International Journal of Approximate Reasoning 90 (2017) 359–373

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Compositional models for credal sets ☆

Jiřina Vejnarová

Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodárenskou věží 4, Prague, Czech Republic

ARTICLE INFO

Article history:
Received 31 December 2016
Received in revised form 15 July 2017
Accepted 14 August 2017
Available online 23 August 2017

Keywords:
Imprecise probabilities
Credal sets
Multidimensional models
Conditional independence

ABSTRACT

We present the composition operator, already known from probability, possibility, evidence and valuation-based systems theories, for credal sets. We prove that the proposed definition preserves all the properties enabling us to design compositional models in a way analogous to those in the above-mentioned theories. A special kind of compositional models, the so-called perfect sequences of credal sets, is studied in more detail and (among others) its relationship to perfect sequences of probability distributions is revealed. The theoretical results are illustrated by numerous simple examples.

1. Introduction

In the late 1990s, a new approach to efficient representation of multidimensional probability distributions was introduced with the aim to be an alternative to Graphical Markov Modelling. This approach is based on a simple idea: a multidimensional distribution is *composed* from a system of low-dimensional distributions by repetitive application of a special composition operator, which is also the reason why such models are called *compositional models*.

Later, these compositional models were also introduced in possibility theory [17,18] (here the models are parameterised by a continuous t-norm), and almost ten years ago in evidence theory as well [8,9]. In all these frameworks, the original idea is kept, but there exist some slight differences among these frameworks.

In this paper we present a composition operator for credal sets that fixes an issue known about the previously proposed operator [20]. The primary goal of this paper is to show that the revised composition operator keeps the basic properties of its counterparts in other frameworks, and therefore it enables us to introduce compositional models for multidimensional credal sets, as (to a certain extent) already suggested in [21].

This paper is organised as follows. In Section 2 we summarise the basic concepts and notation. The definition of the operator of composition is presented in Section 3, which is also devoted to its basic properties and a few illustrative examples. In Section 4 we focus on perfect sequences of credal sets and their properties.

2. Basic concepts and notation

In this Section we will briefly recall basic concepts and notation necessary for understanding the paper.

☆ This paper is part of the Virtual special issue on Soft methods in probability and statistics, Edited by Barbara Vantaggi, Maria Brigida Ferraro, Paolo Giordani.

E-mail address: vejnar@utia.cas.cz.

http://dx.doi.org/10.1016/j.ijar.2017.08.007
0888-613X/

Table 1
Credal set describing G and W in Example 1.

$\mathcal{M}(GW)$	P^1		P^2		P^3		P^4	
W	0	1	0	1	0	1	0	1
$G=0$	0.18	0.27	0.12	0.18	0.09	0.36	0.06	0.24
$G=1$	0.44	0.11	0.56	0.14	0.33	0.22	0.42	0.28

Table 2
Marginal credal set describing G in Example 1.

$\mathcal{M}(G)$	P^1	P^2	P^3	P^4
$G=0$	0.45	0.3	0.45	0.3
$G=1$	0.55	0.7	0.55	0.7

2.1. Variables and distributions

For an index set $N = \{1, 2, \dots, n\}$, let $\{X_i\}_{i \in N}$ be a system of variables, each X_i having its values in a finite set $\mathbf{X}_i$; by $\mathbf{X}_N = \mathbf{X}_1 \times \mathbf{X}_2 \times \ldots \times \mathbf{X}_n$ we denote the Cartesian product of these sets.

In this paper we will deal with groups of variables on the Cartesian product's subspaces. Let us note that, throughout the paper, X_K will denote a group of variables $\{X_i\}_{i \in K}$ with values in $\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i$.

Any group of variables X_K can be described by a *probability distribution* (sometimes also called *probability (mass) function*)

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

such that

$$\sum_{x_K \in \mathbf{X}_K} P(x_K) = 1.$$

Having two probability distributions P_1 and P_2 of X_K, we say that P_1 is *absolutely continuous* with respect to P_2 (and denote $P_1 \ll P_2$) if, for any $x_K \in \mathbf{X}_K$,

$$P_2(x_K) = 0 \Longrightarrow P_1(x_K) = 0.$$

This concept plays an important role in the definition of the composition operator.

2.2. Credal sets

A *credal set* $\mathcal{M}(X_K)$ describing a group of variables X_K is usually defined as a closed convex set of probability measures describing the values of these variables. In order to simplify the expression of operations with credal sets, it is often considered ([14]) that a credal set is the set of probability distributions associated to the probability measures in it. The reason is quite simple – to work with point functions is usually more convenient than to work with set functions.

Under such considerations, a credal set can be expressed as a *convex hull* (denoted by CH) of its extreme distributions (ext)

$$\mathcal{M}(X_K) = \mathrm{CH}\{ext(\mathcal{M}(X_K))\}.$$

In this paper we will consider only credal sets with a finite number of extreme points.

Consider a credal set $\mathcal{M}(X_K)$. For each $L \subset K$, its *marginal credal set* $\mathcal{M}(X_L)$ is obtained by element-wise marginalisation, i.e.,

$$\mathcal{M}(X_L) = \mathrm{CH}\{P^{\downarrow L} : P \in \mathrm{ext}(\mathcal{M}(X_K))\}, \tag{1}$$

where $P^{\downarrow L}$ denotes the marginal distribution of P on $\mathbf{X}_L$.

Besides marginalisation, we will also need the opposite operation, called vacuous extension. *Vacuous extension* of a credal set $\mathcal{M}(X_L)$ describing X_L to a credal set $\mathcal{M}(X_K) = \mathcal{M}(X_L)^{\uparrow K}$ $(L \subset K)$ is the maximal credal set describing X_K such that $\mathcal{M}(X_K)^{\downarrow L} = \mathcal{M}(X_L)$.

Example 1. Let $\mathcal{M}(G, W)$ be a credal set describing binary variables G and W as suggested by Table 1 (their interpretation can be found in Section 4.5). The marginal credal set describing G is then contained in Table 2. Let us note that the first and third marginals coincide (as do the remaining two). One can see that its vacuous extension $\mathcal{M}(G)^{\uparrow GW}$ (described by Table 3) is much bigger than $\mathcal{M}(G, W)$. More precisely, any extreme point of $\mathcal{M}(G, W)$ lies inside $\mathcal{M}(G)^{\uparrow GW}$.

To show that $\mathcal{M}(G)^{\uparrow GW}$ is maximal, let us suppose that there exists a credal set $\mathcal{M}'(GW)$ containing $\mathcal{M}(G)^{\uparrow GW}$ and $\mathcal{M}(G) = \mathcal{M}'(G)$. Then $\mathcal{M}'(GW)$ must contain at least one $p = (p_1, p_2, p_3, p_4) \notin \mathcal{M}(X_1 X_2)$. Nevertheless, it means that either $p_1 + p_2 < 0.3$ or $p_1 + p_2 > 0.45$ (from which analogous inequalities for $p_3 + p_4$ follow). However, $p^{\downarrow\{1\}} \notin \mathcal{M}(G)$ and therefore $\mathcal{M}(G)^{\uparrow GW}$ is maximal. □

Table 3
Vacuous extension of $\mathcal{M}(G)$ – $\mathcal{M}(G)^{\uparrow GW}$ in Example 1.

$\mathcal{M}(GW)$	P^1		P^2		P^3		P^4	
W	0	1	0	1	0	1	0	1
$G=0$	0.45	0	0	0.45	0.45	0	0	0.45
$G=1$	0.55	0	0.55	0	0	0.55	0	0.55
	P^5		P^6		P^7		P^8	
W	0	1	0	1	0	1	0	1
$G=0$	0.3	0	0	0.3	0.3	0	0	0.3
$G=1$	0.7	0	0.7	0	0	0.7	0	0.7

Having two credal sets $\mathcal{M}_1$ and $\mathcal{M}_2$ describing X_K and X_L, respectively (assuming that $K, L \subseteq N$), we say that these credal sets are *projective* if their marginals describing common variables coincide, i.e., if

$$\mathcal{M}_1(X_{K\cap L}) = \mathcal{M}_2(X_{K\cap L}).$$

Let us note that if K and L are disjoint, then $\mathcal{M}_1$ and $\mathcal{M}_2$ are always projective, as $\mathcal{M}_1(X_\emptyset) = \mathcal{M}_2(X_\emptyset) \equiv 1$.

