

# Conditioning in Decomposable Compositional Models in Valuation-Based Systems

Radim Jiroušek<sup>1</sup> and Prakash P. Shenoy<sup>2</sup>

<sup>1</sup> Faculty of Management at Univ. of Economics, Jindřichův Hradec, Czech Republic  
radim@utia.cas.cz

<sup>2</sup> School of Business at University of Kansas  
Lawrence, KS, USA  
pshenoy@ku.edu

**Abstract.** Valuation-based systems (VBS) can be considered as a generic uncertainty framework that has many uncertainty calculi, such as probability theory, a version of possibility theory where combination is the product  $t$ -norm, Spohn's epistemic belief theory, and Dempster-Shafer belief function theory, as special cases. In this paper, we focus our attention on conditioning, which is defined using the combination, marginalization, and removal operators of VBS. We show that conditioning can be expressed using the composition operator. We define decomposable compositional models in the VBS framework. Finally, we show that conditioning in decomposable compositional models can be done using local computation. Since all results are obtained in the VBS framework, they hold in all calculi that fit in the VBS framework.

**Keywords:** valuation-based systems, probability theory, possibility theory, Dempster-Shafer belief function theory, Spohn's epistemic belief theory, conditionals, compositional models, decomposable compositional models, conditioning in decomposable compositional models.

## 1 Introduction to Valuation-Based Systems

Valuation-based systems (VBS) were introduced in [9] as a generic uncertainty calculus that has many uncertainty calculi, such as probability theory, a version of possibility theory [2] with the product  $t$ -norm, Spohn's epistemic belief theory [11], and D-S belief function theory [1,8], as special cases. In this section, we formally introduce the VBS framework. Most of the material in this section is taken from [9].

VBS consists of two parts — a static part that is concerned with representation of knowledge, and a dynamic part that is concerned with reasoning with knowledge.

The static part consists of objects called variables and valuations. Let  $\Phi$  denote a finite set whose elements are called *variables*. Elements of  $\Phi$  are denoted by upper-case Roman alphabets such as  $X, Y, Z$ , etc. Subsets of  $\Phi$  are denoted by lower-case Roman alphabets such as  $r, s, t$ , etc.

Let  $\Psi$  denote a set whose elements are called *valuations*. Elements of  $\Psi$  are denoted by lower-case Greek alphabets such as  $\rho, \sigma, \tau$ , etc. Each valuation is associated with a subset of variables, and represents some knowledge about the variables in the subset. Thus, we will say that  $\rho$  is a valuation for  $r$ , where  $r \subseteq \Phi$ .

We can depict VBS graphically by graphs called valuation networks. Consider a finite set of valuations  $A \subset \Psi$  defining a valuation-based system. A corresponding *valuation network* (VN) is a bi-partite graph with variables and valuations as nodes, and there is an edge between each valuation and the variables in the subset associated with it. An example is shown in Figure 1. In this example,  $\Phi = \{D, G, B\}$ ,  $A = \{\delta, \gamma, \beta\}$ , where  $\delta$  is a valuation for  $\{D\}$ ,  $\gamma$  is a valuation for  $\{D, G\}$ , and  $\beta$  is a valuation for  $\{D, B\}$ .

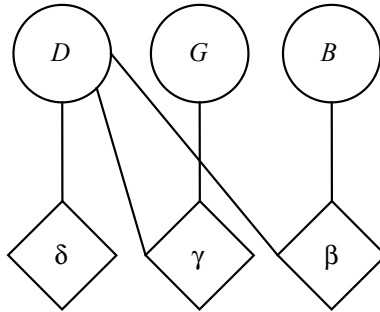


Fig. 1. A valuation network

We will identify a subset of valuations  $\Psi_n \subset \Psi$ , whose elements are called *normal* valuations. Normal valuations are valuations that are coherent in some sense. In probability theory, normal valuations are probability potentials whose values add to one. In D-S belief function theory, normal valuations are basic probability assignment potentials whose values for non-empty subsets add to one (or their corresponding commonality potentials).

The dynamic part of VBS consists of several operators that are used to make inferences from the knowledge encoded in a VBS. We will define three basic operators: combination, marginalization, and removal, and their properties.

