalter: Local Computations with probabilites on graphical structures and their application to expert systems. J. Roy. Statist. Soc. B *50* (1988), 157–227.

[25] S. L. Lauritzen: Graphical Models. Clarendon Press, Oxford 1996.

[26] T. Szántai and E. Kovács: Hypergraphs as a mean of discovering the dependence structure of a discrete multivariate probability distribution. In: Proc. Conference Applied Mathematical Programming and Modelling (APMOD), Bratislava 2008, Ann. Oper. Res. *193* (2012), 1, 71–90. DOI:10.1007/s10479-010-0814-y

[27] J. Whittaker: Graphical Models in Applied Multivariate Statistics. John Wiley and Sons, 1990.

Edith Kovács, Department of Differential Equations, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3-9, Budapest, 1111. Hungary.

e-mail: kovacsea@math.bme.hu

Tamás Szántai, Department of Differential Equations, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3-9, Budapest, 1111. Hungary.

e-mail: szantai@math.bme.hu