Conditional Independence and Factorization of Multidimensional Models

Radim Jiroušek

***Abstract*— In this paper, three different frameworks for uncertainty description are considered: probability and possibility theories, and Dempster-Shafer theory of belief functions. For all of them special operators of composition are introduced, which enable, among others, defining the concept of factorization (used here as an alternative notion for conditional independence) meeting all the semigraphoid axioms. It is showed that whilst for probability and possibility theories factorization and conditional independence coincide, they differ from each other for belief functions. Since the introduced factorization manifests most of the properties required for the concept of conditional independence, the question arises whether it would be useful to substitute the often used concept of the conditional independence with the factorization introduced in this paper.**

I. Introduction

PERHAPS the main reason, why about 25 years ago the concept of conditional independence got into the center of research of so many scholars, is the fact that it enables efficient representation of multidimensional probability distributions: multidimensional models. Namely, given its conditional independence structure (i.e. a list of all the conditional independence relations that hold true) the considered probability distribution is uniquely specified by a system of its marginal distributions or conditional low-dimensional distributions. Naturally, such a system of low-dimensional distributions can be represented by a much smaller number of parameters (probabilities) than the considered multidimensional distribution.

So, under the assumption of validity of all the conditional independence relations that can be read from the respective acyclic directed graph, the distribution represented by a Bayesian network is uniquely specified by a system of conditional distributions. Similarly, for graphical models, assuming that all the conditional independence relations specified by the separation criterion hold true, one gets that the respective multidimensional probability distribution is uniquely determined by its marginal distributions corresponding to cliques of the underlying graph.

The main difference between these two just mentioned models consists in the fact that whereas for Bayesian networks there exists an explicit formula how to compute the respective multidimensional distribution (it is a simple product of the considered conditional distributions) no such a formula exists for general (cyclic) graphical models; linear programming or iterative fitting approaches must be used. An existence of such an explicit formula is usually based on the fact that variables X and Y are conditionally independent given variable Z if and only if $p(X,Y,Z) = p(X,Z) \cdot p(Y|Z)$. The mentioned unique specification by its marginals and the existence of the explicit formula inspired us about 10 years ago to introduce an *operator of composition* that, if applied iteratively, describes how multidimensional distributions can be computed from its low-dimensional marginals. Later, the operator of composition was introduced in possibility theory and recently also for belief functions. In this paper we briefly recall all three definitions and their connection to the concept of conditional independence. In the main part of the paper we shall show that only 6 simple properties must hold for the operator of composition to guarantee that the corresponding relation of conditional independence, we will call it *factorization* here, meets the *semigraphoid* properties, which are generally accepted as axioms, which should hold true for any relation of conditional independence (irrelevance, non-interactivity).

R. Jiroušek is with the Faculty of Management, University of Economics, and the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, (email: radim@utia.cas.cz).

This work was supported by MŠMT ČR under grants 1M0572 and 2C06019.

II. Operators of Composition - Notation

In the whole paper we shall deal with a finite number of variables $X_1, X_2, \dots, X_n$ each of which is specified by a finite set $\mathbf{X}_i$ of its values. So, we will consider multidimensional set (space)

$$\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \dots \times \mathbf{X}_n,$$

and its *subspaces*. For $K \subset N = \{1, 2, \dots, n\}$, $\mathbf{X}_K$ denotes a Cartesian product of those $\mathbf{X}_i$, for which $i \in K$:

$$\mathbf{X}_K = \bigtimes_{i \in K} \mathbf{X}_i,$$

and $X_K = \{X_i\}_{i \in K}$ denotes the set of the respective variables.

A *projection* of $x = (x_1, x_2, \dots, x_n) \in \mathbf{X}_N$ into $\mathbf{X}_K$ will be denoted $x^{\downarrow K}$, i.e. for $K = \{i_1, i_2, \dots, i_\ell\}$

$$x^{\downarrow K} = (x_{i_1}, x_{i_2}, \dots, x_{i_\ell}) \in \mathbf{X}_K.$$

Analogously, for $K \subset L \subseteq N$ and $A \subset \mathbf{X}_L$, $A^{\downarrow K}$ will denote a *projection* of A into $\mathbf{X}_K$:

$$A^{\downarrow K} = \{y \in \mathbf{X}_K : \exists x \in A \ \ (y = x^{\downarrow K})\}.$$

Let us remark that we do not exclude situations when $K = \emptyset$. In this case $A^{\downarrow \emptyset} = \emptyset$.

In addition to the projection, in this text we will need also the opposite operation which will be called extension. By

978-1-4244-1819-0/08/$25.00©2008 IEEE

an *extension* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ we will understand a set

$$A \otimes B = \{x \in \mathbf{X}_{K\cup L} : x^{\downarrow K} \in A \ \ \& \ \ x^{\downarrow L} \in B\}.$$

Notice that if K and L are disjoint then their extension is just their Cartesian product

$$A \otimes B = A \times B.$$

If $K \cap L \neq \emptyset$ and $A^{\downarrow K\cap L} \cap B^{\downarrow K\cap L} = \emptyset$ then also $A \otimes B = \emptyset$.

A. Probability Distributions

Let us start considering probability distributions on $\mathbf{X}_N$ and its subspaces $\mathbf{X}_K$ ($K \subseteq N$). For $L \subseteq K$ and a probability distribution p on $\mathbf{X}_K$, i.e.

$$p : \mathbf{X}_K \longrightarrow [0, 1],$$

for which

$$\sum_{x\in\mathbf{X}_K} p(x) = 1,$$

symbol $p^{\downarrow L}$ will denote its *marginal* distribution defined on $\mathbf{X}_L$. It is defined for all $x \in \mathbf{X}_L$ by the expression

$$p^{\downarrow L}(x) = \sum_{y\in\mathbf{X}_K : y^{\downarrow L}=x} p(y).$$

As usually, we will also speak about probability $p(A)$ of a set $A \subseteq \mathbf{X}_K$, which is a sum of probabilities of the respective points

$$p(A) = \sum_{x\in A} p(x).$$

Realize that $p(\emptyset) = 0$, but $p^{\downarrow\emptyset}(\emptyset) = 1$.

Consider three disjoint sets $I, J, K \subset N$ ($I \neq \emptyset \neq J$). We say that for distribution p groups of variables X_I and X_J are *conditionally independent given variables* X_K (in symbol $X_I \perp\!\!\!\perp X_J|X_K[p]$) if for all $x \in \mathbf{X}_{I\cup J\cup K}$ the following equality holds true

$$p^{\downarrow I\cup J\cup K}(x) \cdot p^{\downarrow K}(x^{\downarrow K}) = p^{\downarrow I\cup K}(x^{\downarrow I\cup K}) \cdot p^{\downarrow J\cup K}(x^{\downarrow J\cup K}).$$

It is well known that this is equivalent to the fact that

$$p^{\downarrow I\cup J\cup K}(x) = p^{\downarrow I\cup K}(x^{\downarrow I\cup K}) \cdot p^{\downarrow J\cup K}(x^{\downarrow J}|x^{\downarrow K}).$$

From two low-dimensional distributions p_1 and p_2 one can get a distribution of a higher dimension with the help of the following operator of composition.