**Combination.** The first operator is the *combination* operator  $\oplus: \Psi \times \Psi \rightarrow \Psi_n$ , which represents aggregation of knowledge. It has the following properties.

1. (*Domain*) If  $\rho$  is a valuation for  $r$ , and  $\sigma$  is a valuation for  $s$ , then  $\rho \oplus \sigma$  is a normal valuation for  $r \cup s$ .
2. (*Commutativity*)  $\rho \oplus \sigma = \sigma \oplus \rho$ .
3. (*Associativity*)  $\rho \oplus (\sigma \oplus \tau) = (\rho \oplus \sigma) \oplus \tau$ .

The domain property expresses the fact that if  $\rho$  represents some knowledge about variables in  $r$ , and  $\sigma$  represents some knowledge about variables in  $s$ ,

then  $\rho \oplus \sigma$  represents the aggregated knowledge about variables in  $r \cup s$ . The commutativity and associativity properties reflect the fact that the sequence in which knowledge is aggregated makes no difference in the aggregated result. In probability theory, combination of two valuations is pointwise multiplication followed by normalization assuming normalization is possible. If normalization is not possible, this means that the knowledge encoded in  $\rho$  and  $\sigma$  are completely inconsistent. Henceforth, for the sake of simplicity, we will assume that we don't have inconsistent valuations.

The set of all normal valuations with the combination operator  $\oplus$  forms a commutative semigroup. We will let  $\iota_\emptyset$  denote the (unique) identity valuation of this semigroup. Thus, for any normal valuation  $\rho$ ,  $\rho \oplus \iota_\emptyset = \rho$ . It is easy to see that the domain of  $\iota_\emptyset$  is  $\emptyset$ , hence the notation.

The set of all normal valuations for  $s$  with the combination operator  $\oplus$  also forms a commutative semigroup (which is different from the semigroup discussed in the previous paragraph). Let  $\iota_s$  denote the (unique) identity for this semigroup. Thus, for any normal valuation  $\sigma$  for  $s$ ,  $\sigma \oplus \iota_s = \sigma$ .

It is important to note that in most uncertainty calculi, in general,  $\rho \oplus \rho \neq \rho$ . Thus, it is important to ensure that we do not double count knowledge when double counting matters, i.e., it is okay to double count knowledge  $\rho$  that is idempotent, i.e.,  $\rho \oplus \rho = \rho$ . In representing our knowledge as valuations from  $\Psi$ , we have to ensure that there is no double counting of non-idempotent knowledge.

**Marginalization.** Another operator is marginalization  $-X: \Psi \rightarrow \Psi$ , which allows us to coarsen knowledge by marginalizing  $X$  out of the domain of a valuation. It has the following properties.

1. (*Domain*) If  $\rho$  is a valuation for  $r$ , and  $X \in r$ , then  $\rho^{-X}$  is a valuation for  $r \setminus \{X\}$ .
2. (*Normal*)  $\rho^{-X}$  is normal if and only if  $\rho$  is normal.
3. (*Order does not matter*) If  $\rho$  is a valuation for  $r$ ,  $X \in r$ , and  $Y \in r$ , then  $(\rho^{-X})^{-Y} = (\rho^{-Y})^{-X}$ , which we will denote by  $\rho^{-\{X,Y\}}$ .
4. (*Local computation*) If  $\rho$  and  $\sigma$  are valuations for  $r$  and  $s$ , respectively,  $X \in r$ , and  $X \notin s$ , then  $(\rho \oplus \sigma)^{-X} = (\rho^{-X}) \oplus \sigma$ .

The domain property is self-explanatory. Marginalization preserves normal (and non-normal) property of valuations. The order does not matter property dictates that when we coarsen knowledge by marginalizing out several variables, the order in which the variables are marginalized does not matter in the final result. Thus, if  $\rho$  is a normal valuation for  $r$ , then  $\rho^{-r} = \iota_\emptyset$ , which is the only normal valuation for  $\emptyset$ . Sometimes, we will let  $\rho^{\downarrow r \setminus \{X,Y\}}$  denote  $\rho^{-\{X,Y\}}$ . Thus, the “ $-$ ” notation is useful when we wish to emphasize the variables being marginalized, whereas the “ $\downarrow$ ” notation is useful when we wish to emphasize the variables that remain after the marginalization operation. In probability theory, marginalization of variable  $X$  corresponds to addition over the state space of  $X$ .