2.3. Independence

Among the numerous definitions of independence for credal sets studied in [3], two of them appear as the most appropriate to be applied in real situations: epistemic independence and strong independence.

Epistemic independence can be intuitively expressed as follows: two variables are *epistemically independent* if information about one of them cannot change our state of knowledge about the other. It is a very natural definition, often used in a precise probability setting to explain the concept of independence; moreover, it possesses a clear behavioural interpretation. However, as noticed in [14], it gives rise to very complex problems in inference.

Strong independence, on the other hand, is the most usual concept in credal networks, and therefore it seems to be the most appropriate for our multidimensional models.

We say that (groups of) variables X_K and X_L (K and L disjoint) are *strongly independent* with respect to $\mathcal{M}(X_{K\cup L})$ iff (in terms of probability distributions)

$$\mathcal{M}(X_{K\cup L}) = \mathrm{CH}\{P_1 \cdot P_2 : P_1 \in \mathcal{M}(X_K), P_2 \in \mathcal{M}(X_L)\}.$$

2.4. Conditional independence

Again, there exist several generalisations of this notion to conditional independence, as one can see in [14]. The following definition (in [14], called *conditional independence in distribution*), is suggested by the authors as the most appropriate for the marginal problem; hence it seems to be a suitable concept in our case as well, since the operator of composition can also be used as a tool for solving the marginal problem, as shown (in the framework of possibility theory) e.g., in [18].

Given three groups of variables X_K, X_L and X_M (K, L, M being mutually disjoint subsets of N such that K and L are nonempty), we say that X_K and X_L are *conditionally independent* given X_M under global set $\mathcal{M}(X_{K\cup L\cup M})$ (to simplify the notation, we will denote this relationship by $K \perp\!\!\!\perp L|M$) iff

$$\mathcal{M}(X_{K\cup L\cup M}) = \mathrm{CH}\{(P_1 \cdot P_2)/P_1^{\downarrow M} : P_1 \in \mathcal{M}(X_{K\cup M}), P_2 \in \mathcal{M}(X_{L\cup M}), P_1^{\downarrow M} = P_2^{\downarrow M}\}.$$

This definition is a generalisation of stochastic conditional independence: if $\mathcal{M}(X_{K\cup L\cup M})$ is a singleton, then $\mathcal{M}(X_{K\cup M})$ and $\mathcal{M}(X_{L\cup M})$ are also (projective) singletons and the definition reduces to the definition of stochastic conditional independence.

It has been proven in [14] that this independence concept satisfies the so-called semigraphoid properties [11], analogous to stochastic conditional independence, and it is the reason it is also suitable for our models.

The other two concepts, called *conditional independence on decomposition* and *causal irrelevance*, respectively, are stronger than the definition introduced above. Adopting any of these concepts would substantially influence the properties of compositional models (or, at least, their relation to conditional independence). Furthermore, none of these concepts satisfies the semigraphoid properties. The former fails contraction, while the latter is not symmetric (for more details see [14]).

3. Composition operator

In this section we will introduce the definition of composition operator for credal sets. The concept of the composition operator is presented first in a precise probability framework, as it seems to be useful for its better understanding.

Table 4
Two marginal credal sets describing variables R and T – Example 2.

$\mathcal{M}_1(R)$	P_1^1	P_1^2
$R=0$	0.2	0.5
$R=1$	0.8	0.5

$\mathcal{M}_2(T)$	P_2^1	P_2^2
$T=0$	0.5	0.8
$T=1$	0.5	0.2

3.1. Composition operator of probability distributions

Let us recall the definition of composition of two probability distributions [5]. Consider two index sets $K, L \subset N$. We do not put any restrictions on K and L; they may, but need not, be disjoint, and one may be a subset of the other. Let P_1 and P_2 be two probability distributions of (groups of) variables X_K and X_L; then

$$(P_1 \triangleright P_2)(X_{K\cup L}) = \frac{P_1(X_K)\cdot P_2(X_L)}{P_2(X_{K\cap L})}, \tag{2}$$

whenever $P_1(X_{K\cap L}) \ll P_2(X_{K\cap L})$; otherwise, it remains undefined.

It is a specific property of composition operator for probability distributions — the operator is always defined in other settings [18,9]. In Section 3.5 we shall see that the credal sets framework is not an exception, either.

To make the concept clearer, let us list the results of its application in a few basic situations:

1. If $K\cap L=\emptyset$, then $(P_1 \triangleright P_2)(X_{K\cup L}) = P_1(X_K)\cdot P_2(X_L)$;
2. If $K=L$, then $(P_1 \triangleright P_2)(X_{K\cup L}) = P_1(X_K)$;
3. $(P_1 \triangleright P_2)(X_{K\cup L}) \neq (P_2 \triangleright P_1)(X_{K\cup L})$ in general; and
4. $(P_1 \triangleright P_2)(X_{K\cup L}) = (P_2 \triangleright P_1)(X_{K\cup L}) \Longleftrightarrow P_1(X_{K\cap L}) = P_2(X_{K\cap L})$.

Analogous results for credal sets will be proven/demonstrated by examples in Sections 3.3 and 3.4, respectively.

3.2. Definition

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be credal sets describing X_K and X_L, respectively. Our goal is to define a new credal set, denoted by $\mathcal{M}_1 \triangleright \mathcal{M}_2$, which will be describing $X_{K\cup L}$ and will contain all of the information contained in $\mathcal{M}_1$ and, as much as possible, in $\mathcal{M}_2$. The required properties are met by Definition 1, as we shall prove in Section 3.4.

Definition 1. For two credal sets $\mathcal{M}_1$ and $\mathcal{M}_2$ describing X_K and X_L, their *composition* $\mathcal{M}_1 \triangleright \mathcal{M}_2$ is defined as a convex hull of probability distributions P obtained as follows. For each pair of distributions $P_1 \in \mathcal{M}_1(X_K)$ and $P_2 \in \mathcal{M}_2(X_L)$ such that $P_2^{\downarrow K\cap L} \in argmin\{Q_2 \in \mathcal{M}_2(X_{K\cap L}) : d(Q_2, P_1^{\downarrow K\cap L})$, the distribution P is obtained by one of the following rules:

[a] If $P_1^{\downarrow K\cap L} \ll P_2^{\downarrow K\cap L}$

$$P(X_{K\cup L}) = \frac{P_1(X_K)\cdot P_2(X_L)}{P_2^{\downarrow K\cap L}(X_{K\cap L})}.$$

[b] Otherwise

$$P(X_{K\cup L}) \in \mathrm{ext}\{P_1^{\uparrow K\cup L}(X_K)\}.$$

Function d used in this definition is a suitable distance function (e.g., the Kullback–Leibler divergence, total variation or another f-divergence [16]). It should be noted that most of the important results presented in this paper are not affected by a specific choice of the distance function, as they are obtained for (at least) projective credal sets.

Application of Definition 1 to a few simple examples (with the goal to clarify the notion) will be the topic of the next subsection.

3.3. Examples

Let us now illustrate the application of the operator of composition and its properties in three examples. The first one shows what happens when $K\cap L=\emptyset$.

Example 2. Let $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$ be two credal sets describing R and T, respectively, with the values shown in Table 4 (interpretation of both R and T can again be found in Section 4.5).

$\mathcal{M}_1 \triangleright \mathcal{M}_2$ is obtained via [a] in Definition 1 and its values are contained in Table 5. One can easily see that both $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(R) = \mathcal{M}_1(R)$ and $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(T) = \mathcal{M}_2(T)$ hold true. $\square$

Table 5
$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(RT)$ – Example 2.