Definition 1: Consider arbitrary two distributions p_1 and p_2 defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}$, respectively ($K_1 \neq \emptyset \neq K_2$). If $p_1^{\downarrow K_1\cap K_2}$ is dominated by $p_2^{\downarrow K_1\cap K_2}$, i.e.

$$\forall z \in \mathbf{X}_{K_1\cap K_2} \quad p_2^{\downarrow K_1\cap K_2}(z) = 0 \Longrightarrow p_1^{\downarrow K_1\cap K_2}(z) = 0,$$

then $p_1 \triangleright p_2$ is for all $x \in \mathbf{X}_{K\cup L}$ defined by the expression

$$(p_1 \triangleright p_2)(x) = \frac{p_1(x^{\downarrow K_1}) \cdot p_2(x^{\downarrow K_2})}{p_2^{\downarrow K_1\cap K_2}(x^{\downarrow K_1\cap K_2})}.$$

Otherwise the composition $p_1 \triangleright p_2$ remains undefined.

What is the result of composition of two distributions p_1 and p_2? The basic answer to this question is given by the following simple assertion.

Lemma 1: Consider three probability distributions p_1, p_2, p_3, defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \mathbf{X}_{K_3}$, respectively. If $p_1 \triangleright p_2$ is defined (in case of point (iv) we assume that $(p_1 \triangleright p_2) \triangleright p_3$ is defined) then

(i) *$p_1 \triangleright p_2$ is a probability distribution on $\mathbf{X}_{K_1\cup K_2}$;*
(ii) *$(p_1 \triangleright p_2)^{\downarrow K_1} = p_1$;*
(iii) *$p_1 \triangleright p_2 = p_2 \triangleright p_1 \iff p_1^{\downarrow K_1\cap K_2} = p_2^{\downarrow K_1\cap K_2}$;*
(iv) *$K_1 \supseteq (K_2 \cap K_3) \implies (p_1 \triangleright p_2) \triangleright p_3 = (p_1 \triangleright p_3) \triangleright p_2$;*
(v) *$K_2 \supseteq L \supseteq (K_1 \cap K_2) \implies p_1 \triangleright p_2 = (p_1 \triangleright p_2^{\downarrow L}) \triangleright p_2$;*
(vi) *$(K_1 \cup K_2) \supseteq L \supseteq K_1$*
$\implies (p_1 \triangleright p_2)^{\downarrow L} = p_1 \triangleright p_2^{\downarrow K_2\cap L}$.

All these properties were proved in our preceding papers [4], [6], nevertheless most of them follow immediately from the fact that if $p_1 \triangleright p_2$ is defined then

$$(p_1 \triangleright p_2)(x) = p_1(x^{\downarrow K_1}) \cdot p_2(x^{\downarrow K_2\setminus K_1}|x^{\downarrow K_1\cap K_2}).$$

From this equality one can also immediately see the property, which is of great importance from the point of view of this paper, and which is expressed by the following assertion (its proof can also be found in [4]).

Lemma 2: Let I, J, K be disjoint subsets of N, I and J be nonempty. For a probability distribution p defined on $\mathbf{X}_N$

$$p^{\downarrow I\cup J\cup K} = p^{\downarrow I\cup K} \triangleright p^{\downarrow J\cup K}$$

if and only if

$$X_I \perp\!\!\!\perp X_J|X_K[p].$$

The lemma reads that for probability distributions we could define a concept of conditional independence on the basis of factorization: Variables X_I and X_J are conditionally independent given variables X_K for distribution p if and only if the marginal distribution $p^{\downarrow I\cup J\cup K}$ *factorizes* in the following sense

$$p^{\downarrow I\cup J\cup K} = p^{\downarrow I\cup K} \triangleright p^{\downarrow J\cup K}.$$

B. Possibility Distributions

To distinguish possibility distributions from probability ones, we will denote possibility distributions by Greek character π (with possible indices). In analogy to a probability distribution, possibility distribution π on $\mathbf{X}_K$ is also a mapping

$$\pi : \mathbf{X}_K \longrightarrow [0, 1].$$

In this paper we will consider only *normal* possibility distributions, i.e. distributions π for which

$$\max_{x\in\mathbf{X}_K} \pi(x) = 1.$$

In a way closely connected with the notion of normalization is the most important difference between probabilistic and possibilistic settings, which concerns marginalization.

Marginalization in possibility theory differs from that in the probabilistic framework in using maximization instead of summation. For $J \subset K$ a marginal possibility distribution $\pi^{\downarrow J}$ of distribution π (which is assumed to be defined on $\mathbf{X}_K$) is defined for all $x \in \mathbf{X}_L$ by the formula

$$\pi^{\downarrow L}(x) = \max_{y \in \mathbf{X}_K : y^{\downarrow L} = x} \pi(y).$$

In analogy to this, possibility $\pi(A)$ of a set $A \subseteq \mathbf{X}_K$ is got from the respective possibility distribution π defined on $\mathbf{X}_K$ in the following way

$$\pi(A) = \max_{x \in A} \pi(x).$$

Since conditioning as well as the concept of independence in possibility theory are closely connected with t-norms, it is quite natural that also operator of composition is parameterized by a t-norm.

Definition 2: A *triangular norm* (or a *t-norm*) T is a binary operator on $[0,1]$ (i.e. $T : [0,1]^2 \to [0,1]$) satisfying the following three conditions:

- for any $x \in [0,1]$, $T(1,x) = x$;
- for any $x_1, x_2, y_1, y_2 \in [0,1]$ such that $x_1 \leq x_2, y_1 \leq y_2$
$$T(x_1, y_1) \leq T(x_2, y_2);$$
- for any $x, y, z \in [0,1]$, $T(T(x,y),z) = T(x, T(y,z))$, and $T(x,y) = T(y,x)$.

