Two Composition Operators for Belief Functions Revisited

Radim Jiroušek[1,2], Václav Kratochvíl[1,2], and Prakash P. Shenoy[3]

[1]Faculty of Management, University of Economics, Jindřichův Hradec, Czechia
[2]Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czechia
[1,2] *{radim, velorex}@utia.cas.cz*
[3]School of Business, University of Kansas, Lawrence, KS, USA
[3] *pshenoy@ku.edu*

June 10, 2022

Abstract

In probability theory, compositional models are as powerful as Bayesian networks. However, the relation between belief-function graphical models and the corresponding compositional models is much more complicated due to several reasons. One of them is that there are two composition operators for belief functions. This paper deals with their main properties and presents sufficient conditions under which they yield the same results.

1 Introduction

Two different composition operators for belief functions are defined in the literature (Jiroušek et al., 2007; Jiroušek and Shenoy, 2014). Surprisingly, for directed graphical belief function models, e.g., Almond's 'Captain's problem' (Almond, 1995), the corresponding compositional models are the same (regardless of the operator used). This unexpected finding is surprising since the two operators are designed based on different ideas and for different purposes. Historically, the first operator (Jiroušek et al., 2007), called the *f-composition* operator here, is designed to represent multivariate basic probabilistic assignments (BPA) using the lowest number of parameters. The second operator (Jiroušek et al., 2007) is consistent with the Dempster-Shafer (D-S) theory of evidence, and therefore, we call it the *d-composition* operator. The d-composition operator introduces conditional independence relations among the variables, similar to the probabilistic composition operator. Thus, the class of d-compositional models is equivalent to the class of directed probabilistic graphical models. In general, this is not true for f-compositional models.

The idea behind the f-composition operator is to decrease the number of parameters necessary for representing multidimensional belief functions. For example, instead of representing one three-dimensional BPA, one represents only two two-dimensional BPAs. In the case of binary variables it means that one can use only $2 \times 2^{(2^2)} = 32$ instead of $2^{(2^3)} = 128$ parameters. Naturally, there is no free lunch, and one has to pay for it by restricting the class of such BPAs. One has to give up the possibility of using belief functions whose BPAs are not factorizable.

In probability theory, there is a *factorization lemma*[1] that says if a joint probability distribution P for variables X, Y, and Z can be expressed in the form of a product of two factors $\phi_1(X,Y)$ and $\phi_2(Y,Z)$, then X and Z are conditionally independent given Y, written as $X \perp\!\!\!\perp Z|Y$. Therefore, for a probabilistic compositional model, one can identify the induced conditional independence relations based on the factors in the model. There is a similar result for the D-S theory (Shenoy, 1994) and thus, for the d-composition operator. However, it is not clear what belief function theory corresponds to the f-composition operator, and therefore the problem of identification of the conditional independence relations for f-composition models is not obvious. However, there are other problems associated with the d-composition operator. The result of d-composition is sometimes undefined (see Definition 2 in Section 3).

Thus, it is not clear which composition operator is better. The user should choose the one which suits better the purpose of the application. Both of them satisfy the properties expected from composition operators (described in Section 3). Both composition operators have corresponding inverse decomposition operators. One of them corresponds to the notion of conditional independence, the other to a specific way of factorization. We will not study decompositions explicitly in this paper.

An outline of the remainder of the paper is as follows. Section 2 introduce the necessary concepts and notation from belief function theory. Section 3 contains definitions of the two composition operators. Section 4 has the main result of this paper. Section 5 illustrates the main result using Almond's captain's problem (Almond, 1995). Section 6 has a summary and some concluding remarks.

2 Belief Functions

Let $\mathcal{W}$ denote a set of variables each with finite number of states. For $X \in \mathcal{W}$, let Ω_X denote the set of states of variable X. A *basic probability assignment* (BPA) for variables $\mathcal{U} \subseteq \mathcal{W}$ (or equivalently, a BPA on the Cartesian product $\Omega_{\mathcal{U}} = \times_{X \in \mathcal{U}} \Omega_X$) is a mapping $m_{\mathcal{U}} : 2^{\Omega_{\mathcal{U}}} \to [0,1]$, such that $\sum_{\mathbf{a} \subseteq \Omega_{\mathcal{U}}} m_{\mathcal{U}}(\mathbf{a}) = 1$ and $m_{\mathcal{U}}(\emptyset) = 0$.

Consider a BPA $m_{\mathcal{U}}$. If the set of the corresponding variables is clear from the context, we omit the subscript $\mathcal{U}$. We say $\mathbf{a}$ is a *focal element* of m if $m(\mathbf{a}) > 0$. If m has only one

[1] It is not the same as the factorization lemma from the theory of categories.

focal element, we say m is *deterministic.* If this focal element is $\Omega_{\mathcal{U}}$, i.e., $m(\Omega_{\mathcal{U}}) = 1$, we say that m is *vacuous.*

Given a BPA m, the same information can be expressed by the corresponding *commonality* function (which is also defined on the power set 2^{Ω}):

$$Q_m(\mathbf{a}) = \sum_{\mathbf{b}\subseteq\Omega:\, \mathbf{b}\supseteq\mathbf{a}} m(\mathbf{b}). \tag{1}$$