$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(RT)$	P^1		P^2		P^3		P^4	
T	0	1	0	1	0	1	0	1
$R=0$	0.1	0.1	0.16	0.04	0.25	0.25	0.4	0.4
$R=1$	0.4	0.4	0.64	0.16	0.25	0.25	0.1	0.1

Table 6
Credal set $\mathcal{M}_1(X_1X_2)$ – Example 3.

$\mathcal{M}_1(X_1X_2)$	P_1^1		P_1^2		P_1^3		P_1^4	
X_2	0	1	0	1	0	1	0	1
$X_1=0$	0.2	0.8	0.1	0.4	0.25	0.25	0	0
$X_1=1$	0	0	0.1	0.4	0.25	0.25	0.5	0.5

Table 7
Credal set $\mathcal{M}_2(X_2X_3)$ – Example 3.

$\mathcal{M}_2(X_2X_3)$	P^1		P^2		P^3		P^4	
X_3	0	1	0	1	0	1	0	1
$X_2=0$	0	0.3	0.2	0.1	0.5	0	0.2	0.3
$X_2=1$	0	0.7	0.4	0.3	0.5	0	0.2	0.3

Table 8
Marginal credal sets of non-projective credal sets – Example 3.

$\mathcal{M}_1(X_2)$	P_1^1	P_1^2
$X_2=0$	0.2	0.5
$X_2=1$	0.8	0.5

$\mathcal{M}_2(X_2)$	P_2^1	P_2^2
$X_2=0$	0.3	0.5
$X_2=1$	0.7	0.5

Table 9
Credal set $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$ – Example 3.

$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$		P^1		P^2		P^3		P^4	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0	0.3	0.2	0.1	0	0.1	0.07	0.03
	$X_2=1$	0	0.7	0.4	0.3	0	0.3	0.17	0.13
$X_1=1$	$X_2=0$	0	0	0	0	0	0.2	0.13	0.07
	$X_2=1$	0	0	0	0	0	0.4	0.23	0.17
		P^5		P^6		P^7		P^8	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0.25	0	0.1	0.15	0	0	0	0
	$X_2=1$	0.25	0	0.1	0.15	0	0	0	0
$X_1=1$	$X_2=0$	0.25	0	0.1	0.15	0.5	0	0.2	0.3
	$X_2=1$	0.25	0	0.1	0.15	0.5	0	0.2	0.3
		P^9		P^{10}		P^{11}		P^{12}	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0	0.2	0.13	0.07	0	0.1	0.07	0.03
	$X_2=1$	0	0.8	0.46	0.34	0	0.4	0.23	0.17
$X_1=1$	$X_2=0$	0	0	0	0	0	0.1	0.07	0.03
	$X_2=1$	0	0	0	0	0	0.4	0.23	0.17

The next example illustrates application of Definition 1 to two credal sets that are not projective.

Example 3. Let $\mathcal{M}_1(X_1X_2)$ and $\mathcal{M}_2(X_2X_3)$ be two credal sets describing pairs of binary variables X_1X_2 and X_2X_3, respectively. Their extreme points are contained in Tables 6 and 7.

These two credal sets are not projective, as can be seen from Table 8 containing both $\mathcal{M}_1(X_2)$ and $\mathcal{M}_2(X_2)$. Therefore $\mathcal{M}_2(X_2) \subset \mathcal{M}_1(X_2)$.

Definition 1 in this case leads (using total variation) to $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$ (see Table 9), while $(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$ is a smaller credal set, being a convex hull of distributions contained in Table 10. $\square$

Table 10
Credal set $(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$ – Example 3.

$(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$		P^1		P^2		P^3		P^4	
X_3		0	1	0	1	0	1	0	1
$X_1 = 0$	$X_2 = 0$	0	0.3	0.2	0.1	0	0.1	0.07	0.03
	$X_2 = 1$	0	0.7	0.4	0.3	0	0.3	0.17	0.13
$X_1 = 1$	$X_2 = 0$	0	0	0	0	0	0.2	0.13	0.07
	$X_2 = 1$	0	0	0	0	0	0.4	0.23	0.17
		P^5		P^6		P^7		P^8	
X_3		0	1	0	1	0	1	0	1
$X_1 = 0$	$X_2 = 0$	0.25	0	0.1	0.15	0	0	0	0
	$X_2 = 1$	0.25	0	0.1	0.15	0	0	0	0
$X_1 = 1$	$X_2 = 0$	0.25	0	0.1	0.15	0.5	0	0.2	0.3
	$X_2 = 1$	0.25	0	0.1	0.15	0.5	0	0.2	0.3

Table 11
Marginal credal sets $\mathcal{M}_1(X_1)$ and $\mathcal{M}_2(X_2)$ – Example 4.

$\mathcal{M}_1(X_1)$	P_1^1	P_1^2
$X_1 = 0$	0.2	0.7
$X_1 = 1$	0.8	0.3

$\mathcal{M}_2(X_2)$	P_2^1
$X_2 = 0$	0.5
$X_2 = 1$	0.5

Table 12
Credal set $\mathcal{M}_3(X_1X_2)$ – Example 4.

$\mathcal{M}_3(X_1X_2)$	P_3^1		P_3^2		P_3^3		P_3^4	
X_2	0	1	0	1	0	1	0	1
$X_1 = 0$	1	0	0	1	0	0	0	0
$X_1 = 1$	0	0	0	0	1	0	0	1

Table 13
$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2)$ (and also $((\mathcal{M}_1 \triangleright \mathcal{M}_2) \triangleright \mathcal{M}_3)(X_1X_2)$) – Example 4.

$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2)$	P^1		P^2	
X_2	0	1	0	1
$X_1 = 0$	0.1	0.1	0.35	0.35
$X_1 = 1$	0.4	0.4	0.15	0.15

Table 14
$(\mathcal{M}_2 \triangleright \mathcal{M}_3)(X_1X_2)$ – Example 4.

$(\mathcal{M}_2 \triangleright \mathcal{M}_3)(X_1X_2)$	P^1		P^2		P^3		P^4	
X_2	0	1	0	1	0	1	0	1
$X_1 = 0$	0.5	0.5	0.5	0	0	0.5	0	0
$X_1 = 1$	0	0	0	0.5	0.5	0	0.5	0.5

This difference deserves an explanation. $\mathcal{M}_2 \triangleright \mathcal{M}_1$ is smaller (more precise) than $\mathcal{M}_1 \triangleright \mathcal{M}_2$, which corresponds to the idea that we want $\mathcal{M}_2 \triangleright \mathcal{M}_1$ to keep all the information contained in $\mathcal{M}_2$. Therefore, we disregard those distributions from $\mathcal{M}_1$ that do not correspond to any from $\mathcal{M}_2$, although these distributions are taken into account when composing $\mathcal{M}_1 \triangleright \mathcal{M}_2$.

This is an example of a typical property of the operator of composition — it is not commutative. Neither is it associative, as can be seen from the following simple example.

Example 4. Let $\mathcal{M}_1(X_1)$ and $\mathcal{M}_2(X_2)$ be two credal sets describing X_1 and X_2, respectively, with the values shown in Table 11, and $\mathcal{M}_3(X_1X_2)$ be another credal set describing X_1X_2. Its values can be found in Table 12.

Then $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2)$ (Table 13) is obtained via option [a] in Definition 1 and $((\mathcal{M}_1 \triangleright \mathcal{M}_2) \triangleright \mathcal{M}_3)(X_1X_2) = (\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2)$ is true according to Property 2 in Lemma 1.

On the other hand, $(\mathcal{M}_2 \triangleright \mathcal{M}_3)(X_1X_2)$ is composed via option [b] in Definition 1. The result is shown in Table 14. Now, computing $(\mathcal{M}_1 \triangleright (\mathcal{M}_2 \triangleright \mathcal{M}_3))(X_1X_2)$ we obtain, again via option [b] in Definition 1, a credal set with the distributions shown in Table 15. It evidently differs from $((\mathcal{M}_1 \triangleright \mathcal{M}_2) \triangleright \mathcal{M}_3)(X_1X_2)$. $\square$

Table 15
$(\mathcal{M}_1 \triangleright (\mathcal{M}_2 \triangleright \mathcal{M}_3))(X_1X_2)$ – Example 4.