In this paper we shall deal only with *continuous t*-norms, i.e. with t-norms which are continuous functions. The reader not familiar with t-norms can consider in the following text only the simplest t-norms: either *Gödel's t-norm* $T_G(x,y) = \min(x,y)$, or *product t-norm* $T_p(x,y) = x \cdot y$;

Consider possibility distribution π defined on $\mathbf{X}_K$ and two disjoint nonempty subsets I, J of K. The conditional distribution $\pi^{\downarrow I \cup K}\left(x^{\downarrow I}|_T x^{\downarrow J}\right)$ is defined by

$$\pi^{\downarrow I \cup K}\left(x^{\downarrow I}|_T x^{\downarrow J}\right) = \sup\{z \in [0,1] : T(z, \pi^{\downarrow I}\left(x^{\downarrow I}\right)) \leq \pi^{\downarrow I \cup J}\left(x^{\downarrow I \cup J}\right)\}.$$

In the case that

$$\pi^{\downarrow I \cup J}\left(x^{\downarrow I \cup J}\right) = T\left(\pi^{\downarrow I}\left(x^{\downarrow I}\right), \pi^{\downarrow J}\left(x^{\downarrow J}\right)\right),$$

the two groups of variables X_I and X_J are said (for distribution π) to be *T-independent*.

To be able to introduce conditional independence let us consider three disjoint subsets I, J and K ($I \neq \emptyset \neq J$). For a possibility distribution π two groups of variables X_I and X_J are *conditionally T-independent* given the third group X_K if

$$\pi^{\downarrow I \cup J \cup K}(x^{\downarrow I \cup J \cup K}) = T(\pi^{\downarrow I \cup K}(x^{\downarrow I \cup K}), \pi^{\downarrow J \cup K}(x^{\downarrow J}|_T x^{\downarrow K})).$$

This property will be denoted by $X_I \perp\!\!\!\perp_T X_J | X_K[\pi]$.

In [15], Vejnarová defined a possibilistic version of the operator of composition.

Definition 3: Consider arbitrary two possibility distributions π_1 and π_2 defined on $\mathbf{X}_{K_1}$ and $\mathbf{X}_{K_2}$, respectively ($K_1 \neq \emptyset \neq K_2$). For an arbitrary continuous t-norm T their *composition* $\pi_1 \triangleright_T \pi_2$ is defined for all $x \in \mathbf{X}_{K \cup L}$ by the following expression

$$(\pi_1 \triangleright_T \pi_2)(x) = T(\pi_1(x^{\downarrow K_1}), \pi_2(x^{\downarrow K_2 \setminus K_1}|x^{\downarrow K_1 \cap K_2})).$$

Let us highlight the main difference between probabilistic and possibilistic operators of composition: whereas in probability theory the operator of composition may be undefined for a couple of probability distributions, in possibility theory the result of composition is always defined.

From the point of view of this paper the most important are the properties of the possibilistic operator of composition formulated in the following two Lemmas, which were proved by Vejnarová in [15], [16] and [17].

Lemma 3: For arbitrary three possibility distributions π_1, π_2, π_3, defined on $\mathbf{X}_{K_1}, \mathbf{X}_{K_2}, \mathbf{X}_{K_3}$, respectively, and a continuous t-norm T the following six properties hold true:

(i) $\pi_1 \triangleright_T \pi_2$ *is a possibility distribution on* $\mathbf{X}_{K_1 \cup K_2}$;

(ii) $(\pi_1 \triangleright_T \pi_2)^{\downarrow K_1} = \pi_1$;

(iii) $\pi_1 \triangleright_T \pi_2 = \pi_2 \triangleright_T \pi_1 \iff \pi_1^{\downarrow K_1 \cap K_2} = \pi_2^{\downarrow K_1 \cap K_2}$;

(iv) $K_1 \supseteq (K_2 \cap K_3)$
$\implies (\pi_1 \triangleright_T \pi_2) \triangleright_T \pi_3 = (\pi_1 \triangleright_T \pi_3) \triangleright_T \pi_2$;

(v) $K_2 \supseteq L \supseteq (K_1 \cap K_2)$
$\implies \pi_1 \triangleright_T \pi_2 = (\pi_1 \triangleright_T \pi_2^{\downarrow L}) \triangleright_T \pi_2$;

(vi) $(K_1 \cup K_2) \supseteq L \supseteq K_1$
$\implies (\pi_1 \triangleright_T \pi_2)^{\downarrow L} = \pi_1 \triangleright_T \pi_2^{\downarrow K_2 \cap L}$.

Lemma 4: Let I, J, K be disjoint subsets of N, I and J be nonempty. For a possibility distribution π defined on $\mathbf{X}_N$

$$\pi^{\downarrow I \cup J \cup K} = \pi^{\downarrow I \cup K} \triangleright_T \pi^{\downarrow J \cup K}$$

if and only if

$$X_I \perp\!\!\!\perp_T X_J | X_K[\pi].$$

The last lemma reformulates for possibility distributions what was said for probability distributions: variables X_I and X_J are conditionally T-independent given variables X_K for possibility distribution π if and only if the marginal distribution $\pi^{\downarrow I \cup J \cup K}$ *factorizes* in the following sense

$$\pi^{\downarrow I \cup J \cup K} = \pi^{\downarrow I \cup K} \triangleright_T \pi^{\downarrow J \cup K}.$$

C. Belief Functions

A belief function is defined with the help of a *basic* (*probability or belief*) *assignment* m on $\mathbf{X}_N$, i.e.

$$m : \mathcal{P}(\mathbf{X}_N) \longrightarrow [0,1]$$

for which

$$\sum_{A \subseteq \mathbf{X}_N} m(A) = 1.$$

Therefore, for the sake of simplicity, we will not speak about belief functions but about basic assignments: We shall marginalize and compose basic assignments. For each $K \subset N$ *marginal basic assignment* of m is defined (for each $B \subseteq \mathbf{X}_K$):

$$m^{\downarrow K}(B) = \sum_{A \subseteq \mathbf{X}_N : A^{\downarrow K} = B} m(A).$$

TABLE I

BASIC ASSIGNMENTS m_1 AND m_2.

$A \subseteq \mathbf{X}_1$	$m_1(A)$
$\{a\}$	0.2
$\{\bar{a}\}$	0.3
$\{a\bar{a}\}$	0.5

$B \subseteq \mathbf{X}_2$	$m_2(B)$
$\{b\}$	0.6
$\{\bar{b}\}$	0
$\{a\bar{b}\}$	0.4

An operator of composition was for basic assignments defined in [8] by the following definition.

Definition 4: For two arbitrary basic assignments m_1 on $\mathbf{X}_{K_1}$ and m_2 on $\mathbf{X}_{K_2}$ ($K_1 \neq \emptyset \neq K_2$) a *composition* $m_1 \triangleright m_2$ is defined for each $C \subseteq \mathbf{X}_{K_1 \cup K_2}$ by one of the following expressions:

[a] if $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) > 0$ and $C = C^{\downarrow K_1} \otimes C^{\downarrow K_2}$ then

$$(m_1 \triangleright m_2)(C) = \frac{m_1(C^{\downarrow K_1}) \cdot m_2(C^{\downarrow K_2})}{m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2})};$$

[b] if $m_2^{\downarrow K_1 \cap K_2}(C^{\downarrow K_1 \cap K_2}) = 0$ and $C = C^{\downarrow K_1} \times \mathbf{X}_{K_2 \setminus K_1}$ then

$$(m_1 \triangleright m_2)(C) = m_1(C^{\downarrow K_1});$$

[c] in all other cases $(m_1 \triangleright m_2)(C) = 0$.