Whenever a commonality function is given, it is possible to reconstruct the corresponding BPA m:

$$m(\mathbf{a}) = \sum_{\mathbf{b}\subseteq\Omega:\mathbf{b}\supseteq\mathbf{a}} (-1)^{|\mathbf{b}\setminus\mathbf{a}|} Q_m(\mathbf{b}). \tag{2}$$

For BPA $m_{\mathcal{V}}$, we often consider its *marginal* $m_{\mathcal{V}}^{\downarrow\mathcal{U}}$ for $\mathcal{U} \subseteq \mathcal{V}$. A similar notation is used also for *projections* of states. If $a \in \Omega_{\mathcal{V}}$, $a^{\downarrow\mathcal{U}}$ denotes the element of $\Omega_{\mathcal{U}}$, which is obtained from a by omitting the values of variables in $\mathcal{V} \setminus \mathcal{U}$. For $\mathbf{a} \subseteq \Omega_{\mathcal{V}}$,

$$\mathbf{a}^{\downarrow\mathcal{U}} = \{a^{\downarrow\mathcal{U}} : a \in \mathbf{a}\}.$$

Using this notation, the marginal $m_{\mathcal{V}}^{\downarrow\mathcal{U}}$ of BPA $m_{\mathcal{V}}$ for $\mathcal{U} \subseteq \mathcal{V}$ is defined as follows:

$$m_{\mathcal{V}}^{\downarrow\mathcal{U}}(\mathbf{b}) = \sum_{\mathbf{a}\subseteq\Omega_{\mathcal{V}}:\, \mathbf{a}^{\downarrow\mathcal{U}}=\mathbf{b}} m_{\mathcal{V}}(\mathbf{a}).$$

for all $\mathbf{b} \subseteq \Omega_{\mathcal{U}}$.

The projection of sets enables us to define a *join* of two sets. Consider two arbitrary sets $\mathcal{U}$ and $\mathcal{V}$ of variables (they may be disjoint or overlapping, or one may be a subset of the other). Consider two sets $\mathbf{a} \subseteq \Omega_{\mathcal{U}}$ and $\mathbf{b} \subseteq \Omega_{\mathcal{V}}$. Their join is defined as

$$\mathbf{a} \bowtie \mathbf{b} = \{c \in \Omega_{\mathcal{U}\cup\mathcal{V}} : c^{\downarrow\mathcal{U}} \in \mathbf{a} \;\;\&\;\; c^{\downarrow\mathcal{V}} \in \mathbf{b}\}.$$

Notice that if $\mathcal{U}$ and $\mathcal{V}$ are disjoint, then $\mathbf{a} \bowtie \mathbf{b} = \mathbf{a} \times \mathbf{b}$. If $\mathcal{U} = \mathcal{V}$, then $\mathbf{a} \bowtie \mathbf{b} = \mathbf{a} \cap \mathbf{b}$. In general, for $\mathbf{c} \subseteq \Omega_{\mathcal{U}\cup\mathcal{V}}$, $\mathbf{c}$ is a subset of $\mathbf{c}^{\downarrow\mathcal{U}} \bowtie \mathbf{c}^{\downarrow\mathcal{V}}$, which may be a proper subset. If $\mathbf{c}^{\downarrow\mathcal{U}\cap\mathcal{V}}$ is a singleton subset, then $\mathbf{c} = \mathbf{c}^{\downarrow\mathcal{U}} \bowtie \mathbf{c}^{\downarrow\mathcal{V}}$.

To construct multidimensional models from low-dimensional building blocks, we need some operators connecting two low-dimensional BPAs into one BPA. One possibility is the classical Dempster's combination rule, which is used to combine distinct belief functions. Consider two BPAs $m_{\mathcal{U}}$ and $m_{\mathcal{V}}$ for arbitrary sets of variables $\mathcal{U}$ and $\mathcal{V}$. Dempster's combination, denoted by $\oplus$, is defined as follows (Shafer, 1976):

$$(m_{\mathcal{U}} \oplus m_{\mathcal{V}})(\mathbf{c}) = \frac{1}{1-K} \sum_{\mathbf{a}\subseteq\Omega_{\mathcal{U}},\mathbf{b}\subseteq\Omega_{\mathcal{V}}:\mathbf{a}\bowtie\mathbf{b}=\mathbf{c}} m_{\mathcal{U}}(\mathbf{a}) \cdot m_{\mathcal{V}}(\mathbf{b}), \tag{3}$$

for each $\mathbf{c} \subseteq \Omega_{\mathcal{U}\cup\mathcal{V}}$, where

$$K = \sum_{\mathbf{a}\subseteq\Omega_{\mathcal{U}},\mathbf{b}\subseteq\Omega_{\mathcal{V}}:\mathbf{a}\bowtie\mathbf{b}=\emptyset} m_{\mathcal{U}}(\mathbf{a}) \cdot m_{\mathcal{V}}(\mathbf{b}). \tag{4}$$

K can be interpreted as the amount of conflict between $m_{\mathcal{U}}$ and $m_{\mathcal{V}}$. If $K = 1$, we say $m_{\mathcal{U}}$ and $m_{\mathcal{V}}$ are in *total conflict* and their Dempster's combination is undefined.

3 Composition Operator

The following definition answers the question: What do we mean by a belief function composition operator?