$(\mathcal{M}_1 \triangleright (\mathcal{M}_2 \triangleright \mathcal{M}_3))(X_1X_2)$	P^1		P^2		P^3		P^4	
X_2	0	1	0	1	0	1	0	1
$X_1 = 0$	0.2	0	0.2	0	0	0.2	0	0.2
$X_1 = 1$	0.8	0	0	0.8	0.8	0	0	0.8
	P^5		P^6		P^7		P^8	
X_2	0	1	0	1	0	1	0	1
$X_1 = 0$	0.7	0	0.7	0	0	0.7	0	0.7
$X_1 = 1$	0.3	0	0	0.3	0.3	0	0	0.3

3.4. Basic properties

In the following Lemma we prove that the composition operator possesses basic properties required above.

Lemma 1. *For two credal sets $\mathcal{M}_1$ and $\mathcal{M}_2$ describing X_K and X_L, respectively, the following properties hold true:*

1. *$\mathcal{M}_1 \triangleright \mathcal{M}_2$ is a credal set describing $X_{K\cup L}$.*
2. *$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_K) = \mathcal{M}_1(X_K)$.*
3. *$\mathcal{M}_1 \triangleright \mathcal{M}_2 = \mathcal{M}_2 \triangleright \mathcal{M}_1$ iff $\mathcal{M}_1(X_{K\cap L}) = \mathcal{M}_2(X_{K\cap L})$.*

Proof.

1. To prove that $\mathcal{M}_1 \triangleright \mathcal{M}_2$ is a credal set describing $X_{K\cup L}$, it is sufficient to take into consideration that it is the convex hull of probability distributions on $\mathbf{X}_{K\cup L}$, which is obvious from both [a] and [b] in Definition 1.
2. As marginalisation of a credal set is element-wise, it is sufficient to prove that, for any $P \in (\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L})$, $P^{\downarrow K} = P_1 \in \mathcal{M}_1(X_K)$ holds. In the case [a] it immediately follows from the results obtained for precise probabilities (see, e.g., [5]). In the case [b] it is obvious, as any P belongs to a vacuous extension of $P_1 \in \mathcal{M}_1(X_K)$ to $X_{K\cup L}$.
3. First, let us assume that

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L}) = (\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_{K\cup L}).$$

Then also its marginals must be identical; in particular

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cap L}) = (\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_{K\cap L}).$$

Taking into account Assertion 2 of this Lemma we obtain

$$\begin{aligned}(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cap L}) &= \left(((\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L}))^{\downarrow K}\right)^{\downarrow K\cap L} = ((\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_K))^{\downarrow K\cap L} \\ &= (\mathcal{M}_1(X_K))^{\downarrow K\cap L} = \mathcal{M}_1(X_{K\cap L})\end{aligned}$$

and similarly

$$(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_{K\cap L}) = \mathcal{M}_2(X_{K\cap L}),$$

from which the desired equality immediately follows.
Let us, on the other hand, suppose $\mathcal{M}_1(X_{K\cap L}) = \mathcal{M}_2(X_{K\cap L})$. In this case only option [a] in Definition 1 is applied, and for any distribution P in $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L})$ there exist $P_1 \in \mathcal{M}_1(X_K)$ and $P_2 \in \mathcal{M}_2(X_L)$ such that $P_1^{\downarrow K\cap L} = P_2^{\downarrow K\cap L}$ and $P = (P_1 \cdot P_2)/P_2^{\downarrow K\cap L}$. But simultaneously (due to projectivity) $P = (P_1 \cdot P_2)/P_1^{\downarrow K\cap L}$, which is an element of $(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_{K\cup L})$. Hence

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L}) = (\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_{K\cup L}),$$

as desired. □

3.5. Relationship to probabilistic case

As said in the Introduction, the operator of composition was originally introduced in (precise) probability theory. Nevertheless, any probability distribution may be viewed also as a singleton credal set (i.e., a credal set containing a single point). One would expect that the operator of composition we have introduced in this paper coincides with the probabilistic one if applied to a singleton credal set. And it is the case, as can be seen from the following Lemma.

Lemma 2. *Let $\mathcal{M}_1(X_K)$ and $\mathcal{M}_2(X_L)$ be two singleton credal sets (i.e., probability distributions) describing X_K and X_L, respectively, where $\mathcal{M}_1(X_{K\cap L})$ is absolutely continuous with respect to $\mathcal{M}_2(X_{K\cap L})$. Then $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L})$ is also a singleton credal set.*

Table 16
Singleton credal sets $\mathcal{M}_1(X_1X_2)$ and $\mathcal{M}_2(X_2X_3)$ – Example 5.

$\mathcal{M}_1(X_1X_2)$	P_1	
X_2	0	1
$X_1=0$	0.25	0.25
$X_1=1$	0.25	0.25

$\mathcal{M}_2(X_2X_3)$	P_2	
X_3	0	1
$X_2=0$	0.5	0.5
$X_2=1$	0	0

Table 17
$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3) = \mathcal{M}_1(X_1X_2)^{\uparrow\{1,2,3\}}$ – Example 5.

$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$		P^1		P^2		P^3		P^4	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0.25	0	0.25	0	0.25	0	0.25	0
	$X_2=1$	0	0.25	0	0.25	0	0.25	0	0.25
$X_1=1$	$X_2=0$	0	0.25	0	0.25	0.25	0	0.25	0
	$X_2=1$	0	0.25	0.25	0	0	0.25	0.25	0
		P^5		P^6		P^7		P^8	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0.25	0	0.25	0	0.25	0	0.25	0
	$X_2=1$	0.25	0	0.25	0	0.25	0	0.25	0
$X_1=1$	$X_2=0$	0	0.25	0	0.25	0.25	0	0.25	0
	$X_2=1$	0	0.25	0.25	0	0	0.25	0.25	0
		P^9		P^{10}		P^{11}		P^{12}	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0	0.25	0	0.25	0	0.25	0	0.25
	$X_2=1$	0	0.25	0	0.25	0	0.25	0	0.25
$X_1=1$	$X_2=0$	0	0.25	0	0.25	0.25	0	0.25	0
	$X_2=1$	0	0.25	0.25	0	0	0.25	0.25	0
		P^{13}		P^{14}		P^{15}		P^{16}	
X_3		0	1	0	1	0	1	0	1
$X_1=0$	$X_2=0$	0	0.25	0	0.25	0	0.25	0	0.25
	$X_2=1$	0.25	0	0.25	0	0.25	0	0.25	0
$X_1=1$	$X_2=0$	0	0.25	0	0.25	0.25	0	0.25	0
	$X_2=1$	0	0.25	0.25	0	0	0.25	0.25	0

Proof. Let us suppose that $\mathcal{M}_1 \triangleright \mathcal{M}_2$ is not a singleton, i.e., it contains at least two different points. Due to the condition of absolute continuity, both these points can be expressed in the form

$$P^i(X_{K\cup L}) = P_1^i(X_K) \cdot P_2^i(X_L)/(P_2^i)^{\downarrow K\cap L}(X_{K\cap L}),$$

$i=1,2$. As $P^1(X_{K\cup L}) \neq P^2(X_{K\cup L})$, it is evident that either $P_1^1(X_K) \neq P_1^2(X_K)$ or $P_2^1(X_L)/(P_2^1)^{\downarrow K\cap L}(X_{K\cap L}) \neq P_2^2(X_L)/(P_2^2)^{\downarrow K\cap L}(X_{K\cap L})$ (and therefore also $P_2^1(X_L) \neq P_2^2(X_L)$), or both. In any case, it is a contradiction as both credal sets $\mathcal{M}_1(X_K)$ and $\mathcal{M}_2(X_L)$ are singletons. □

The reader should, however, realise that the definition of the operator of composition for singleton credal sets is not completely equivalent to the definition of composition for probabilistic distributions. They equal each other only in the case when the probabilistic version is defined. This is ensured in Lemma 2 by assuming absolute continuity. If it does not hold, the probabilistic operator is not defined, while its credal version introduced in this paper is always defined (analogous to an evidential operator of composition). Nevertheless, in this case, the result is not a singleton credal set. We shall illustrate this fact by a simple example.