Making inferences in VBS means finding the (posterior) marginal of the joint valuation for some variables of interest, i.e., computing  $(\oplus \Lambda)^{\downarrow \{Z\}}$ , where  $\Lambda$  includes valuations that represent observations and independent pieces of evidence,

and  $Z$  denotes an unobserved variable of interest. When we have many variables in  $\Phi$ , it may be computationally intractable to compute explicitly the joint valuation  $(\oplus \Lambda)$ . The local computation property allows us to compute the marginal  $(\rho \oplus \sigma)^{-X}$  without having to explicitly compute  $\rho \oplus \sigma$ . Notice that the combination in  $(\rho^{-X}) \oplus \sigma$  is on a smaller set of variables  $(r \cup s \setminus \{X\})$  compared to  $\rho \oplus \sigma$  (which is on  $r \cup s$ ). By repeatedly using this property for all variables being marginalized in some sequence, we get the so-called variable elimination algorithm for computing marginals. Also, if we wish to compute several marginals, then it is useful to cache the intermediate results in the computation of one marginal so that these can be re-used in the computation of other marginals. A binary join tree is a data structure that is useful for this purpose. For details, see [10].

While the combination and marginalization operators suffice for the problem of making inferences, there is yet another operator, called removal, that is useful for defining conditionals, and for defining the composition operator.

**Removal.** The removal operator  $\ominus: \Psi \times \Psi_n \rightarrow \Psi_n$  represents removing knowledge in the second valuation from the knowledge in the first valuation. The properties of the removal operator are as follows.

1. (*Domain*): Suppose  $\sigma$  is a valuation for  $s$  and  $\rho$  is a normal valuation for  $r$ . Then  $\sigma \ominus \rho$  is a normal valuation for  $r \cup s$ .
2. (*Identity*): For each normal valuation  $\rho$  for  $r$ ,  $\rho \oplus \rho \ominus \rho = \rho$ . Thus,  $\rho \ominus \rho$  acts as an identity for  $\rho$ , and we denote  $\rho \ominus \rho$  by  $\iota_\rho$ . Thus,  $\rho \oplus \iota_\rho = \rho$ .
3. (*Combination and Removal*): Suppose  $\pi$  and  $\theta$  are valuations, and suppose  $\rho$  is a normal valuation. Then,  $(\pi \oplus \theta) \ominus \rho = \pi \oplus (\theta \ominus \rho)$ .

We call  $\sigma \ominus \rho$  the valuation resulting after removing  $\rho$  from  $\sigma$ . Notice that the removal operator cannot be extended as an operator  $\ominus: \Psi \times \Psi \rightarrow \Psi_n$  because of the identity property, which defines the removal operator as an inverse of the combination operator. In probability theory removal is pointwise division followed by normalization (here, division of any real number by zero results in zero, by definition).

It is important to note that given a normal valuation  $\rho$  for  $r$ , we have a number of (may be different) identity valuations  $\iota$  such that  $\rho \oplus \iota = \rho$ . So far we have explicitly mentioned  $\iota_\emptyset$ ,  $\iota_r$  and  $\iota_\rho$ . However, and it will be shown in Lemma 1, for any  $s \subseteq r$ ,  $\iota_{\rho \uparrow s}$  acts also as an identity valuation for  $\rho$ , i.e.,

$$\rho \oplus \iota_{\rho \uparrow s} = \rho.$$

We reproduce some important results about the removal operator from [9].

**Proposition 1.** *Suppose  $\pi, \sigma, \theta$  are valuations, and  $\rho$  is a normal valuation for  $r$  and  $X \notin r$ . Then,*