Definition 1 *By an* composition operator $\triangleright$ *we mean any binary operator satisfying the following four axioms. Consider arbitrary three BPAs* $m_{\mathcal{T}}$, $m_{\mathcal{U}}$, *and* $m_{\mathcal{V}}$.

A1 (Domain): $m_{\mathcal{T}} \triangleright m_{\mathcal{U}}$ *is a BPA for* $\mathcal{T} \cup \mathcal{U}$.

A2 (Composition preserves first marginal): $(m_{\mathcal{T}} \triangleright m_{\mathcal{U}})^{\downarrow \mathcal{T}} = m_{\mathcal{T}}$.

A3 (Commutativity under consistency): *If* $m_{\mathcal{T}}$ *and* $m_{\mathcal{U}}$ *are consistent, i.e.,* $m_{\mathcal{T}}^{\downarrow \mathcal{T} \cap \mathcal{U}} = m_{\mathcal{U}}^{\downarrow \mathcal{T} \cap \mathcal{U}}$, *then* $m_{\mathcal{T}} \triangleright m_{\mathcal{U}} = m_{\mathcal{U}} \triangleright m_{\mathcal{T}}$.

A4 (Restricted associativity): *If* $\mathcal{T} \supset (\mathcal{U} \cap \mathcal{V})$, *or,* $\mathcal{U} \supset (\mathcal{T} \cap \mathcal{V})$, *then* $(m_{\mathcal{T}} \triangleright m_{\mathcal{U}}) \triangleright m_{\mathcal{V}} = m_{\mathcal{T}} \triangleright (m_{\mathcal{U}} \triangleright m_{\mathcal{V}})$.

Notice that axioms *A1*, *A3*, *A4* guarantee that the composition operator uniquely reconstructs BPA $m_{\mathcal{T} \cup \mathcal{V}}$ from its marginals, if there exists a *lossless* decomposition of $m_{\mathcal{T} \cup \mathcal{V}}$ into $m_{\mathcal{T}}$ and $m_{\mathcal{V}}$. Surprisingly, it is axiom *A4*, which guarantees that no necessary information from $m_{\mathcal{V}}$ is lost. Axiom *A2* solves the problem arising when non-consistent basic assignments are composed. Generally, there are two ways of coping with this problem. Either find a compromise (a mixture of inconsistent pieces of knowledge) or give preference to one of the sources. The solution expressed by axiom *A2* is superior to the other two from a computational point of view.

The following assertion (for proofs see (Jiroušek and Shenoy, 2014)) summarizes the main properties of composition operators. Based on these, efficient computational procedures were designed.

Proposition 1 *For arbitrary BPAs* $m_{\mathcal{T}}, m_{\mathcal{U}}, m_{\mathcal{V}}$ *the following statements hold.*

1. (Reduction:) *If* $\mathcal{U} \subseteq \mathcal{T}$, *then* $m_{\mathcal{T}} \triangleright m_{\mathcal{U}} = m_{\mathcal{T}}$.
2. (Stepwise composition): *If* $(\mathcal{T} \cap \mathcal{U}) \subseteq \mathcal{V} \subseteq \mathcal{U}$, *then* $(m_{\mathcal{T}} \triangleright m_{\mathcal{U}}^{\downarrow \mathcal{V}}) \triangleright m_{\mathcal{U}} = m_{\mathcal{T}} \triangleright m_{\mathcal{U}}$.
3. (Exchangeability): *If* $\mathcal{U} \supset (\mathcal{T} \cap \mathcal{V})$, *then* $(m_{\mathcal{T}} \triangleright m_{\mathcal{U}}) \triangleright m_{\mathcal{V}} = (m_{\mathcal{T}} \triangleright m_{\mathcal{V}}) \triangleright m_{\mathcal{U}}$.
4. (Simple marginalization): *If* $(\mathcal{T} \cap \mathcal{U}) \subseteq \mathcal{V} \subseteq (\mathcal{T} \cup \mathcal{U})$, *then* $(m_{\mathcal{T}} \triangleright m_{\mathcal{U}})^{\downarrow \mathcal{V}} = m_{\mathcal{T}}^{\downarrow \mathcal{T} \cap \mathcal{V}} \triangleright m_{\mathcal{U}}^{\downarrow \mathcal{U} \cap \mathcal{V}}$.

Before defining a composition operator for the D-S theory, notice that Dempster's combination rule is *not* a composition operator. Though it satisfies the first axiom (Domain), it does not satisfy the remaining three axioms. Whereas Dempster's rule is commutative and associative, a composition operator only satisfies these properties in special situations. On the other hand, Dempster's rule does not preserve the first marginal. Dempster's rule is designed to combine distinct pieces of evidence, whereas composition is designed to combine marginals that may not be independent. Nevertheless, as shown below, Dempster's rule can be used to define a composition operator.