Example 5. Let $\mathcal{M}_1(X_1X_2)$ and $\mathcal{M}_2(X_2X_3)$ be two singleton credal sets describing variables X_1X_2 and X_2X_3, respectively. Their values are shown in Table 16. Let us compute $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$. As $\mathcal{M}_1(X_2) = \{[0.5, 0.5]\}$, while $(\mathcal{M}_2(X_2) = \{[1, 0]\})$, it is evident that $\mathcal{M}_1$ is not absolutely continuous with respect to $\mathcal{M}_2$. Therefore we have, via option [b] in Definition 1:

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3) = \mathcal{M}_1(X_1X_2)^{\uparrow\{1,2,3\}}.$$

The extreme points of this credal set, which is evidently not a singleton, are described in Table 17.

Let us remark that $(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$, in contrast to $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2X_3)$, is a singleton credal set (see Table 18), because $\mathcal{M}_2(X_2)$ is absolutely continuous with respect to $\mathcal{M}_1(X_2)$. □

Table 18
$(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$ – Example 5.

$(\mathcal{M}_2 \triangleright \mathcal{M}_1)(X_1X_2X_3)$	$X_3 = 0$		$X_3 = 1$	
X_2	0	1	0	1
$X_1 = 0$	0.25	0	0.25	0
$X_1 = 1$	0.25	0	0.25	0

3.6. Relationship to strong independence

The following Theorem, proven in [20], expresses the relationship between strong independence and the operator of composition. It is, together with Lemma 1, the most important assertion enabling us to introduce multidimensional models.

Theorem 1. *Let $\mathcal{M}$ be a credal set describing $X_{K\cup L}$ with marginals $\mathcal{M}(X_K)$ and $\mathcal{M}(X_L)$. Then*

$$\mathcal{M}(X_{K\cup L}) = (\mathcal{M}^{\downarrow K} \triangleright \mathcal{M}^{\downarrow L})(X_{K\cup L}) \tag{3}$$

iff

$$(K \setminus L) \perp\!\!\!\perp (L \setminus K) | (K \cap L). \tag{4}$$

Proof. Let us suppose that (3) holds. Since $\mathcal{M}_1(X_K)$ and $\mathcal{M}_2(X_L)$ are projective, option [a] in Definition 1 is applied and therefore

$$\mathcal{M}(X_{K\cup L}) = \{(P_1 \cdot P_2)/P_2^{\downarrow K\cap L} : P_1 \in \mathcal{M}(X_K), P_2 \in \mathcal{M}(X_L), P_1^{\downarrow K\cap L} = P_2^{\downarrow K\cap L}\}).$$

Proving (4) means finding, for any P from $\mathcal{M}(X_{K\cup L})$, a pair of projective distributions P_1 and P_2 from $\mathcal{M}(X_K)$ and $\mathcal{M}(X_L)$, respectively, such that $P = (P_1 \cdot P_2)/P_1{}^{\downarrow K\cap L}$. But due to the condition of projectivity, $\mathcal{M}(X_{K\cup L})$ consists of exactly this type of distributions.

Let, on the other hand, (4) be satisfied. Then any P from $\mathcal{M}(X_{K\cup L})$ can be expressed as a conditional product of its marginals, namely

$$P = (P^{\downarrow K} \cdot P^{\downarrow K})/P^{\downarrow K\cap L},$$

$P^{\downarrow K} \in \mathcal{M}(X_K)$ and $P^{\downarrow L} \in \mathcal{M}(X_L)$. Therefore,

$$\mathcal{M}(X_{K\cup L}) = \{(P^{\downarrow K} \cdot P^{\downarrow K})/P^{\downarrow K\cap L} : P^{\downarrow K} \in \mathcal{M}_1(X_K), P^{\downarrow L} \in \mathcal{M}_2(X_L))\},$$

which concludes the proof. $\square$

4. Compositional models

In this section we will consider repetitive application of the operator of composition with the goal to create a multidimensional model. Since the operator is neither commutative nor associative, we have to specify in which order the low-dimensional credal sets should be composed together. To make the formulae more transparent, we will omit parentheses in the case that the operator is to be applied from left to right, i.e., in what follows

$$\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3 \triangleright \ldots \triangleright \mathcal{M}_{n-1} \triangleright \mathcal{M}_n = (\ldots((\mathcal{M}_1 \triangleright \mathcal{M}_2) \triangleright \mathcal{M}_3) \triangleright \ldots \triangleright \mathcal{M}_{n-1}) \triangleright \mathcal{M}_n. \tag{5}$$

Furthermore, we will always assume that $\mathcal{M}_i$ is a credal set describing X_{K_i}.

4.1. Perfect sequences

The reader familiar with probabilistic, possibilistic or evidential compositional models knows that one of the most important notions in this theory is that of the so-called *perfect sequence*, which will now also be introduced for credal sets.

Definition 2. A generating sequence of credal sets $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ is called *perfect* if

$$\begin{aligned}
\mathcal{M}_1 \triangleright \mathcal{M}_2 &= \mathcal{M}_2 \triangleright \mathcal{M}_1,\\
\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3 &= \mathcal{M}_3 \triangleright (\mathcal{M}_1 \triangleright \mathcal{M}_2),\\
&\vdots\\
\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_n &= \mathcal{M}_n \triangleright (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{n-1}).
\end{aligned}$$

Table 19
$\mathcal{M}_3(RT)$ – Example 6.

$\mathcal{M}_3(RT)$	P_3^1		P_3^2		P_3^3		P_3^4	
T	0	1	0	1	0	1	0	1
$R=0$	0.2	0	0.2	0	0.5	0	0.5	0
$R=1$	0.3	0.5	0.6	0.2	0	0.5	0.3	0.2

Table 20
Two marginals of credal set $\mathcal{M}_3(RT)$ – Example 6.

$\mathcal{M}_3(R)$	P_1^1	P_1^2
$R=0$	0.2	0.5
$R=1$	0.8	0.5

$\mathcal{M}_3(T)$	P_2^1	P_2^2
$T=0$	0.5	0.8
$T=1$	0.5	0.2

It is evident that the necessary condition for perfectness is pairwise projectivity of low-dimensional credal sets. However, the following example demonstrates the fact that it need not be sufficient.

Example 6. Let $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$ be the same as in Example 2 and let $\mathcal{M}_3(RT)$ be defined by Table 19. It is evident that $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ are pairwise projective, as the marginals of $\mathcal{M}_3$ contained in Table 20 coincide with $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$, respectively, and $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$ are trivially projective, as already mentioned above. But they do not form a perfect sequence because

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3)(X_1X_2) = (\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_1X_2),$$

which is equal to the convex hull of distributions in Table 5, while

$$(\mathcal{M}_3 \triangleright (\mathcal{M}_1 \triangleright \mathcal{M}_2))(X_1X_2) = \mathcal{M}_3(X_1X_3),$$

which is different. □

Therefore a stronger condition, expressed by the following assertion, must be fulfilled.

Lemma 3. *A generating sequence $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ is perfect* iff *the pairs of credal sets $\mathcal{M}_j$ and $(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})$ are projective, i.e., if*

$$\mathcal{M}_j(X_{K_j\cap(K_1\cup\ldots\cup K_{j-1})}) = (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})(X_{K_j\cap(K_1\cup\ldots\cup K_{j-1})}),$$

for all $j = 2, 3, \ldots, n$.