Since one of the goals of this paper is to show that this operator enables us to define for belief functions an innovative version of the relation of conditional independence, we will, in agreement with the requirements of an anonymous referee, illustrate its properties on a couple of examples.

Example 1: Consider two basic assignments m_1 and m_2 on $\mathbf{X}_1 = \{a, \bar{a}\}$ and $\mathbf{X}_2 = \{b, \bar{b}\}$, respectively, which are specified in Table I. Since, in this case, m_1 and m_2 are defined for disjoint sets of variables ($K_1 \cap K_2$ is empty), composition simplifies to the expression

$$(m_1 \triangleright m_2)(C) = m_1(C^{\downarrow\{1\}}) \cdot m_2(C^{\downarrow\{2\}}),$$

which is to be understood exactly in the sense of Definition 4: for all C such that $C = C^{\downarrow\{1\}} \otimes C^{\downarrow\{2\}}$ it is defined by the product $m_1(C^{\downarrow\{1\}}) \cdot m_2(C^{\downarrow\{2\}})$, for all the other C it is 0 (see also Table II).

With respect to the main purpose of this paper, it is relevant to notice the following fact. Consider, for example, set $C = \{ab, \bar{a}\bar{b}\} \neq C^{\downarrow 1} \otimes C^{\downarrow 2} = \{a, \bar{a}\} \otimes \{b, \bar{b}\} = \mathbf{X}_1 \otimes \mathbf{X}_2$. In this case Definition 4 assigns $m_1 \triangleright m_2(C) = 0$. If any positive value were assigned to this set C, it would express that one gives a part of her belief either $a \wedge b$ or $\bar{a} \wedge \bar{b}$. This means that one believes that there is a type of dependence between variables X_1 and X_2.

Using Table II, where the values of $m_1 \triangleright m_2$ are presented, the reader can easily check also other properties expected for the composition; for example that $m_1 = (m_1 \triangleright m_2)^{\downarrow\{1\}}$, and since m_1 and m_2 are trivially projective (consistent) also $m_2 = (m_1 \triangleright m_2)^{\downarrow\{2\}}$ (see property (iii) of Lemma 5 below). ■

TABLE II

BASIC ASSIGNMENT $m_1 \triangleright m_2$ FROM EXAMPLE 1.

$C \subseteq \mathbf{X}_{\{1,2\}}$	$C = C^{\downarrow\{1\}} \otimes C^{\downarrow\{2\}}$	$(m_1 \triangleright m_2)(C)$
$\{ab\}$	$\{a\} \otimes \{b\}$	0.12
$\{a\bar{b}\}$	$\{a\} \otimes \{\bar{b}\}$	0
$\{\bar{a}b\}$	$\{\bar{a}\} \otimes \{b\}$	0.18
$\{\bar{a}\bar{b}\}$	$\{\bar{a}\} \otimes \{\bar{b}\}$	0
$\{ab, a\bar{b}\}$	$\{a\} \otimes \mathbf{X}_2$	0.08
$\{ab, \bar{a}b\}$	$\mathbf{X}_1 \otimes \{b\}$	0.3
$\{ab, \bar{a}\bar{b}\}$		0
$\{a\bar{b}, \bar{a}b\}$		0
$\{a\bar{b}, \bar{a}\bar{b}\}$	$\mathbf{X}_1 \otimes \{\bar{b}\}$	0
$\{\bar{a}b, \bar{a}\bar{b}\}$	$\{\bar{a}\} \otimes \mathbf{X}_2$	0.12
$\{ab, a\bar{b}, \bar{a}b\}$		0
$\{ab, a\bar{b}, \bar{a}\bar{b}\}$		0
$\{ab, \bar{a}b, \bar{a}\bar{b}\}$		0
$\{a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$		0
$\{ab, a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$	$\mathbf{X}_1 \otimes \mathbf{X}_2$	0.2

TABLE III

BASIC ASSIGNMENTS $m_1(x_{\{1,2\}})$ AND $m_2(x_{\{2,3\}})$.

$C \subseteq \mathbf{X}_{\{1,2\}}$	$m_1(C)$
$\{ab\}$	0.1
$\{a\bar{b}\}$	0.5
$\{\bar{a}b\}$	0.2
$\{\bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}\}$	0
$\{ab, \bar{a}b\}$	0
$\{ab, \bar{a}\bar{b}\}$	0
$\{a\bar{b}, \bar{a}b\}$	0
$\{a\bar{b}, \bar{a}\bar{b}\}$	0
$\{\bar{a}b, \bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}, \bar{a}b\}$	0
$\{ab, a\bar{b}, \bar{a}\bar{b}\}$	0
$\{ab, \bar{a}b, \bar{a}\bar{b}\}$	0
$\{a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$	0
$\{ab, a\bar{b}, \bar{a}b, \bar{a}\bar{b}\}$	0.2

$C \subseteq \mathbf{X}_{\{2,3\}}$	$m_2(C)$
$\{bc\}$	0
$\{b\bar{c}\}$	0
$\{\bar{b}c\}$	0.3
$\{\bar{b}\bar{c}\}$	0.1
$\{bc, b\bar{c}\}$	0
$\{bc, \bar{b}c\}$	0
$\{bc, \bar{b}\bar{c}\}$	0.1
$\{b\bar{c}, \bar{b}c\}$	0
$\{b\bar{c}, \bar{b}\bar{c}\}$	0
$\{\bar{b}c, \bar{b}\bar{c}\}$	0.1
$\{bc, b\bar{c}, \bar{b}c\}$	0
$\{bc, b\bar{c}, \bar{b}\bar{c}\}$	0
$\{bc, \bar{b}c, \bar{b}\bar{c}\}$	0.3
$\{b\bar{c}, \bar{b}c, \bar{b}\bar{c}\}$	0
$\{bc, b\bar{c}, \bar{b}c, \bar{b}\bar{c}\}$	0.1

Example 2: Consider three binary variables X_1, X_2, X_3 with $\mathbf{X}_1 = \{a, \bar{a}\}$, $\mathbf{X}_2 = \{b, \bar{b}\}$, $\mathbf{X}_3 = \{c, \bar{c}\}$, and two 2-dimensional basic assignments m_1 and m_2 as specified in Table III.

Notice that these two assignments are not projective; for this see their one-dimensional marginals in Table IV. Therefore, because of property (iii) of Lemma 5 presented below $m_1 \triangleright m_2 \neq m_2 \triangleright m_1$.