3.1 d-composition

In this paper we follow the idea introduced in (Jiroušek and Shenoy, 2018). It defines d-composition of two BPAs $m_{\mathcal{U}}$, $m_{\mathcal{V}}$ (for any $\mathcal{U}, \mathcal{V}$) as follows.

$$(m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}}) = m_{\mathcal{U}} \oplus m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}},$$

where $\ominus$ denotes the inverse to Dempster's combination rule. $\ominus$ is defined using the corresponding commonality functions. Since it is known that the Dempster's rule can be stated as the product of the corresponding commonality functions (Shafer (1976)), i.e.,

$$Q_{m_1 \oplus m_2} = \frac{1}{1-K}\, Q_{m_1} \cdot Q_{m_2},$$

where K is the normalization factor from Eq. (3) defined by Eq. (4). Thus, $m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}$ was computed as a BPA corresponding to the following commonality function

$$Q_{m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}} = \frac{Q_{m_{\mathcal{V}}}}{Q_{m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}}}.$$

However, as shown in (Jiroušek and Shenoy, 2014), the composition operator $\triangleright_d$ sometimes yields BPAs with negative values (such BPAs are often called pseudo-BPAs).

Example 1 Consider $\Omega_X = \{x, \bar{x}\}$, $\Omega_Y = \{y, \bar{y}\}$, which means that $|2^{\Omega_X}| = 4$, and $|2^{\Omega_{XY}}| = 16$. In this example, consider a BPA m_{XY} for (X, Y) with only two focal elements – see Table 1. In tables, we depict only focal elements, i.e., if $\mathbf{a} \subseteq \Omega$ is not included in the table, then its value is 0.

Table 1: A Simple Example m_{XY}

$\mathbf{a}$	$m_{XY}(\mathbf{a})$
$\{(x,y)\}$	0.9
$\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}$	0.1

Its marginal $m_X = m_{XY}^{\downarrow X}$ has also two focal elements, namely $m_X(\{x\}) = 0.9$ and $m_X(\Omega_X) = 0.1$. Therefore, the corresponding commonality function is as follows: $Q_{m_X}(\{x\}) = 1$, $Q_{m_X}(\{\bar{x}\}) = Q_{m_X}(\Omega_X) = 0.1$. The computation of the corresponding $Q_{m_{XY} \ominus m_X}$ and $m_{XY} \ominus m_X$ can be seen in Table 2. $\square$

To avoid situations when the result of a composition is not a BPA, in this paper, we accept the possibility that the result of the operation of composition is undefined. Another advantage of this approach is that we also avoid the necessity of using commonality functions.

Definition 2 *Suppose $m_{\mathcal{U}}$ and $m_{\mathcal{V}}$ are BPAs. If $m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}$ is a BPA, then the d-composition is defined as follows:*

$$m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}} = m_{\mathcal{U}} \oplus (m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}). \tag{5}$$

If $m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U} \cap \mathcal{V}}$ is not a BPA, then $m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}}$ is undefined.

Table 2: Computation of $(m_{XY} \ominus m_X)(\mathbf{a})$.

$\mathbf{a}$	$Q_{m_{XY}}(\mathbf{a})$	$Q_{m_X}(\mathbf{a}^{\downarrow X})$	$Q_{m_{XY}\ominus m_X}(\mathbf{a}) = \frac{Q_{m_{XY}}(\mathbf{a})}{Q_{m_X}(\mathbf{a}^{\downarrow X})}$	$(m_{XY} \ominus m_X)(\mathbf{a})$
$\{(x,y)\}$	1	1	1	0.9
$\{(x,\bar{y})\}$	0.1	1	0.1	
$\{(\bar{x},\bar{y})\}$	0.1	0.1	1	
$\{(x,y),(x,\bar{y})\}$	0.1	1	0.1	-0.9
$\{(x,y),(\bar{x},\bar{y})\}$	0.1	0.1	1	
$\{(x,\bar{y}),(\bar{x},\bar{y})\}$	0.1	0.1	1	
$\Omega_{X,Y} \setminus \{(\bar{x},y)\}$	0.1	0.1	1	1

Remark 1 A disadvantage of this definition follows from the fact that neither the axioms of Definition 1, nor Properties expressed in Proposition 1 generally hold exactly as they are expressed. Namely, one has to add that they hold under the assumption that the corresponding compositions are defined. As an example, consider the Stepwise composition (Property 2 from Proposition 1) with $\mathcal{T} = \emptyset$: If $\mathcal{U} \subseteq \mathcal{V}$, then $m_{\mathcal{V}}^{\mathcal{U}} \triangleright m_{\mathcal{V}} = m_{\mathcal{V}}$. Naturally, this equality can hold only when $m_{\mathcal{V}}^{\mathcal{U}} \triangleright m_{\mathcal{V}}$ is defined.

3.2 f-composition

The f-composition operator is defined as follows:

Definition 3 *Consider two BPAs $m_{\mathcal{U}}$ and $m_{\mathcal{V}}$. Their f-composition is a BPA $m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}}$ defined for each nonempty $\mathbf{c} \subseteq \Omega_{\mathcal{U}\cup\mathcal{V}}$ by one of the following expressions:*

(i) *If $m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{U}\cap\mathcal{V}}) > 0$ and $\mathbf{c} = \mathbf{c}^{\downarrow\mathcal{U}} \bowtie \mathbf{c}^{\downarrow\mathcal{V}}$, then* $(m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}})(\mathbf{c}) = \dfrac{m_{\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{U}}) \cdot m_{\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{V}})}{m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{U}\cap\mathcal{V}})}$;

(ii) *If $m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{U}\cap\mathcal{V}}) = 0$ and $\mathbf{c} = \mathbf{c}^{\downarrow\mathcal{U}} \times \Omega_{\mathcal{V}\setminus\mathcal{U}}$, then $(m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}})(\mathbf{c}) = m_{\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{U}})$;*

(iii) *In all other cases, $(m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}})(\mathbf{c}) = 0$.*

Remark 2 f-composition is always defined. Notice that if $m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{U}\cap\mathcal{V}}) = 0$ (i.e., the formula in case (**i**) is undefined), then the definition accepts a heuristic solution saying "I do not know".