Proof. This assertion is proved just by a multiple application of Assertion 3 of Lemma 1:

$$\begin{aligned}
&\mathcal{M}_1 \triangleright \mathcal{M}_2 = \mathcal{M}_2 \triangleright \mathcal{M}_1 \quad \Longleftrightarrow \quad \mathcal{M}_1(X_{K_2\cap K_1}) = \mathcal{M}_2(X_{K_2\cap K_1}),\\
&\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3 = \mathcal{M}_3 \triangleright (\mathcal{M}_1 \triangleright \mathcal{M}_2)\\
&\qquad\qquad\qquad\qquad \Longleftrightarrow \quad (\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K_3\cap(K_1\cup K_2)}) = \mathcal{M}_3(X_{K_3\cap(K_1\cup K_2)}),\\
&\qquad\qquad\qquad\qquad\qquad \vdots\\
&\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_n \mathcal{M}_n \triangleright (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}n-1)\\
&\qquad\qquad\qquad\qquad \Longleftrightarrow \quad (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{n-1})(X_{K_n\cap(K_1\cup\ldots\cup K_{n-1})})\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = \mathcal{M}_n(X_{K_n\cap(K_1\cup\ldots\cup K_{n-1})}),
\end{aligned}$$

which finishes the proof. □

4.2. Properties of perfect sequences

From Definition 2 one can hardly see what the properties of the perfect sequences are; the most important among them is expressed by the following characterisation Theorem, which, hopefully, also reveals why we call these sequences perfect.

Theorem 2. *A generating sequence of credal sets $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ is perfect* iff *all the credal sets from this sequence are marginal to the composed credal set $\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_n$:*

$$(\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_n)(X_{K_j}) = \mathcal{M}_j(X_{K_j}), \tag{6}$$

for all $j = 1, \ldots, m$.

Proof. The fact that all credal sets $\mathcal{M}_j$ from a perfect sequence are marginals of $(\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_n)$ follows from the fact that $(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_j)$ is marginal to $(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_n)$ (due to *(ii)* of Lemma 1) and $\mathcal{M}_j$ is marginal to

$$\mathcal{M}_j \triangleright (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1}) = \mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_j.$$

Suppose now that, for all $j = 1, \ldots, n$, $\mathcal{M}_j$ are marginal credal sets to $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_n$. It means that all the credal sets from the sequence are pairwise projective, and that each $\mathcal{M}_j$ is projective with any marginal credal set of $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_n$, and consequently also with $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1}$. So we get that

$$\mathcal{M}_j(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})}) = (\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})})$$

for all $j = 2, \ldots, n$, which is equivalent, due to Lemma 3, to the fact that $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ is perfect. $\square$

This result is quite close to Theorem 5 in [13], expressing the joint coherence of separately coherent conditional lower previsions. Condition (6) is, roughly speaking, substituted by the requirement of coherence of a marginal of the resulting model with the corresponding original coherent conditional lower prevision (for more details see [13]).

The following (almost trivial) assertion, which establishes a sufficient condition for perfectness, resembles assertions concerning decomposable models.

Theorem 3. *Let a generating sequence of pairwise projective credal sets $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ be such that $K_1, K_2, \ldots, K_n$ meets the well-known running intersection property:*

$$\forall j = 2, 3, \ldots, n \ \exists \ell (1 \leq \ell < j) \quad \text{such that } K_j \cap (K_1 \cup \ldots \cup K_{j-1}) \subseteq K_\ell.$$

Then $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_n$ is perfect.

Proof. Due to Lemma 3 it is sufficient to show that, for each $j = 2, \ldots, n$ credal set $\mathcal{M}_j$ and the composed credal set $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1}$ are projective. Let us prove it by induction.

For $j = 2$ the required projectivity is guaranteed by the fact that we assume pairwise projectivity of all $\mathcal{M}_1, \ldots, \mathcal{M}_n$. So we have to prove it for general $j > 2$ under the assumption that the assertion holds for $j - 1$, which means (due to Theorem 2) that all $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_{j-1}$ are marginal to $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1}$. Since we assume that $K_1, \ldots, K_n$ meets the running intersection property, there exists $\ell \in \{1, 2, \ldots j-1\}$ such that $K_j \cap (K_1 \cup \ldots \cup K_{j-1}) \subseteq K_\ell$. Therefore $(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})})$ and $\mathcal{M}_\ell(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})})$ are the same marginals of $\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1}$ and therefore they have to be equal to each other:

$$(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})}) = \mathcal{M}_\ell(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})}).$$

However, we assume that $\mathcal{M}_j$ and $\mathcal{M}_\ell$ are projective and therefore also

$$(\mathcal{M}_1 \triangleright \ldots \triangleright \mathcal{M}_{j-1})(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})}) = \mathcal{M}_j(X_{K_j \cap (K_1 \cup \ldots \cup K_{j-1})}),$$

as desired. $\square$

It should be pointed out at this moment that the running intersection property of $K_1, K_2, \ldots, K_n$ is a sufficient condition which guarantees perfectness of a generating sequence of pairwise projective credal sets. By no means is this condition necessary, as will be shown in the following example.

Example 7. A simple example is given by two credal sets $\mathcal{M}_1$ and $\mathcal{M}_2$ from Example 6 describing X_1 and X_2, respectively, and the third credal set $\tilde{\mathcal{M}}_3 = \mathcal{M}_1 \triangleright \mathcal{M}_2$. Considering sequence $\mathcal{M}_1, \mathcal{M}_2, \tilde{\mathcal{M}}_3$, it is evident that $K_1 = \{1\}, K_2 = \{2\}, K_3 = \{1, 2\}$ do not meet the running intersection property. And yet the sequence $\mathcal{M}_1, \mathcal{M}_2, \tilde{\mathcal{M}}_3$ is perfect because all the credal sets are marginal (or equal) to $\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3$. Notice that if we chose, instead of $\tilde{\mathcal{M}}_3$, any other credal set $\mathcal{M}_3$ describing X_1X_2 different from $\tilde{\mathcal{M}}_3 = \mathcal{M}_1 \triangleright \mathcal{M}_2$, e.g., that from Example 6, the generating sequence $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ would no longer be perfect. $\square$

So we can see that perfectness of a sequence is not only a structural property connected with the properties of $K_1, K_2, \ldots, K_n$; it also depends on specific values of the respective credal sets.

We can summarize, that properties of perfect sequences presented up to this point completely correspond to those possessed by analogous models not only in (precise) probability framework, but also in possibility and evidence theories.

4.3. Perfect sequence as convex hull

In this subsection we will study the relationship between perfect sequences of credal sets and those of probability distribution. Before doing that, let us present a simple Lemma necessary for the proof of the main Theorem.

Lemma 4. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two projective credal sets describing X_K and X_L, respectively. Then*

$$\begin{aligned}&\{\mathrm{ext}((\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_K \cup X_L))\}\\&\quad\subseteq \{P_1 \triangleright P_2 : P_1 \in \mathrm{ext}(\mathcal{M}_1(X_K)),\ P_2 \in \mathrm{ext}(\mathcal{M}_2(X_L)),\ P_1^{\downarrow K\cap L} = P_2^{\downarrow K\cap L}\}.\end{aligned} \tag{7}$$

Proof. By Definition 1, $(\mathcal{M}_1 \triangleright \mathcal{M}_2)(X_{K\cup L})$ is a convex hull of the set of probability distributions from the set in the right hand side of (7), taking into account the definition of composition operator for precise probabilities. Therefore its extreme points must also belong to this set. □

The equality need not hold in (7), as can be seen from the following simple example.

Example 8. Let $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$ be two credal sets describing R and T, respectively, as defined in Example 2. One can easily see that P^3 is not an extreme point of the credal set defined as a convex hull of distributions from Table 5 because it can be obtained as a linear combination of P^1 and P^4. □

Theorem 4. *Let $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_m$ be a perfect sequence of credal sets such that each $\mathcal{M}_i, i = 1, \dots m$, is the convex hull of its extreme points, i.e.,*

$$\mathcal{M}_i(X_{K_i}) = \mathrm{CH}\{P_i : P_i \in \mathrm{ext}(\mathcal{M}_i(X_{K_i}))\}.$$

Then

$$\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_m$$

is a convex hull of all

$$P_1 \triangleright P_2 \triangleright \ldots \triangleright P_m$$

such that each $P_i \in \mathrm{ext}(\mathcal{M}_i(X_{K_i}))$ and $P_1, P_2, \dots, P_m$ form a perfect sequence.