$$(\pi \oplus \theta) \ominus \rho = (\pi \ominus \rho) \oplus \theta, \tag{1}$$

and

$$(\sigma \ominus \rho)^{-X} = \sigma^{-X} \ominus \rho. \tag{2}$$

**Domination.** As defined in the identity property,  $\rho \oplus \iota_\rho = \rho$ . In general, if  $\rho'$  is a normal valuation for  $r$  that is distinct from  $\rho$ , then  $\rho' \oplus \iota_\rho$  may not equal  $\rho'$ . However, there may exist a class of normal valuations for  $r$  such that if  $\rho'$  is in this class, then  $\rho' \oplus \iota_\rho = \rho'$ . Following the terminology in [5], we will call this class of normal valuations as valuations that are *dominated* by  $\rho$ . Thus, if  $\rho$  dominates  $\rho'$ , written as  $\rho \gg \rho'$ , then  $\rho' \oplus \iota_\rho = \rho'$ . In probability theory, if  $\rho$  and  $\rho'$  are normal probability potentials for  $r$  such that  $\rho(\mathbf{x}) = 0 \Rightarrow \rho'(\mathbf{x}) = 0$ , then  $\rho \gg \rho'$ .

**Composition.** A general definition of the composition operator is as follows. Suppose  $\rho$  and  $\sigma$  are normal valuations for  $r$  and  $s$ , respectively, and, to avoid composition of conflicting valuations, suppose that  $\sigma \downarrow^{r \cap s} \gg \rho \downarrow^{r \cap s}$ . The composition of  $\rho$  and  $\sigma$ , written as  $\rho \triangleright \sigma$ , is defined as follows:

$$\rho \triangleright \sigma = \rho \oplus \sigma \ominus \sigma \downarrow^{r \cap s} \tag{3}$$

Unlike the combination operator, the valuations  $\rho$  and  $\sigma$  being composed do not have to be distinct. Intuitively, we adjust for the double-counting of the knowledge in  $\rho$  and  $\sigma$  by removing the knowledge that is double counted.

The most important properties of the composition operator that were proved in [6] are summarized in the following proposition.

**Proposition 2.** *Suppose  $\rho$  and  $\sigma$  are normal valuations for  $r$  and  $s$ , respectively, and suppose that  $\sigma \downarrow^{r \cap s} \gg \rho \downarrow^{r \cap s}$ . Then the following statement hold.*

1. Domain:  $\rho \triangleright \sigma$  is a normal valuation for  $r \cup s$ .
2. Composition preserves first marginal:  $(\rho \triangleright \sigma) \downarrow^r = \rho$ .
3. Non-commutativity: In general,  $\rho \triangleright \sigma \neq \sigma \triangleright \rho$ .
4. Commutativity under consistency: If  $\rho$  and  $\sigma$  have a common marginal for  $r \cap s$ , i.e.,  $\rho \downarrow^{r \cap s} = \sigma \downarrow^{r \cap s}$ , then  $\rho \triangleright \sigma = \sigma \triangleright \rho$ .
5. Non-associativity: Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $\tau \downarrow^{(r \cup s) \cap t} \gg (\rho \triangleright \sigma) \downarrow^{(r \cup s) \cap t}$ . Then, in general,

$$(\rho \triangleright \sigma) \triangleright \tau \neq \rho \triangleright (\sigma \triangleright \tau).$$

6. Associativity under a special condition: Suppose  $\tau$  is a normal valuation for  $t$ , suppose  $\tau \downarrow^{(r \cup s) \cap t} \gg (\rho \triangleright \sigma) \downarrow^{(r \cup s) \cap t}$ , and suppose  $s \supset (r \cap t)$ . Then,

$$(\rho \triangleright \sigma) \triangleright \tau = \rho \triangleright (\sigma \triangleright \tau).$$

7. Composition of marginals: Suppose  $t$  is such that  $(r \cap s) \subseteq t \subseteq s$ . Then

$$(\rho \triangleright \sigma \downarrow^t) \triangleright \sigma = \rho \triangleright \sigma.$$

## 2 Conditionals

Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r$  and  $s$  are disjoint subsets of  $t$ . We call  $\tau \downarrow^{(r \cup s)} \ominus \tau \downarrow^r$  the *conditional for  $s$  given  $r$  with respect to  $\tau$* . To simplify notation, we will let  $\tau(s|r)$  denote  $\tau \downarrow^{(r \cup s)} \ominus \tau \downarrow^r$ . Also, if  $r = \emptyset$ , let  $\tau(s)$  denote  $\tau(s|\emptyset)$ .

The following proposition is taken from [9].