4 Properties of Composition Operators

First, we prove the following simple assertion characterizing $m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}}$. A similar result is stated by Shenoy (1994) in the context of valuation-based systems.

Proposition 2 *Consider nonempty sets of variables $\mathcal{U} \subsetneq \mathcal{V}$ and BPA $m_{\mathcal{V}}$.*
If $(m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow\mathcal{U}\cap\mathcal{V}})$ is a BPA, then the following two properties hold:

- $m_{\mathcal{V}} = m_{\mathcal{V}}^{\downarrow \mathcal{U}} \oplus (m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U}\cap\mathcal{V}})$;
- $\left(m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U}\cap\mathcal{V}}\right)^{\downarrow \mathcal{U}}$ *is vacuous.*

Proof. The first property is a direct implication of the associativity and commutativity of the Dempster's rule of combination, and the latter one follows immediately from the property

$$\mathcal{W} \supseteq \mathcal{T} \supseteq \mathcal{W} \cap \mathcal{V} \implies (m_{\mathcal{V}} \oplus m_{\mathcal{W}})^{\downarrow \mathcal{T}} = m_{\mathcal{V}} \oplus m_{\mathcal{W}}^{\downarrow \mathcal{T}}$$

called "local computation" (Shenoy and Shafer, 1990). □

These two properties are often expected to hold for the conditional BPA $m_{\mathcal{V}\setminus\mathcal{U}|\mathcal{U}}$. Recall that the conditional BPA was defined by Smets (1978) and Shafer (1982) using so-called conditional embedding. We do not need this notion in this paper, and so we do not present the definition. Nevertheless, it may be an interesting question for the future study to find out under what conditions $m_{\mathcal{V}\setminus\mathcal{U}|\mathcal{U}} = m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U}\cap\mathcal{V}}$.

In Example 1 we presented a simple BPA m_{XY} for which $m_{XY} \ominus m_X$ was not a BPA. It means that there are BPAs that cannot be a second argument of a d-composition. From this, however, one cannot exclude the existence of another BPA for which the properties from Proposition 2 hold. Thus, let us turn back to the above-presented example and show that for m_{XY} from Table 1 such a BPA does not exist.

Example 1 (*Continued.*) Let us assume that there exists two-dimensional BPA $m_{Y|X}$ such that for m_{XY} from Table 1 $m_{XY} = m_{XY}^{\downarrow X} \oplus m_{Y|X}$. Then, under this assumption, for all $\mathbf{a} \subseteq \Omega_{XY}$

$$m_{XY}(\mathbf{a}) = (1-K)^{-1} \sum_{\mathbf{b}\subseteq\Omega_X \& \mathbf{c}\subseteq\Omega_{XY}:\mathbf{b}\bowtie\mathbf{c}=\mathbf{a}} m_{XY}^{\downarrow X}(\mathbf{b}) \cdot m_{Y|X}(\mathbf{c}). \tag{6}$$

Since $\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\} = \{x,\bar{x}\} \bowtie \{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}$, and for no other $\mathbf{b},\mathbf{c}$ their join $\mathbf{b} \bowtie \mathbf{c} = \{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}$, it is clear that $m_{Y|X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}) = (1-K)$ because

$$\begin{aligned}
0.1 &= m_{XY}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}) \\
&= (1-K)^{-1}\ m_{XY}^{\downarrow X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}^{\downarrow X}) \cdot m_{Y|X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}) \\
&= (1-K)^{-1}\ \cdot m_{XY}^{\downarrow X}(\{x,\bar{x}\}) \cdot m_{Y|X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}) \\
&= (1-K)^{-1}\ \cdot 0.1 \cdot m_{Y|X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}).
\end{aligned}$$

Since $\{x\} \bowtie \{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\} = \{(x,y),(x,\bar{y})\}$, it immediately follows from (6) that

$$\begin{aligned}
m_{XY}(\{(x,y),(x,\bar{y})\}) &\geq (1-K)^{-1} m_{XY}^{\downarrow X}(\{x\}) \cdot m_{Y|X}(\{(x,y),(x,\bar{y}),(\bar{x},\bar{y})\}) \\
&= (1-K)^{-1} \cdot 0.9 \cdot (1-K) = 0.9,
\end{aligned}$$

which is in the contradiction with the assumption, because $m_{XY}(\{(x,y),(x,\bar{y})\}) = 0$. □