To determine general 3-dimensional assignment (of binary variables) one has to specify 255 numbers, because $\mathbf{X}_{\{1,2,3\}}$ has $2^8 - 1 = 255$ nonempty subsets. However, when computing $m_1 \triangleright m_2$, most of these 255 values equal 0 because most of these subsets do not meet the condition $C = C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$ and therefore the corresponding value of the assignment $m_1 \triangleright m_2$ is defined by the point [c] of the definition.

What are the subsets for which $C \neq C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$?

TABLE IV

ONE-DIMENSIONAL MARGINAL ASSIGNMENTS $m_1^{\downarrow\{1\}}, m_1^{\downarrow\{2\}}$ AND $m_2^{\downarrow\{2\}}, m_2^{\downarrow\{3\}}$.

$A \subseteq \mathbf{X}_1$	$m_1^{\downarrow\{1\}}(A)$
$\{a\}$	0.6
$\{\bar{a}\}$	0.2
$\{a, \bar{a}\}$	0.2

$A \subseteq \mathbf{X}_2$	$m_1^{\downarrow\{2\}}(A)$
$\{b\}$	0.3
$\{\bar{b}\}$	0.5
$\{b, \bar{b}\}$	0.2

$A \subseteq \mathbf{X}_2$	$m_2^{\downarrow\{2\}}(A)$
$\{b\}$	0
$\{\bar{b}\}$	0.5
$\{b, \bar{b}\}$	0.5

$A \subseteq \mathbf{X}_3$	$m_2^{\downarrow\{3\}}(A)$
$\{c\}$	0.3
$\{\bar{c}\}$	0.3
$\{c, \bar{c}\}$	0.4

TABLE V

BASIC ASSIGNMENT $m_1 \triangleright m_2$ FOR EXAMPLE 2.

	$C \subseteq \mathbf{X}_{\{1,2,3\}}$	$C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$	$(m_1 \triangleright m_2)(C)$
[a]	$\{a\bar{b}c\}$	$\{a\bar{b}\} \otimes \{\bar{b}c\}$	0.3
[a]	$\{a\bar{b}\bar{c}\}$	$\{a\bar{b}\} \otimes \{\bar{b}\bar{c}\}$	0.1
[a]	$\{a\bar{b}c, a\bar{b}\bar{c}\}$	$\{a\bar{b}\} \otimes \{\bar{b}c, \bar{b}\bar{c}\}$	0.1
[b]	$\{abc, ab\bar{c}\}$	$\{ab\} \otimes \mathbf{X}_1$	0.1
[b]	$\{\bar{a}bc, \bar{a}b\bar{c}\}$	$\{\bar{a}b\} \otimes \mathbf{X}_1$	0.2
[a]	$\{abc, \bar{a}bc, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c}\}$	$\mathbf{X}_{\{1,2\}} \otimes \{bc, \bar{b}\bar{c}\}$	0.04
[a]	$\left\{\begin{matrix} abc, \bar{a}bc, a\bar{b}c \\ \bar{a}\bar{b}c, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c} \end{matrix}\right\}$	$\mathbf{X}_{\{1,2\}} \otimes \{bc, \bar{b}c, \bar{b}\bar{c}\}$	0.12
[a]	$\left\{\begin{matrix} abc, \bar{a}bc, a\bar{b}c, \bar{a}\bar{b}c \\ ab\bar{c}, \bar{a}b\bar{c}, a\bar{b}\bar{c}, \bar{a}\bar{b}\bar{c} \end{matrix}\right\}$	$\mathbf{X}_{\{1,2\}} \otimes \mathbf{X}_{\{2,3\}}$	0.04

For example, it is easy to show that all the sets of cardinality 7 belong to this category (hint: show that for any $C \subseteq \mathbf{X}_{\{1,2,3\}}$, for which $|C| = 7$, $C^{\downarrow\{1,2\}} = \mathbf{X}_{\{1,2\}}$ and $C^{\downarrow\{2,3\}} = \mathbf{X}_{\{2,3\}}$).

Since all singletons (one-point-sets) meet the considered equality $C = C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$, all sets C, for which $C \neq C^{\downarrow\{1,2\}} \otimes C^{\downarrow\{2,3\}}$ must have at least two elements: an example is $\{abc, \bar{a}b\bar{c}\}$. As further examples may serve sets $\{ab\bar{c}, \bar{a}bc, \bar{a}b\bar{c}, a\bar{b}c\}$ and $\{\bar{a}bc, a\bar{b}c, ab\bar{c}\}$. A common characteristics of all these sets is that assigning a positive belief to them one introduces a type of conditional relationship between X_1 and X_3 given (at least one) value of X_2.

Let us turn our attention back to computation of $m_1 \triangleright m_2$ for assignments of our example. For this, one immediately notices that point [b] of the definition is used whenever $C \subseteq \mathbf{X}_{\{1,2,3\}}$ is considered for which $C^{\downarrow\{2\}} = b$, since $m_2^{\downarrow\{2\}}(b) = 0$. In fact, we get only 8 subsets, for which assignment $m_1 \triangleright m_2$ is positive - see Table V, where the first column bears the information, which point of the definition is used to compute the respective value. ■

Let us start studying properties of the operator of composition for basic assignments. The reader will perhaps not be surprised if we claim that the operator of composition meets all the six basic properties holding for operators of composition in probabilistic as well as in possibilistic cases. Notice that similarly to the possibilistic version (and in contrast to the probabilistic one), the operator of composition for basic assignments is always defined.

Lemma 5: For arbitrary basic assignments m_1, m_2, m_3 defined on $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$, respectively

(i) $m_1 \triangleright m_2$ *is a basic assignment on* $\mathbf{X}_{K_1 \cup K_2}$*;*

(ii) $(m_1 \triangleright m_2)^{\downarrow K_1} = m_1$*;*

(iii) $m_1 \triangleright m_2 = m_2 \triangleright m_1 \iff m_1^{\downarrow K_1 \cap K_2} = m_2^{\downarrow K_1 \cap K_2}$*;*

(iv) $K_1 \supseteq (K_2 \cap K_3)$
$$\implies (m_1 \triangleright m_2) \triangleright m_3 = (m_1 \triangleright m_3) \triangleright m_2;$$

(v) $K_2 \supseteq L \supseteq (K_1 \cap K_2)$
$$\implies m_1 \triangleright m_2 = (m_1 \triangleright m_2^{\downarrow L}) \triangleright m_2;$$

(vi) $(K_1 \cup K_2) \supseteq L \supseteq K_1$
$$\implies (m_1 \triangleright m_2)^{\downarrow L} = m_1 \triangleright m_2^{\downarrow K_2 \cap L}.$$

All these properties were proved in [8]. The only exception is property (v); however its proof is rather technical and we omit it due to the lack of space.