**Proposition 3.** *Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r$ ,  $s$ , and  $u$  are disjoint subsets of  $t$ . Then the following statements hold.*

1.  $\tau(s) = \tau \downarrow s$ .
2.  $\tau(r) \oplus \tau(s|r) = \tau(r \cup s)$ .
3.  $\tau(s|r) \oplus \tau(u|r \cup s) = \tau(s \cup u|r)$ .
4. *Suppose  $X \in s$ . Then,  $\tau(s|r)^{-X} = \tau(s \setminus \{X\}|r)$ .*
5.  $\tau(r) \oplus (\tau(s|r)^{-s}) = \tau(r)$ .
6.  $\tau(s|r)$  is a normal valuation for  $r \cup s$ .

**Lemma 1.** *Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r \subseteq t$ . Then*

$$\tau \oplus \iota_{\tau(r)} = \tau.$$

*Proof.* In the following equations, we use only the associativity and commutativity properties of combination, and property 2 of Proposition 3.

$$\begin{aligned} \tau \oplus \iota_{\tau(r)} &= (\tau(r) \oplus \tau(t \setminus r|r)) \oplus \iota_{\tau(r)} = (\tau(t \setminus r|r) \oplus \tau(r)) \oplus \iota_{\tau(r)} \\ &= \tau(t \setminus r|r) \oplus (\tau(r) \oplus \iota_{\tau(r)}) = \tau(t \setminus r|r) \oplus \tau(r) = \tau. \end{aligned}$$

□

Lemma 1 will help us to prove the following lemma, which expresses conditioning using the composition operator.

**Lemma 2.** *Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r$  and  $s$  are nonempty disjoint subsets of  $t$  such that  $r \cup s = t$ . Then*

$$\tau(s|r) = \iota_{\tau(r)} \triangleright \tau.$$

*Proof.*

$$\iota_{\tau(r)} \triangleright \tau = \iota_{\tau(r)} \oplus \tau \ominus \tau \downarrow r = \tau \ominus \tau \downarrow r = \tau(s|r).$$

□

**Conditional Independence for Variables.** Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r$ ,  $s$ , and  $v$  are disjoint subsets of  $t$ . We say  $r$  is *conditionally independent of  $s$  given  $v$  with respect to  $\tau$* , written as  $r \perp\!\!\!\perp_{\tau} s \mid v$ , if  $\tau \downarrow (r \cup s \cup v)$  factors into valuations  $\alpha$  for  $r \cup v$  and  $\beta$  for  $s \cup v$ , i.e.,  $\tau \downarrow r \cup s \cup v = \alpha \oplus \beta$ .

Some observations. First, while  $\tau$  has to be necessarily normal, valuations  $\alpha$  and  $\beta$  do not have to be normal. Second, the definition of conditional independence does not involve the removal operator, only the combination and marginalization operators. However, we can characterize conditional independence in terms of conditionals, which are defined using the removal operator. This is done in Proposition 4 below. Third, if  $s = \emptyset$ , then  $r \perp\!\!\!\perp_{\tau} \emptyset \mid v$  since we can let  $\alpha = \tau \downarrow (r \cup v)$  and  $\beta = \iota_v$ . This property is called *trivial independence* by Geiger and Pearl [3]. Fourth, if  $v = \emptyset$ , then we say  $r$  and  $s$  are *independent* with respect to  $\tau$ , written as  $r \perp\!\!\!\perp_{\tau} s$ , if  $\tau \downarrow (r \cup s) = \alpha \oplus \beta$ , where  $\alpha$  is a valuation for  $r$  and  $\beta$  is a valuation for  $s$ . Thus, independence is a special case of conditional independence.

The following result is proved in [9].

**Proposition 4.** *Suppose  $\tau$  is a normal valuation for  $t$ , and suppose  $r, s$ , and  $v$  are disjoint subsets of  $t$ . The following statements are equivalent.*

1.  $r \perp\!\!\!\perp_{\tau} s \mid v$ .
2.  $\tau(r \cup s \cup v) = \tau(v) \oplus \tau(r|v) \oplus \tau(s|v)$ .
3.  $\tau(r \cup s|v) = \tau(r|v) \oplus \tau(s|v)$ .
4.  $\tau(r \cup s \cup v) \oplus \tau(v) = \tau(r \cup v) \oplus \tau(s \cup v)$ .
5.  $\tau(r \cup s \cup v) = \tau(r|v) \oplus \tau(s \cup v)$ .
6.  $\tau(r|s \cup v) = \tau(r|v) \oplus \iota_{\tau(s \cup v)}$ .
7.  $\tau(r|s \cup v) = \alpha \oplus \iota_{\tau(s \cup v)}$ , where  $\alpha$  is a valuation for  $r \cup v$ .