To simplify the notation, and to make it a bit more lucid, let us denote in the rest of this section $m_{\mathcal{V}|\mathcal{U}} = m_{\mathcal{V}} \ominus m_{\mathcal{V}}^{\downarrow \mathcal{U}\cap\mathcal{V}}$. Moreover, in connection with Definition 2, we will identify situations when BPA $m_{\mathcal{V}|\mathcal{U}\cap\mathcal{V}}$ exists and is, in a way, "adapted" to BPA $m_{\mathcal{U}}$. We will say that $m_{\mathcal{V}|\mathcal{U}\cap\mathcal{V}}$ is *tight* with respect to $m_{\mathcal{U}}$ if for all couples of focal elements $\mathbf{a}$ and $\mathbf{b}$ ($\mathbf{a}$ is a focal element of $m_{\mathcal{U}}$, and $\mathbf{b}$ is a focal element of $m_{\mathcal{V}|\mathcal{U}\cap\mathcal{V}}$) the following condition holds:

$$\text{for } \forall\, b \in \mathbf{b},\ \exists\, a \in \mathbf{a},\ \text{such that } \{a\} \bowtie \{b\} \neq \emptyset. \tag{7}$$

Proposition 3 *Let two BPAs $m_{\mathcal{U}}, m_{\mathcal{V}}$ are such that $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ exists. If $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ is tight with respect to $m_{\mathcal{U}}$, then*

$$m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}} = m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}}.$$

Proof. Recall that for BPA $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$, the existence of which is assumed,

$$m_{\mathcal{V}} = m_{\mathcal{V}}^{\downarrow \mathcal{V}\cap\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}, \tag{8}$$

and that the d-composition is defined

$$m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}} = m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}.$$

What are the focal elements of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$? Let $\mathbf{a}$ and $\mathbf{b}$ be arbitrary focal elements of $m_{\mathcal{U}}$ and $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$, respectively. Due to Proposition 2, $\left(m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}\right)^{\downarrow \mathcal{V}\cap\mathcal{U}}$ is vacuous, $\mathbf{b}^{\downarrow \mathcal{V}\cap\mathcal{U}} = \Omega_{\mathcal{V}\cap\mathcal{U}}$, and $\mathbf{c} = \mathbf{a} \bowtie \mathbf{b} \neq \emptyset$ is a focal element of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$. Therefore, when computing the Dempster's rule of combination $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$, the corresponding coefficient of conflict (see Eq. 4)

$$K = \sum_{\mathbf{a}\subseteq\Omega_{\mathcal{U}},\mathbf{b}\subseteq\Omega_{\mathcal{V}}:\mathbf{a}\bowtie\mathbf{b}=\emptyset} m_{\mathcal{U}}(\mathbf{a}) \cdot m_{\mathcal{V}\setminus\mathcal{U}|\mathcal{V}\cap\mathcal{U}}(\mathbf{b}) = 0. \tag{9}$$

The question is whether for a focal element $\mathbf{c}$ of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ it may happen that $\mathbf{c} = \mathbf{a} \bowtie \mathbf{b}$, and either $\mathbf{a} \neq \mathbf{c}^{\downarrow\mathcal{U}}$, or $\mathbf{b} \neq \mathbf{c}^{\downarrow\mathcal{V}}$. Since $\mathbf{b}^{\downarrow \mathcal{V}\cap\mathcal{U}} = \Omega_{\mathcal{V}\cap\mathcal{U}}$, for $\forall\, a \in \mathbf{a}$, $\exists\, b \in \mathbf{b}$, $\{a\} \bowtie \{b\}$ is a singleton from $\mathbf{c}^{\downarrow\mathcal{U}} \bowtie \mathbf{c}^{\downarrow\mathcal{V}}$ and therefore $\mathbf{a} \subseteq \mathbf{c}^{\downarrow\mathcal{U}}$. Similarly, the assumption that $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ is tight with respect to $m_{\mathcal{U}}$ guarantees that $\mathbf{b} \subseteq \mathbf{c}^{\downarrow\mathcal{V}}$. For all $c \in \mathbf{a} \bowtie \mathbf{b}$, $c^{\downarrow\mathcal{U}} \in \mathbf{a}$ from the definition of a join, and therefore $\mathbf{a} \supseteq \mathbf{c}^{\downarrow\mathcal{U}}$. Analogously, $c^{\downarrow\mathcal{V}} \in \mathbf{b}$ yields $\mathbf{b} \supseteq \mathbf{c}^{\downarrow\mathcal{V}}$. So, we have proven that each focal element $\mathbf{c}$ of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ is created by a single pair of focal elements $\mathbf{c}^{\downarrow\mathcal{U}}$ of $m_{\mathcal{U}}$ and $\mathbf{c}^{\downarrow\mathcal{V}}$ of $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$. Therefore (using definition from Eq. (3) and Eq. (9)),

$$\begin{aligned}(m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}})(\mathbf{c}) \\ = \sum_{\mathbf{a}\subseteq\Omega_{\mathcal{U}},\mathbf{b}\subseteq\Omega_{\mathcal{V}}:\mathbf{a}\bowtie\mathbf{b}=\mathbf{c}} m_{\mathcal{U}}(\mathbf{a}) \cdot m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}(\mathbf{b}) = m_{\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{U}}) \cdot m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{V}}).\end{aligned} \tag{10}$$

In the same, way we get from Eq. (8) also

$$m_{\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{V}}) = (m_{\mathcal{V}}^{\downarrow \mathcal{V}\cap\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}})(\mathbf{c}^{\downarrow\mathcal{V}}) = m_{\mathcal{V}}^{\downarrow \mathcal{V}\cap\mathcal{U}}(\mathbf{c}^{\downarrow \mathcal{V}\cap\mathcal{U}}) \cdot m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{V}}), \tag{11}$$

which gives that, under the given assumptions,

$$m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{V}}) = \frac{m_{\mathcal{V}}(\mathbf{c}^{\downarrow\mathcal{V}})}{m_{\mathcal{V}}^{\downarrow\mathcal{V}\cap\mathcal{U}}(\mathbf{c}^{\downarrow\mathcal{V}\cap\mathcal{U}})}. \tag{12}$$