Answering the question, what is the relationship between the factorization with the help of the operator of composition and the concept of conditional independence, is in this case much more difficult than in the previous two subsections. One of the reasons is the fact that the notion of the conditional independence for belief functions was introduced in several different ways. Perhaps the most frequent (and maybe also with the greatest number of supporters) is the one, which can be easily defined with the help of *commonality function*. Using notation of Studený [11], commonality function Com_m is defined for basic assignment m (assuming that m is defined on $\mathbf{X}_N$) for each $A \subset \mathbf{X}_N$ by a simple formula

$$Com_m(A) = \sum_{B \supseteq A} m(B).$$

Yaghlane *et al.* [2] define the concept of conditional noninteractivity (as well as Shenoy defines his concept of conditional independence) in the way that variables X_I and variables X_J are *conditionally non-interactive* given variables X_K if and only if for all $A \subseteq \mathbf{X}_N$

$$Com_{m^{\downarrow I \cup J \cup K}}(A^{\downarrow I \cup J \cup K}) \cdot Com_{m^{\downarrow K}}(A^{\downarrow K}) = Com_{m^{\downarrow I \cup K}}(A^{\downarrow I \cup K}) \cdot Com_{m^{\downarrow J \cup K}}(A^{\downarrow J \cup K}).$$

In this paper we shall denote this property by

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K.$$

Though for basic assignments it does not hold true that $X_I \perp\!\!\!\perp_{[m]} X_J | X_K$ if and only if the basic marginal assignment $m^{\downarrow I \cup J \cup K}$ factorizes in the following sense

$$m^{\downarrow I \cup J \cup K} = m^{\downarrow I \cup K} \triangleright m^{\downarrow J \cup K},$$

still there are properties which reflect a similarity of these two notions. First, Yaghlane *et al.* in [2] showed that if $X_I \perp\!\!\!\perp_{[m]} X_J | X_K$ then all focal elements of $m^{\downarrow I \cup J \cup K}$ (i.e. sets $A \subseteq \mathbf{X}_{I \cup J \cup K}$, for which $m^{\downarrow I \cup J \cup K}(A) > 0$) are *Z-layered rectangles*, which are nothing else than sets $A \subseteq \mathbf{X}_{I \cup J \cup K}$, which can be expressed as an extension of its respective projections:

$$A = A^{\downarrow I \cup K} \otimes A^{\downarrow J \cup K}.$$

Therefore, combining the mentioned Yaghlane *et al.* property with Definition 4 we get the following simple assertion.

Assertion: Consider a basic assignment m on $\mathbf{X}_N$ and three disjoint subsets $I, J, K \subset N$ ($I \neq \emptyset \neq J$). If $A \subseteq \mathbf{X}_{I\cup J\cup K}$ is a focal element of $m^{\downarrow I\cup J\cup K}$ and $A \neq A^{\downarrow I\cup K} \otimes A^{\downarrow J\cup K}$ then neither of the following two expressions holds true:

$$X_I \perp\!\!\!\perp_{[m]} X_J | X_K,$$

and

$$m^{\downarrow I\cup J\cup K} = m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}.$$

So, the first property connecting the concepts of conditional non-interactivity and factorization is that any of them guarantees that the focal elements of the respective basic assignment can be expressed as an extension of its corresponding projections (Z-layered rectangles in the language of Yaghlane *et al.*).

It is well known that if all focal elements of a basic assignment m are *singletons*, i.e. if $m(A) > 0$ implies that $|A| = 1$, then this basic assignment corresponds to a probability distribution, and it is why some authors call it *Bayesian basic assignment*. In [11] Studený claims that for Bayesian basic assignments the concept of conditional non-interactivity coincides with the concept of conditional independence of the corresponding probability distribution. In [8], we proved that if we compose by the operator of composition two Bayesian basic assignments, such that the corresponding probability distributions may be composed by the probabilistic operator of composition (i.e. the composition is defined) then the resulting distribution is again Bayesian.

Lemma 6: Let m_1 and m_2 be Bayesian basic assignments on $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively, for which

$$m_2{}^{\downarrow K\cap L}(A) = 0 \implies m_1{}^{\downarrow K\cap L}(A) = 0 \qquad (1)$$

for any $A \subseteq \mathbf{X}_{K\cup L}$. Then $m_1 \triangleright m_2$ is a Bayesian basic assignment.

From this, it is obvious that considering Bayesian assignment m, for composition of its marginals $m^{\downarrow I\cup K} \triangleright m^{\downarrow J\cup K}$ only case [a] of the definition is applied.[1] Therefore, comparing Definitions 1 and 4 we see that the result is always defined (composed distributions, being marginals of the same distribution are consistent) and is the same for both the Definitions.

These considerations result in a second property connecting the concepts of conditional non-interactivity and factorization: for Bayesian basic assignments they coincide with probabilistic conditional independence.

In conclusion of this section we will show the difference between the compared concepts. Namely, in [2] the authors admit that their concept of conditional non-interactivity (as showed by Studený) is *not consistent with marginalization*. What does it mean and the fact that the factorization considered in this paper does not suffer from this imperfectness will be visible from the following example, which is borrowed from [2].

Example 3: Consider three binary variables X_1, X_2, X_3 with $\mathbf{X}_1 = \{a, \bar{a}\}$, $\mathbf{X}_2 = \{b, \bar{b}\}$, $\mathbf{X}_3 = \{c, \bar{c}\}$ as in Example 2, and two basic assignments $m_1(x_{\{1,3\}}), m_2(x_{\{2,3\}})$, each of which having only two focal elements:

$$m_1(\{a\bar{c}, \bar{a}\bar{c}\}) = 0.5 \quad m_1(\{a\bar{c}, \bar{a}c\}) = 0.5$$
$$m_1(\{b\bar{c}, \bar{b}\bar{c}\}) = 0.5 \quad m_1(\{b\bar{c}, \bar{b}c\}) = 0.5$$

For them, it is showed in [2] that there does not exist basic assignment m on $\mathbf{X}_{\{1,2,3\}}$ such that m_1, m_2 are its marginals (i.e. $m^{\downarrow\{1,3\}} = m_1, m^{\downarrow\{2,3\}} = m_2$) and $X_1 \perp\!\!\!\perp_{[m]} X_2 | X_3$.