Proof. Let us prove the assertion by induction. For $m = 2$ it is obvious as it follows directly from Definition 1. Let us suppose that

$$\begin{aligned}&\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_j\\&\quad= \mathrm{CH}\{P_1 \triangleright P_2 \triangleright \ldots \triangleright P_j,\ P_i \in \mathrm{ext}(\mathcal{M}_i),\ P_1, P_2, \dots, P_j \text{ is perfect}\}\end{aligned}$$

for $2 \le j < m$ and prove that

$$\begin{aligned}&\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_{j+1}\\&\quad= \mathrm{CH}\{P_1 \triangleright P_2 \triangleright \ldots \triangleright P_{j+1},\ P_i \in \mathrm{ext}(\mathcal{M}_i),\ P_1, P_2, \dots, P_{j+1} \text{ is perfect}\}\end{aligned} \tag{8}$$

holds as well.

By convention (5)

$$\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_j \triangleright \mathcal{M}_{j+1} = (\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_j) \triangleright \mathcal{M}_{j+1}$$

and since $\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \ldots \triangleright \mathcal{M}_j$ and $\mathcal{M}_{j+1}$ are projective, we can apply Definition 1 to these credal sets and obtain

$$\begin{aligned}&(\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_j) \triangleright \mathcal{M}_{j+1}\\&\quad= \mathrm{CH}\{Q_j \cdot \frac{P_{j+1}}{P_{j+1}^{\downarrow(K_1\cup\cdots\cup K_j)\cap K_{j+1}}},\ Q_j \in \mathrm{ext}(\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_j),\\&\qquad\qquad P_{j+1} \in \mathrm{ext}(\mathcal{M}_{j+1}),\ Q_j^{\downarrow(K_1\cup\cdots\cup K_j)\cap K_{j+1}} = P_{j+1}^{\downarrow(K_1\cup\cdots\cup K_j)\cap K_{j+1}}\}.\end{aligned}$$

However, due to Lemma 4

$$Q_j \in \{P_1 \triangleright P_2 \triangleright \ldots \triangleright P_j,\ P_i \in \mathrm{ext}(\mathcal{M}_i),\ P_1, P_2, \dots, P_j \text{ is perfect}\}.$$

Table 21
Credal set $\mathcal{M}_3(RTG)$.

$\mathcal{M}_3(RTG)$		P_3^1		P_3^2		P_3^3	
G		0	1	0	1	0	1
$R=0$	$T=0$	0	0.1	0	0.16	0	0.4
	$T=1$	0.1	0.3	0.01	0.03	0.075	0.025
$R=1$	$T=0$	0.05	0.05	0.32	0.32	0.2	0.2
	$T=1$	0.3	0.1	0.04	0.12	0.025	0.075

Let us denote by $P_1^*, P_2^*, \ldots, P_j^*$ a perfect sequence such that

$$Q_j = P_1^* \triangleright P_2^* \triangleright \ldots \triangleright P_j^*.$$

Then, due to Lemma 3 (applied to precise probability distributions) $P_1^*, P_2^*, \ldots, P_j^*, P_{j+1}$ form a perfect sequence. Therefore

$$\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_{j+1} \subseteq \mathrm{CH}\{P_1 \triangleright P_2 \triangleright \ldots \triangleright P_{j+1}, P_i \in \mathrm{ext}(\mathcal{M}_i), P_1, P_2, \ldots, P_{j+1} \text{ is perfect}\}.$$

Let, on the other hand, $P_1, P_2, \ldots, P_{j+1}$ be a perfect sequence of distributions such that each $P_i \in \mathrm{ext}(\mathcal{M}_i)$. Then

$$P_1 \triangleright P_2 \triangleright \ldots \triangleright P_{j+1} \in \mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_{j+1},$$

and therefore also

$$\mathrm{CH}\{P_1 \triangleright P_2 \triangleright \ldots \triangleright P_{j+1}, P_1, P_2, \ldots, P_{j+1} \text{ is perfect}\} \subseteq \mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \cdots \triangleright \mathcal{M}_{j+1}.$$

Therefore (8) is satisfied. □

Let us note, that an analogous result was obtained in the framework of possibility theory for product t-norm (for more details see [19]).

4.4. Inference from credal compositional models

Up to now, little work was done in the direction of inference on even (precise) probabilistic compositional models. Some preliminary results can be found in [2,6,7]. The situation is even more complicated in our models with inherent imprecision.

One can profit from Theorem 4 and make inference from perfect sequences of probability distributions, but still, no universal procedure exists. In a special case when the conditions of Theorem 3 are satisfied, local computations by Lauritzen and Spiegelhalter [12] for any of these perfect sequences can be applied. Finally a convex hull of the probabilities of interest must be constructed.

This approach is demonstrated in a small illustrative example, which is the content of the next subsection. Nevertheless, we know that it can hardly be used in practical situations and the need for an efficient computational procedure is obvious.

4.5. Illustrative example

Our example, simplified from [10], will describe relationships among events influencing the fact whether Joan goes for her regular evening walk or not. All the possible situations will be described with the help of the following four variables:

W — describes whether Joan goes for her evening walk;
R — corresponds to the evening weather conditions: if it rains or not;
G — describes one of Joan's friends' intention to come to pay her a visit;
T — expresses Joan's attitude to the evening TV programme.

The overall model will be composed from some credal sets described in the preceding parts of this paper, more precisely from credal sets $\mathcal{M}_1(R)$ and $\mathcal{M}_2(T)$ from Example 2, $\mathcal{M}_4(G, W) = \mathcal{M}(G, W)$ from Example 1 and $\mathcal{M}_3(R, T, G)$ defined by Table 21.

These credal sets form a perfect sequence $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{M}_4$, as $\mathcal{M}_1 \triangleright \mathcal{M}_2$ is marginal to $\mathcal{M}_3$ and $\mathcal{M}_3$ and $\mathcal{M}_4$ are projective. The credal set $\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3 \triangleright \mathcal{M}_4(R, T, G, W)$ is the convex hull of the distributions from Table 22. This credal set can be expressed as a convex hull of $P_1^{i_1} \triangleright P_2^{i_2} \triangleright P_3^{i_3} \triangleright P_4^{i_4}$, where any $P_1^{i_1}, P_2^{i_2}, P_3^{i_3}, P_4^{i_4}$ form a perfect sequence and $P_j^{i_j} \in ext(\mathcal{M}_j)$ is true for any $P_j^{i_j}$. We have six perfect sequences in this example, namely,

Table 22
Joint model "evening walk".

$\mathcal{M}_1 \triangleright \mathcal{M}_2 \triangleright \mathcal{M}_3 \triangleright \mathcal{M}_4$		P^1				P^2			
		$G=0$		$G=1$		$G=0$		$G=1$	
W		0	1	0	1	0	1	0	1
$R=0$	$T=0$	0	0	0.08	0.02	0	0	0.06	0.04
	$T=1$	0.04	0.06	0.24	0.06	0.02	0.08	0.18	0.12
$R=1$	$T=0$	0.02	0.03	0.04	0.01	0.01	0.04	0.03	0.02
	$T=1$	0.12	0.18	0.08	0.02	0.06	0.24	0.06	0.04
		P^3				P^4			
		$G=0$		$G=1$		$G=0$		$G=1$	
W		0	1	0	1	0	1	0	1
$R=0$	$T=0$	0	0	0.128	0.032	0	0	0.096	0.064
	$T=1$	0.004	0.006	0.024	0.006	0.002	0.008	0.018	0.012
$R=1$	$T=0$	0.128	0.192	0.256	0.064	0.064	0.256	0.192	0.128
	$T=1$	0.048	0.072	0.032	0.008	0.024	0.096	0.024	0.016
		P^5				P^6			
		$G=0$		$G=1$		$G=0$		$G=1$	
W		0	1	0	1	0	1	0	1
$R=0$	$T=0$	0	0	0.32	0.08	0	0	0.24	0.16
	$T=1$	0.03	0.045	0.02	0.005	0.015	0.06	0.015	0.01
$R=1$	$T=0$	0.08	0.12	0.16	0.04	0.04	0.16	0.12	0.08
	$T=1$	0.01	0.015	0.06	0.015	0.005	0.02	0.045	0.03

$$\begin{array}{ll} P_1^1, P_2^1, P_3^1, P_4^1; & P_1^1, P_2^1, P_3^1, P_4^3; \\ P_1^1, P_2^2, P_3^2, P_4^1; & P_1^1, P_2^2, P_3^2, P_4^3; \\ P_1^2, P_2^2, P_3^3, P_4^2; & P_1^2, P_2^2, P_3^3, P_4^4, \end{array} \tag{9}$$

where $P_1^{i_1}$ and $P_2^{i_2}$ are probabilities from Table 4, $P_3^{i_3}$ from Table 21, and $P_4^{i_4}$ from Table 1.