Substituting Eq. (12) into Eq. (10), we get exactly the formula from case **(i)** of Definition 3. The fact that case **(ii)** of this definition never creates a focal element of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ follows from the fact that each couple of focal elements $\mathbf{a}$ and $\mathbf{b}$ ($\mathbf{a}$ is a focal element of $m_{\mathcal{U}}$, and $\mathbf{b}$ is a focal element of $m_{\mathcal{V}|\mathcal{U}\cap\mathcal{V}}$) gives rise of a focal element $\mathbf{a} \bowtie \mathbf{b}$ of $m_{\mathcal{U}} \oplus m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$. Thus, whenever case **(ii)** of Definition 3 is used (under the assumptions of this assertion), then it assigns zero. $\square$

Corollary Let two BPAs $m_{\mathcal{U}}, m_{\mathcal{V}}$ are such that $m_{\mathcal{V}|\mathcal{V}\cap\mathcal{U}}$ exists. If $m_{\mathcal{V}}^{\downarrow\mathcal{V}\cap\mathcal{U}}$ is vacuous, or, if $\mathcal{V}\cap\mathcal{U} = \emptyset$, then

$$m_{\mathcal{U}} \triangleright_f m_{\mathcal{V}} = m_{\mathcal{U}} \triangleright_d m_{\mathcal{V}}.$$

Example 2 In this example we show that, generally, d-composition and f-composition of two BPAs may differ from each other. Consider three binary variables X, Y, Z, and m_{XY} and $m_{Z|Y}$ from Table 3.

Table 3: Example when $m_{Z|Y}$ is not tight with respect to m_{XY}.

$\mathbf{a}$	$m_{XY}(\mathbf{a})$
$\{(x,y)\}$	1.00

$\mathbf{a}$	$m_{Z|Y}(\mathbf{a})$
$\{(\bar{y},\bar{z}),(y,z)\}$	1.00

$\mathbf{a}$	$(m_{XY} \triangleright_d m_{Z|Y})(\mathbf{a})$
$\{(x,y,z)\}$	1.00

$\mathbf{a}$	$(m_{XY} \triangleright_f m_{Z|Y})(\mathbf{a})$
$\{(x,y,\bar{z}),(x,y,z)\}$	1.00

Notice that in this example, $m_{Z|Y}$ is not tight with respect to m_{XY} because for $(\bar{y},\bar{z}) \in \{(\bar{y},\bar{z}),(y,z)\}$ there is no element $a \in \{(x,y)\}$ such that $a \bowtie (\bar{y},\bar{z}) \neq \emptyset$.

5 Almond's Captain's Problem

Let us briefly replicate the Captain's problem from the book by Almond (1995). As said in Section 1, this example motivated this research. Namely, when being converted into the form of a compositional model, it defined the same eight-dimensional BPA regardless of the used composition operator.

For the detailed story, we refer the reader to the original book (Almond, 1995), or the paper (Jiroušek et al., 2022) published in this proceedings. The problem concerns the relation of eight variables presented in Table 4. Their mutual relations are in this paper described in a slightly different way than in the cited book. Here we use three

Table 4: Variables for the Captain's decision.

Variable	# states	States	Description
L	2	true, false	Loading is delayed?
F	2	true, false	Weather forecast is foul?
W	2	true, false	Weather in route is foul?
M	2	true, false	Maintenance is done?
R	2	true, false	Ship needs repairs at sea?
D	4	0, 1, 2, 3	Departure delay (in days)
S	4	0, 1, 2, 3	Sailing delay (in days)
A	7	0, 1, 2, 3, 4, 5, 6	Arrival delay (in days)

prior (one-dimensional) belief functions and five low-dimensional (conditional) BPAs (see Table 5). The resulting eight-dimensional BPA is, for example, given by the formula

$$m_{\{L\}} \triangleright m_{\{F\}} \triangleright m_{\{M\}} \triangleright m_{\{D,F,L,M\}} \triangleright m_{\{F,W\}} \triangleright m_{\{M,R\}} \triangleright m_{\{S,W,R\}} \triangleright m_{\{A,D,S\}}. \tag{13}$$

The ordering of low-dimensional BPAs in Eq. (13) is compatible with the directed graphical model that underlies the Captain's problem in the sense that the conditional for a variable should be composed only after the conditional associated with its parents. Formally, and not using graphs, this property can be formulated that for any compositional model $m_{\mathcal{U}_1} \triangleright m_{\mathcal{U}_2} \triangleright \ldots \triangleright m_{\mathcal{U}_k}$ corresponding to a directed graphical model, for all $j = 2, \ldots, k$, set $\mathcal{U}_j \setminus (\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_{j-1})$ must be singleton, i.e., $|\mathcal{U}_j \setminus (\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_{j-1})| = 1$. Thus, there are other sequences in which the low-dimensional BPAs may be composed without influencing the resulting eight-dimensional one. All of them may be got from Formula (13) by the application of axioms $A3$, $A4$ from Definition 1, and Property 3 from Proposition 1. Another equivalent one is, e.g.,

$$m_{\{F\}} \triangleright m_{\{F,W\}} \triangleright m_{\{M\}} \triangleright m_{\{M,R\}} \triangleright m_{\{S,W,R\}} \triangleright m_{\{L\}} \triangleright m_{\{D,F,L,M\}} \triangleright m_{\{A,D,S\}}. \tag{14}$$