### 3 Decomposable Compositional Models in VBS

In probability theory, inference with Bayesian networks is usually based on the idea of *local computation* of Lauritzen and Spiegelhalter [7]. This idea can be briefly expressed as follows. A Bayesian network is first transformed into a decomposable model (using well-known operations *moralization* and *triangulation* of a directed graph), and the required posterior marginal is then computed by a process exploiting the “tree” structure of decomposable models. Therefore, it is not surprising that we speak about decomposable compositional models in the VBS framework.

The tree structure of decomposable models is expressed as a running intersection property. We say that a sequence of sets  $s_1, s_2, \dots, s_n$  meets *running intersection property* (RIP) if for each  $j = 2, 3, \dots, n$  there exists a  $k < j$  such that

$$s_j \cap (s_1 \cup \dots \cup s_{j-1}) = s_j \cap s_k.$$

Decomposable compositional models are formed by multiple applications of the composition operator. Since it is not always associative (property 5 of Proposition 2), we use the following convention. If we do not specify an order using brackets, the operators will always be performed from left to right, i.e.,  $\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \tau^{\downarrow s_3} \triangleright \dots \triangleright \tau^{\downarrow s_n}$  denotes  $(\dots((\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2}) \triangleright \tau^{\downarrow s_3}) \triangleright \dots \triangleright \tau^{\downarrow s_n})$ .

**Definition 1.** *Suppose  $\tau$  is a normal valuation for  $t$ . We say  $\tau$  is decomposable if there exists a sequence  $(s_1, s_2, \dots, s_n)$  of subsets of  $t$  such that it meets RIP and*

$$\tau = \tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_n}.$$

*In this case we also say that  $\tau$  is decomposable with respect to the sequence  $(s_1, s_2, \dots, s_n)$ .*

It is well-known that if a sequence  $(s_1, s_2, \dots, s_n)$  meets RIP, then we can find another sequence starting with, say  $s_j$ , that also meets RIP. More precisely, for each  $j = 1, 2, \dots, n$  there exists (at least one) permutation  $(s_{\ell_1}, s_{\ell_2}, \dots, s_{\ell_n})$  that meets RIP and such that  $s_{\ell_1} = s_j$ . Therefore, the following assertion is of great importance.

**Theorem 1.** *If  $\tau$  is decomposable with respect to  $(s_1, s_2, \dots, s_n)$ , and  $(s_{j_1}, s_{j_2}, \dots, s_{j_n})$  is a permutation of  $(s_1, s_2, \dots, s_n)$  such that it meets RIP, then  $\tau$  is decomposable with respect to  $(s_{j_1}, s_{j_2}, \dots, s_{j_n})$ , i.e.,*

$$\tau = \tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_n} = \tau^{\downarrow s_{j_1}} \triangleright \tau^{\downarrow s_{j_2}} \triangleright \dots \triangleright \tau^{\downarrow s_{j_n}}.$$

*Proof.* The proof of this assertion is based on an important result concerning decomposable graphs, which follows from the results of S. Haberman ([4], Lemma 2.8) saying that the system of subsets (more exactly, *multiset*)

$$\{s_2 \cap s_1, s_3 \cap (s_1 \cup s_2), s_4 \cap (s_1 \cup s_2 \cup s_3), \dots, s_n \cap (s_1 \cup \dots \cup s_{n-1})\}$$

does not depend on the selected RIP ordering of the sequence  $(s_1, s_2, \dots, s_n)$ . Taking into account the running intersection property, we know that each element of this multiset is an intersection of two sets from the sequence  $(s_1, s_2, \dots, s_n)$ . Therefore, the above mentioned property can be expressed as follows: For any pair of distinct sets  $s_i, s_j$  from a system  $\{s_1, s_2, \dots, s_n\}$ , which can be ordered to meet RIP, the number of times the set  $s_i \cap s_j$  appears in the sequence

$$s_{j_2} \cap s_{j_1}, s_{j_3} \cap (s_{j_1} \cup s_{j_2}), s_{j_4} \cap (s_{j_1} \cup s_{j_2} \cup s_{j_3}), \dots, s_{j_n} \cap (s_{j_1} \cup \dots \cup s_{j_{n-1}})$$

does not depend on the RIP ordering  $(s_{j_1}, s_{j_2}, \dots, s_{j_n})$ .