Since

$$m_1^{\downarrow\{3\}}(\{\bar{c}\}) = m_2^{\downarrow\{3\}}(\{\bar{c}\}) = .5,$$

$$m_1^{\downarrow\{3\}}(\{c, \bar{c}\}) = m_2^{\downarrow\{3\}}(\{c, \bar{c}\}) = .5,$$

basic assignments m_1 and m_2 are projective (consistent), and therefore their composition (and due to property (iii) it does not matter whether we consider $m = m_1 \triangleright m_2$ or $m = m_2 \triangleright m_1$) is an assignment having both m_1 and m_2 for its marginals. As the reader easily verifies, it is an assignment with also only two focal elements

$$(m_1 \triangleright m_2)(\{ab\bar{c}, a\bar{b}\bar{c}, \bar{a}b\bar{c}, \bar{a}\bar{b}\bar{c}\}) = .5,$$
$$(m_1 \triangleright m_2)(\{ab\bar{c}, \bar{a}\bar{b}c\}) = .5.$$

Now, let us show that $X_1 \not\perp\!\!\!\perp_{[m_1 \triangleright m_2]} X_2 | X_3$. For this, it is enough to show for one set $A \subseteq \{a, \bar{a}\} \times \{b, \bar{b}\} \times \{c, \bar{c}\}$ the following equality does not hold true

$$Com_m(A) \cdot Com_{m^{\downarrow\{3\}}}(A^{\downarrow\{3\}})$$
$$= Com_{m^{\downarrow\{1,3\}}}(A^{\downarrow\{1,3\}}) \cdot Com_{m^{\downarrow\{2,3\}}}(A^{\downarrow\{2,3\}}).$$

Let us consider $A = \{ab\bar{c}, a\bar{b}\bar{c}, \bar{a}b\bar{c}, \bar{a}\bar{b}\bar{c}\}$ and compute the necessary commonality functions

$$Com_m(\{ab\bar{c}, a\bar{b}\bar{c}, \bar{a}b\bar{c}, \bar{a}\bar{b}\bar{c}\}) = m(\{ab\bar{c}, a\bar{b}\bar{c}, \bar{a}b\bar{c}, \bar{a}\bar{b}\bar{c}\}) = 0.5,$$
$$Com_{m^{\downarrow\{1,3\}}}(\{a\bar{c}, \bar{a}\bar{c}\}) = m^{\downarrow\{1,3\}}(\{a\bar{c}, \bar{a}\bar{c}\}) = 0.5,$$
$$Com_{m^{\downarrow\{2,3\}}}(\{b\bar{c}, \bar{b}\bar{c}\}) = m^{\downarrow\{2,3\}}(\{b\bar{c}, \bar{b}\bar{c}\}) = 0.5$$

and, eventually

$$Com_{m^{\downarrow\{3\}}}(\{\bar{c}\}) = m^{\downarrow\{3\}}(\{\bar{c}\}) + m^{\downarrow\{3\}}(\{c, \bar{c}\}) = 1.0.$$

So we see that though the basic assignment m factorizes in the sense that $m = m^{\downarrow\{1,3\}} \triangleright m^{\downarrow\{2,3\}}$, $X_1 \not\perp\!\!\!\perp_{[m]} X_2 | X_3$. ■

III. General Properties of the Operator

This section will be devoted to properties of the operator of composition which do not depend on the framework in which the operator is defined. The studied characteristics can be deduced with the help of properties (i)-(vi) (see Lemmas 1, 3 and 5), which hold for all three versions of

[1] In fact, case [b] is used when $m^{\downarrow K\cap L}(A^{\downarrow K\cap L}) = 0$ but then the results equals zero, which is the same value, which would be received by application of rule [a].

the operator. Therefore, in this section we will consider an *object* $\wp$, which may be either probability distributions, or possibility distributions or basic assignments. Recalling that properties (i)-(vi) hold for probability distributions only if $p_1 \triangleright p_2$ (or $(p_1 \triangleright p_2) \triangleright p_3$ in case of point (iv)) is defined, we have to realize that if we compose marginals $p^{\downarrow I}, p^{\downarrow J}$ of a probability distribution p then the composition $p^{\downarrow I} \triangleright p^{\downarrow J}$ is always defined. Moreover, we also must not forget that in case when $\wp$ is a possibility distribution then the operator of composition $\triangleright$ is parameterized by a continuous t-norm.

Let us consider an arbitrary object $\wp$ defined on $\mathbf{X}_N$ and show that the ternary relation of factorization defined (for disjoint $I, J, L \subset N$, $I \neq \emptyset \neq J$) in the following way

$$X_I \perp\!\!\!\perp_\wp X_J | X_L \iff \wp^{\downarrow I\cup J\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup L}$$

is a *semigraphoid*, i.e. it meets the four semigraphoid axioms listed below. In what follows, each axiom is reformulated into the language of composition and the corresponding theorem is proved.

A. Symmetry

$$X_I \perp\!\!\!\perp_\wp X_J | X_L \Longrightarrow X_J \perp\!\!\!\perp_\wp X_I | X_L$$

Theorem 1: If $\wp^{\downarrow I\cup J\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup L}$ *then also* $\wp^{\downarrow I\cup J\cup L} = \wp^{\downarrow J\cup L} \triangleright \wp^{\downarrow I\cup L}$.

Proof: The assertion follows immediately from the fact that marginals $\wp^{\downarrow I\cup L}$ and $\wp^{\downarrow J\cup L}$ are consistent, and therefore property (iii) may be applied

$$\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup L} = \wp^{\downarrow J\cup L} \triangleright \wp^{\downarrow I\cup L}.$$

■

B. Decomposition

$$X_I \perp\!\!\!\perp_\wp X_{J\cup K} | X_L \Longrightarrow X_I \perp\!\!\!\perp_\wp X_K | X_L$$

Theorem 2: If $\wp^{\downarrow I\cup J\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}$ *then also* $\wp^{\downarrow I\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}$.

Proof: The assertion will be got just by application of marginalization property (vi)

$$\begin{aligned} \wp^{\downarrow I\cup K\cup L} &= \left(\wp^{\downarrow I\cup J\cup K\cup L}\right)^{\downarrow I\cup K\cup L} \\ &= \left(\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}\right)^{\downarrow I\cup K\cup L} \\ &= \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}. \end{aligned}$$

■

C. Weak Union

$$X_I \perp\!\!\!\perp_\wp X_{J\cup K} | X_L \Longrightarrow X_I \perp\!\!\!\perp_\wp X_J | X_{K\cup L}$$

Theorem 3: If $\wp^{\downarrow I\cup J\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}$ *then also* $\wp^{\downarrow I\cup J\cup K\cup L} = \wp^{\downarrow I\cup K\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}$.