To make inference from this model, it is necessary to do it for any of the six models

$$P_1^{i_1}(R) \cdot P_2^{i_2}(T) \cdot \frac{P_3^{i_3}(RTG)}{P_3^{i_3}(RT)} \cdot \frac{P_4^{i_4}(GW)}{P_4^{i_4}(G)}$$

formed by these perfect sequences as suggested in the preceding subsection. It follows from Theorem 1 that $R \perp\!\!\!\perp T$ and $RT \perp\!\!\!\perp W|G$. These independence relationships can be utilized to simplify the inference.

Our goal is to find the probability that Joan invited a guest knowing that it was not raining, the TV programme was good and she did not go for a walk. Any of these conditionals will be expressed in the form

$$P_i(G=1|R=0,T=1,W=0)$$
$$= \frac{P_3^{i_3}(R=0,T=1,G=1) \cdot \frac{P_4^{i_4}(G=1,W=0)}{P_4^{i_4}(G=1)}}{P_3^{i_3}(R=0,T=1,G=1) \cdot \frac{P_4^{i_4}(G=1,W=0)}{P_4^{i_4}(G=1)} + P_3^{i_3}(R=0,T=1,G=0) \cdot \frac{P_4^{i_4}(G=0,W=0)}{P_4^{i_4}(G=0)}},$$

$i=1,\ldots,6$. Computing this expression, we will obtain the following values: 0.86, 0.9, 0.86, 0.9, 0.4, 0.5. Therefore $P(G=1|R=0,T=1,W=0) \in [0.4, 0.9]$ (and hence $P(G=0|R=0,T=1,W=0) \in [0.1, 0.6]$). Similarly we can proceed for other combinations of values.

5. Conclusions

We have defined the composition operator for credal sets, manifesting all the main characteristics of its probabilistic pre-image. Even more, there is one point in which the credal set operator of composition is superior to the probabilistic one: thanks to the ability of credal sets to model total ignorance, the composition operator is always defined for credal sets, which is not the case in the (precise) probabilistic framework.

We have proved the basic properties of the operator (including the relationship to strong independence) necessary for the introduction of compositional models and their most important special case, *perfect sequence models*. We have also found the relationship between perfect sequences of credal sets and those of their extreme points.

Naturally, there are still many open problems to be solved. The most important one is the problem of effective finding of the nearest probability distributions (if there is no projective) and, in general, a design of efficient computational procedures for this type of model with special attention paid to inference.

At this moment we know very little about the relationship between the compositional models developed for credal sets and those introduced in possibility theory [17,18] and in evidence [8] theories (apart from formal analogies). Another interesting class of problems are similarities and differences between the described compositional models and other multidimensional models within the framework of credal sets such as [1,4,15].

Acknowledgements

The support of Grant GAČR 16-12010S is gratefully acknowledged.

References

[1] A. Benavoli, A. Antonucci, Aggregating imprecise probabilistic knowledge, in: T. Augustin, F. Coolen, S. Moral, M.C.M. Troffaes (Eds.), ISIPTA '09: Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications, SIPTA, 2009, pp. 31–40.
[2] Vl. Bína, R. Jiroušek, Effective algorithm of marginalization in multidimensional compositional models, in: T. Kroupa, J. Vejnarová (Eds.), Proceedings of 8th Czech–Japan Seminar on Data Analysis and Decision Making under Uncertainty, Oeconomica, Tøe, 2005, pp. 1–13.
[3] I. Couso, S. Moral, P. Walley, Examples of independence for imprecise probabilities, in: G. De Cooman, F.G. Cozman, S. Moral, P. Walley (Eds.), Proceedings of ISIPTA'99, Ghent, 1999, pp. 121–130.
[4] F.G. Cozman, Credal networks, AI Jānaru 120 (2000) 199–233.
[5] R. Jiroušek, Composition of probability measures on finite spaces, in: D. Geiger, P.P. Shenoy (Eds.), Proc. of UAI'97, Morgan Kaufmann Publ., San Francisco, California, 1997, pp. 274–281.
[6] R. Jiroušek, Vl. Bína, Theoretical foundations for marginalization in compositional models, in: J. Watada (Ed.), Proceedings of 9th Czech–Japan Seminar on Data Analysis and Decision Making under Uncertainty, Kitakyushu, Japan, 2006, pp. 224–231.
[7] R. Jiroušek, V. Kratochvíl, Marginalization algorithm for compositional models, in: B. Bouchon-Meunier, R.R. Yager (Eds.), Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems, EDK, Paris, 2006, pp. 2300–2307.
[8] R. Jiroušek, J. Vejnarová, Compositional models and conditional independence in evidence theory, Int. J. Approx. Reason. 52 (2011) 316–334.
[9] R. Jiroušek, J. Vejnarová, M. Daniel, Compositional models for belief functions, in: G. De Cooman, J. Vejnarová, M. Zaffalon (Eds.), Proc. of ISIPTA'07, Mat-fyz Press, Praha, 2007, pp. 243–252.
[10] R. Jiroušek, J. Vejnarová, J. Gemela, Possibilistic belief network constructed by operators of composition and its application to financial analysis, in: R.P. Srivastava, T.J. Mock (Eds.), Belief Functions in Business Decisions, Physica-Verlag, Heidelberg, 2002, pp. 252–280.
[11] S.L. Lauritzen, Graphical Models, Oxford University Press, 1996.
[12] S.L. Lauritzen, D.J. Spiegelhalter, Local computations with probabilities on graphical structures and their application to expert systems, J. R. Stat. Soc. B 50 (1988) 157–224.
[13] E. Miranda, Updating coherent previsions on finite spaces, Fuzzy Sets Syst. 160 (2009) 1286–1307.
[14] S. Moral, A. Cano, Strong conditional independence for credal sets, Ann. Math. Artif. Intell. 35 (2002) 295–321.
[15] S. Moral, J. Sagrado, Aggregation of imprecise probabilities, in: Aggregation and Fusion of Imperfect Information, 1997, pp. 162–188.
[16] I. Vajda, Theory of Statistical Inference and Information, Kluwer Academic Publishers, Dordrecht, 1989.
[17] J. Vejnarová, Composition of possibility measures on finite spaces: preliminary results, in: B. Bouchon-Meunier, R.R. Yager (Eds.), Proc. of IPMU'98, Editions E.D.K, Paris, 1998, pp. 25–30.
[18] J. Vejnarová, On possibilistic marginal problem, Kybernetika 43 (5) (2007) 657–674.
[19] J. Vejnarová, On interpretation of T-product extensions of possibility distributions, Fuzzy Sets Syst. 232 (1) (2013) 3–17.
[20] J. Vejnarová, Operator of composition for credal sets, in: F.G. Cozman, T. Denoeux, S. Destercke, T. Seidenfeld (Eds.), Proc. of ISIPTA'13, 2013, pp. 355–364.
[21] J. Vejnarová, Credal compositional models and credal networks, in: T. Augustin, S. Doria, E. Miranda, E. Quaeghebeur (Eds.), Proc. of ISIPTA'15, 2015, pp. 315–324.