Suppose  $\tau$  is decomposable with respect to  $(s_1, s_2, \dots, s_n)$ . Using the definition of composition, we have:

$$\begin{aligned} \tau = \tau^{\downarrow s_1} \oplus (\tau^{\downarrow s_2} \ominus \tau^{\downarrow s_2 \cap s_1}) \oplus (\tau^{\downarrow s_3} \ominus \tau^{\downarrow s_3 \cap (s_1 \cup s_2)}) \oplus \dots \\ \oplus (\tau^{\downarrow s_n} \ominus \tau^{\downarrow s_n \cap (s_1 \cup \dots \cup s_{n-1})}) \end{aligned}$$

which can be reorganized independently of the RIP ordering (using the properties of combination and removal and Proposition 1) as follows:

$$\begin{aligned} \tau = (\tau^{\downarrow s_1} \oplus \tau^{\downarrow s_2} \oplus \dots \oplus \tau^{\downarrow s_n}) \\ \ominus \tau^{\downarrow s_2 \cap s_1} \ominus \tau^{\downarrow s_3 \cap (s_1 \cup s_2)} \ominus \dots \ominus \tau^{\downarrow s_n \cap (s_1 \cup \dots \cup s_{n-1})} \quad \square \end{aligned}$$

## 4 Conditioning in Decomposable Compositional Models

In this section, we assume that  $\tau$  is a normal valuation for  $t$ , and that it is decomposable with respect to  $(s_1, s_2, \dots, s_n)$ . Suppose we wish to compute the conditional  $\tau(t \setminus \{X\} | \{X\})$ .

First, we have to find an ordering of  $s_1, s_2, \dots, s_n$  such that it meets RIP, and such that the first set from this ordering contains  $X$ . We know from Theorem 1 that  $\tau$  is decomposable also with respect to this new sequence. Therefore, without loss of generality we can assume that it is  $(s_1, s_2, \dots, s_n)$ , which means that we assume  $X \in s_1$ . Thus, using Lemma 2, we compute

$$\tau(t \setminus \{X\} | \{X\}) = \iota_{\tau(X)} \triangleright \tau = \iota_{\tau(X)} \triangleright (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_n}).$$



However, due to property 6 (associativity under a special condition) of Proposition 2, we have

$$\begin{aligned} \iota_{\tau(X)} \triangleright \left( (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_{n-1}}) \triangleright \tau^{\downarrow s_n} \right) \\ = \left( \iota_{\tau(X)} \triangleright (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_{n-1}}) \right) \triangleright \tau^{\downarrow s_n}, \end{aligned} \quad (4)$$

because  $s_1$ , and thus even more  $s_1 \cup \dots \cup s_{n-1}$ , contains  $\{X\} \cap s_n$ . Notice that also the other assumption of associativity under a special condition is fulfilled because

$$\tau^{\downarrow (s_1 \cup \dots \cup s_{n-1}) \cap s_n} \ggg \left( \iota_{\tau(X)} \triangleright (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_{n-1}}) \right)^{\downarrow (s_1 \cup \dots \cup s_{n-1}) \cap s_n}.$$

Repeating the idea behind equality (4), we get

$$\begin{aligned} \iota_{\tau(X)} \triangleright \left( (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_{n-2}}) \triangleright \tau^{\downarrow s_{n-1}} \right) \\ = \left( \iota_{\tau(X)} \triangleright (\tau^{\downarrow s_1} \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_{n-2}}) \right) \triangleright \tau^{\downarrow s_{n-1}}. \end{aligned}$$