Proof: To prove this assertion we have to realize that, due to property (v),

$$\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L} = \left(\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}\right) \triangleright \wp^{\downarrow J\cup K\cup L},$$

and that, because the assumptions of Theorem 2 are fulfilled, also

$$\wp^{\downarrow I\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}.$$

Using these two equalities we finish the proof in a simple way

$$\begin{aligned} \wp^{\downarrow I\cup J\cup K\cup L} &= \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L} \\ &= \left(\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}\right) \triangleright \wp^{\downarrow J\cup K\cup L} \\ &= \wp^{\downarrow I\cup K\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}. \end{aligned}$$

■

D. Contraction

$$X_I \perp\!\!\!\perp_\wp X_K | X_L \;\; \& \;\; X_I \perp\!\!\!\perp_\wp X_J | X_{K\cup L} \Longrightarrow X_I \perp\!\!\!\perp_\wp X_{J\cup K} | X_L$$

Theorem 4: If $\wp^{\downarrow I\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}$, *and* $\wp^{\downarrow I\cup J\cup K\cup L} = \wp^{\downarrow I\cup K\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}$, *then also* $\wp^{\downarrow I\cup J\cup K\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L}$.

Proof: We will follow the same idea as in the preceding proof but in the reverse order. First, we will use property (v) and then both assumptions of this assertion.

$$\begin{aligned} \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup K\cup L} &= \left(\wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow K\cup L}\right) \triangleright \wp^{\downarrow J\cup K\cup L} \\ &= \wp^{\downarrow I\cup K\cup L} \triangleright \wp^{\downarrow J\cup K\cup L} \\ &= \wp^{\downarrow I\cup J\cup K\cup L}. \end{aligned}$$

■

IV. Conclusions

In three different frameworks we introduced an operator of composition, which, when applied to low-dimensional objects, forms a more-dimensional objects of the same type. Therefore, when applied iteratively, the operator of composition enables constructing multidimensional models from a system of low-dimensional objects. In this paper we were interested only in the very basic properties of the operator: especially, in possibility to introduce the ternary relation of factorization:

$$X_I \perp\!\!\!\perp_\wp X_J | X_L \iff \wp^{\downarrow I\cup J\cup L} = \wp^{\downarrow I\cup L} \triangleright \wp^{\downarrow J\cup L}.$$

We showed a close connection between the relations of factorization and conditional independence in probability and possibility theories, and compared these two notions for belief functions. It appeared to be interesting that for proving famous semigraphoid axioms for the concept of factorization we needed only six basic properties (i)-(vi) (though we are not sure, we expect them to be independent).

When comparing factorization with the prevalent concept of conditional independence for belief functions (Yaghlane *et al.* call it conditional non-interactivity) we showed that, though manifesting some equal properties, they differ from each other. We showed that our concept of factorization does not suffer from the insufficiency, which Studený calls *inconsistency with marginalization*. Nevertheless, we do not hide the fact that in contrast with the fact that belief function factorization coincides with probabilistic conditional independence for Bayesian basic assignments, no such a coincidence holds for basic assignments representing possibilistic distributions. This is, however, quite a different (and complex) question (we know that it is under a serious reserch of J. Vejnarová [18]).

REFERENCES

[1] B. Ben Yaghlane, Ph. Smets, and K. Mellouli, "Belief Function Independence: I. The Marginal Case," *Int. J. of Approximate Reasoning*, vol. 29, no. 1, pp. 47–70, 2002.

[2] B. Ben Yaghlane, Ph. Smets, and K. Mellouli, "Belief Function Independence: II. The Conditional Case," *Int. J. of Approximate Reasoning*, vol. 31, no. (1-2), pp. 31–75, 2002.

[3] E. Ben Yaghlane, Ph. Smets, and K. Mellouli, "Directed Evidential Networks with Conditional Belief functions," *Proc. of ECSQARU 2003*, LNAI 2711, Springer, pp. 291–305, 2003.

[4] R. Jiroušek, "Composition of probability measures on finite spaces," *Proc. of the 13th Conf. Uncertainty in Artificial Intelligence UAI'97*, (D. Geiger and P. P. Shenoy, eds.). Morgan Kaufmann Publ., San Francisco, California, pp. 274–281, 1997.

[5] R. Jiroušek "Graph modelling without graphs," *Proc. of the 7th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems IPMU'98*, (B. Bouchon-Meunier, R.R. Yager, eds.). Editions E.D.K. Paris, pp. 809–816, 1988.

[6] R. Jiroušek "Marginalization in composed probabilistic models," *Proc. of the 16th Conf. Uncertainty in Artificial Intelligence UAI'00* (C. Boutilier and M. Goldszmidt, eds.), Morgan Kaufmann Publ., San Francisco, California, pp. 301–308, 2000.

[7] R. Jiroušek, J. Vejnarová, "General framework for multidimensional models," *Int. J. of Intelligent Systems*, vol. 18, pp. 107–127, 2003.

[8] R. Jiroušek, J. Vejnarová and M. Daniel, "Compositional models of belief functions," *Proc. of the 5th Symposium on Imprecise Probabilities and Their Applications* (G. de Cooman, J. Vejnarová and M. Zaffalon, eds.), Charles University Press, Praha, pp. 243–252, 2007.

[9] P. P. Shenoy, "Conditional independence in valuation-based systems," *Int. J. of Approximate Reasoning*, vol. 10, no. 3, pp. 203–234, 1994.

[10] P. P. Shenoy, "Binary join trees for computing marginals in the Shenoy-Shafer architecture," *Int. J. of Approximate Reasoning*, vol. 17, no. (2-3), pp. 239–263, 1997.

[11] M. Studený, "Formal properties of conditional independence in different calculi of AI," *Proceedings of European Conference on Symbolic and quantitative Approaches to Reasoning and Uncertainty ECSQARU'93*, (K. Clarke, R. Kruse and S. Moral, eds.). Springer-Verlag, 1993, pp. 341–351.

[12] M. Studený, "On stochastic conditional independence: the problems of characterization and description," *Annals of Mathematics and Artificial Intelligence,* vol. 35, p. 323-341, 2002.

[13] M. Studený, "Other approaches to the description of conditional independence structures," *Highly Structured Stochastic Systems,* (Green P. J., Hjort N. L., Richardson S., eds.). Oxford University Press, New York, pp 106-108, 2003.

[14] M. Studený. *Probabilistic Conditional Independence Structures.* Springer, London, 2005.

[15] J. Vejnarová, "Composition of possibility measures on finite spaces: preliminary results," *Proceedings of 7th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems IPMU'98,* (B. Bouchon-Meunier, R.R. Yager, eds.). Editions E.D.K. Paris, 1998, pp. 25–30.

[16] J. Vejnarová, "Possibilistic independence and operators of composition of possibility measures," *Prague Stochastics'98,* (M. Hušková , J. Á. Víšek, P. Lachout, eds.) JČMF, 1998, pp. 575–580.

[17] J. Vejnarová, "Operators of composition and their iterations," LISp Report 9803, VŠE Praha, 1998.

[18] J. Vejnarová, Personal communication, 2007.