Not presenting the lists of focal elements of the eight BPAs from Table 5, we cannot show it, but the reader can certainly imagine that verification of the fact that $m_{\{F,W\}}^{\downarrow\{F\}}$ is vacuous (and therefore $m_{\{F,W\}} = m_{\{F,W\}} \ominus m_{\{F,W\}}^{\downarrow\{F\}}$), and that $m_{\{F,W\}}$ is tight with respect to $m_{\{F\}}$ is simple. It is enough to check 2×2 couples of focal elements to verify the latter condition. Thus, it is easy to verify the assumption of Proposition 3, and to show that $m_{\{F\}} \triangleright_f m_{\{F,W\}} = m_{\{F\}} \triangleright_d m_{\{F,W\}}$. The fact that the analogous equality holds for the first three terms of the formula (14) follows directly from Corollary. In a similar way, it is not difficult to show that the eight-dimensional BPA defined by formula (14) does not depend on which operator of composition is used. Nevertheless, one has to realize that it is necessary to show that $m_{\{M,R\}}$ is tight with respect to $m_{\{F\}} \triangleright m_{\{F,W\}} \triangleright m_{\{M\}}$, where the latter BPA (defined as a composition of three low-dimensional BPAs) has 6 focal elements. As a rule, the longer the compositional model, the more focal elements the corresponding BPA has. Thus, Proposition 3 applies to small compositional models, but when one starts considering multidimensional models composed of hundreds of low-dimensional BPAs, its direct application is unrealistic.

Table 5: Low-dimensional BPAs for the Captain's Problem.

Variables	# focal elements	Description
L	3	prior BPA
F	3	prior BPA
M	1	prior BPA: did not perform maintenance before departure
A, D, S	1	rule calculating total delay: A = D + S
D, F, L, M	1	logical function: departure will be delayed one day for each thing wrong
R, S, W	2	noisy logical statement: sailing time increases by one day if something gets wrong
F, W	2	reliability of weather forecast
M, R	9	relationship between maintenance and repairs at sea

6 Summary & Conclusions

The main result of this paper is presented as Proposition 3. It says that, in some situations, the two composition operators yield the same result. It may be interesting because d-composition is generally of much higher computational complexity than f-composition. Nevertheless, Proposition 3 presents only sufficient conditions, not necessary ones. The determination of necessary conditions remains an open problem.

From the exposition, the reader could notice that the analogy between probabilistic and belief-function graphical models is far from being straightforwards. One can always represent any multivariate probability distribution as a directed graphical model (but the directed graphical model may not encode all the conditional independencies in the joint distribution). As shown in Example 1, it is not true for belief functions because there are joint BPAs for which some conditionals do not exist. Similarly, see Remark 1 stepwise composition need not always hold for BPAs. On the other hand, like the d-composition operator, the probabilistic composition operator is not always defined. Surprisingly, f-composition is always defined. It is made possible by the heuristics expressed by case (ii) of Definition 3.

Acknowledgement

This study was financially supported by the Czech Science Foundation under Grant No. 19-06569S, and by the Ronald G. Harper Professorship at the University of Kansas awarded to the third author.

References

R. G. Almond. *Graphical Belief Modeling.* Chapman & Hall, London, UK, 1995.

R. Jiroušek and P. P. Shenoy. Compositional models in valuation-based systems. *International Journal of Approximate Reasoning*, 55(1):277–293, 2014.

R. Jiroušek and P. P. Shenoy. A new definition of entropy of belief functions in the Dempster-Shafer theory. *International Journal of Approximate Reasoning*, 92(1):49–65, 2018.

R. Jiroušek, J. Vejnarová, and M. Daniel. Compositional models for belief functions. In G. de Cooman, J. Vejnarová, and M. Zaffalon, editors, *Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '07)*, pages 243–252, 2007.

R. Jiroušek, V. Kratochvíl, and P. P. Shenoy. Computing the decomposable entropy of graphical belief function models. In *Proc. 12. Workshop on Uncertainty Processing (WUPES 2022)*. MatfyzPress, Praha, Czech Republic, 2022.

G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

G. Shafer. Belief functions and parametric models. *Journal of the Royal Statistical Society, Series B*, 44(3):322–352, 1982.

P. P. Shenoy. Conditional independence in valuation-based systems. *International Journal of Approximate Reasoning*, 10(3):203–234, 1994.

P. P. Shenoy and G. Shafer. Axioms for probability and belief-function propagation. In R. D. Shachter, T. Levitt, J. F. Lemmer, and L. N. Kanal, editors, *Uncertainty in Artificial Intelligence 4*, Machine Intelligence and Pattern Recognition Series, vol. 9, pages 169–198. North-Holland, Amsterdam, 1990.

P. Smets. *Un modele mathematico-statistique simulant le processus du diagnostic medical.* PhD thesis, Free University of Brussels, 1978.