Thus, eventually, after repeating this step  $(n-1)$  times we get

$$\tau(t \setminus \{X\} | \{X\}) = \iota_{\tau(X)} \triangleright \tau = (\iota_{\tau(X)} \triangleright \tau^{\downarrow s_1}) \triangleright \tau^{\downarrow s_2} \triangleright \dots \triangleright \tau^{\downarrow s_n},$$

from which we see that  $\tau(t \setminus \{X\} | \{X\})$  is again a decomposable model with respect to  $(s_1, s_2, \dots, s_n)$ . Let  $\hat{\tau}$  denote  $\tau(t \setminus \{X\} | \{X\})$ . We can compute the marginal valuations of  $\hat{\tau}$  (that are necessary to represent this multidimensional valuation as a compositional model) as follows:

$$\begin{aligned} \hat{\tau}^{\downarrow s_1} &= \iota_{\tau(X)} \triangleright \tau^{\downarrow s_1} \\ \hat{\tau}^{\downarrow s_2} &= \hat{\tau}^{\downarrow s_2 \cap s_1} \triangleright \tau^{\downarrow s_2} \\ &\vdots \\ \hat{\tau}^{\downarrow s_n} &= \hat{\tau}^{\downarrow s_n \cap (s_1 \cup \dots \cup s_{n-1})} \triangleright \tau^{\downarrow s_n}. \end{aligned}$$

Notice that this computation is tractable because, thanks to RIP, at each step  $\hat{\tau}^{\downarrow s_i \cap (s_1 \cup \dots \cup s_{i-1})}$  is easily computable since  $s_i \cap (s_1 \cup \dots \cup s_{i-1})$  must be contained in some  $s_k$  for  $k < i$ .

## 5 Summary and Conclusions

We have described the abstract VBS framework including the composition operator. We have shown that conditioning, which is defined using the combination, marginalization, and removal operators of VBS, can be expressed in terms of the composition operator. We have defined a decomposable compositional model as a special type of a compositional model in the VBS framework. We have shown that for decomposable compositional models, conditional valuations can be computed efficiently using local computation. All of this is done in the abstract VBS

framework. Since the VBS framework applies to many different uncertainty calculi, we have effectively defined decomposable compositional models, and efficient computation of conditionals in decomposable compositional models, for any calculi that fits in the VBS framework. For example, because Spohn's epistemic theory fits in the VBS framework, all results described in this paper applies to this calculus.

**Acknowledgements.** This work has been supported in part by funds from grant GAČR 403/12/2175 to the first author, and from the Ronald G. Harper Distinguished Professorship at the University of Kansas to the second author. We are grateful to Milan Studený for valuable discussions and comments.

## References

1. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* 38, 325–339 (1967)
2. Dubois, D., Prade, H.: *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press (1988)
3. Geiger, D., Pearl, J.: Logical and algorithmic properties of independence and their application to Bayesian networks. *Annals of Mathematics and Artificial Intelligence* 2, 165–178 (1990)
4. Haberman, S.J.: *The Analysis of Frequency Data*. The University of Chicago Press, Chicago (1974)
5. Jiroušek, R.: Composition of probability measures on finite spaces. In: Geiger, D., Shenoy, P.P. (eds.) *Uncertainty in Artificial Intelligence: Proceedings of the 13th Conference (UAI 1997)*, pp. 274–281. Morgan Kaufmann, San Francisco (1997)
6. Jiroušek, R., Shenoy, P.P.: Compositional models in valuation-based systems. Working Paper No. 325, School of Business, University of Kansas, Lawrence, KS (2011)
7. Lauritzen, S.L., Spiegelhalter, D.J.: Local computation with probabilities on graphical structures and their application to expert systems. *Journal of Royal Statistical Society, series B* 50, 157–224 (1988)
8. Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, Princeton (1976)
9. Shenoy, P.P.: Conditional independence in valuation-based systems. *International Journal of Approximate Reasoning* 10, 203–234 (1994)
10. Shenoy, P.P.: Binary join trees for computing marginals in the Shenoy-Shafer architecture. *International Journal of Approximate Reasoning* 17, 239–263 (1997)
11. Spohn, W.: A general non-probabilistic theory of inductive reasoning. In: Shachter, R.D., Levitt, T.S., Lemmer, J.F., Kanal, L.N. (eds.) *Uncertainty in Artificial Intelligence 4 (UAI 1990)*, pp. 274–281. North Holland (1990)