To appear in: International Symposium on Imprecise Probability: Theories and Applications (ISIPTA), Lugano, Switzerland, Jul. 2017.

Efficient Computation of Belief Theoretic Conditionals

Lalintha G. Polpitiya LALINTHA@UMIAMI.EDU
Kamal Premaratne KAMAL@MIAMI.EDU
Manohar N. Murthi MMURTHI@MIAMI.EDU
Dilip Sarkar SARKAR@MIAMI.EDU
University of Miami
Coral Gables, Florida (USA)

Abstract

Dempster-Shafer (DS) belief theory is a powerful general framework for dealing with a wider variety of uncertainties in data. As in Bayesian probability theory, the conditional operation plays a critical role in DS theoretic strategies for evidence updating and fusion. A major limitation associated with the application of DS theoretic techniques for reasoning under uncertainty is the absence of a feasible computational framework to overcome the prohibitive computational burden this conditional operation entails. This paper addresses this critical challenge via a novel generalized conditional computational model — *DS-Conditional-One* — which allows the conditional to be computed in significantly less computational and space complexity. This computational model also provides valuable insight into the DS theoretic conditional itself and can be utilized as a tool for visualizing the conditional computation. We provide a thorough analysis and experimental validation of the utility, efficiency, and implementation of the proposed data structures and algorithms for carrying out both the *Dempster's conditional* and *Fagin-Halpern conditional,* the two most widely utilized DS theoretic conditional strategies.

Keywords: Dempster-Shafer belief theory; Dempster's conditional; Fagin-Halpern conditional; data structures; algorithms; computational complexity.

1. Introduction

The Dempster-Shafer (DS) belief theory (Dempster, 1967, 1968; Shafer, 1976), also referred to as evidence theory, is a powerful and convenient framework that can handle a wide variety of data imperfections (Shafer, 1990; Smets, 1999). With the greater expressiveness and flexibility in evidential reasoning and decision-making that they offer, DS theoretic (DST) methods are finding increased utilization in numerous application scenarios and have generated an active research field (Yager and Liu, 2008; Denœux, 2016).

Motivation. As in the Bayesian methods, the conditional operation plays a pivotal role in DST strategies for evidence updating and fusion, and in general, for reasoning under uncertainty. Among these various notions that have been proposed over the years, perhaps the most widely used DST conditional notion is the *Dempster's conditional* (Shafer, 1976; Klawonn and Smets, 1992; Nguyen and Smets, 1993; Xu and Smets, 1996; Smets, 2002). On the other hand, the *Fagin-Halpern (FH) conditional* can be considered as the most natural generalization of the probabilistic conditional notion because of its close connection with the inner and outer conditional probability measures (Fagin and Halpern, 1990). The recent work on the DST conditional approach (Premaratne et al., 2009; Wickramarathne et al., 2011) is based on this FH conditional.

Challenges. In spite of the advantages they offer, DST implementations in current use are restricted to smaller frames of discernment (FoDs) because of the prohibitive computational burden

that larger FoDs impose on existing methods. While this difficulty has been addressed via several approximation methods (Yager and Liu, 2008; Denœux, 2016), such approaches usually require one to compromise the quality of the generated results for computational efficiency, and some approaches cannot be extended for computing the DST *conditionals.* Exact (or sufficiently precise) computation of conditionals is of paramount importance because the quality of results generated from DST strategies depend directly on the precision of the conditional. A review of current implementations (Yager and Liu, 2008; Augustin et al., 2014; Denœux, 2016; SIPTA, 2017) confirms that work is needed to overcome these computational limitations associated with the DST conditionals. A fast Möbius transform (FMT), which is analogous to the fast Fourier transform (FFT), has been developed and employed for efficient precise computation of DST notions (Thoma, 1989; Kennes, 1992). Polpitiya et al. (2016) proposes several data structures which enable highly efficient exact computation of the DST notions of belief and plausibility, but it does not address the computation of DST conditionals.

As for the Dempster's conditional, perhaps the most thorough discussion for carrying out its precise computation appears in Klawonn and Smets (1992) and Smets (2002). It provides a matrix calculus based algorithm to compute Dempster's conditional masses. However, this approach is feasible only on smaller frames because of the matrix operations it requires. It is not applicable for FH conditional computation. As for the FH conditional, the work in Wickramarathne et al. (2013) provides a method to identify the propositions that retain non-zero support after FH conditioning, but it does not address conditional computation of these propositions.

Contributions. The main contribution of this paper is a completely new generalized model for computing DST *conditionals.* This conditional computational model — *DS-Conditional-One* — offers significantly greater flexibility and computational capability for implementation of DST conditional strategies. We provide the DS-Conditional-One computational model along with its complexity analysis, experimental validation of the utility, efficiency, implementation of the associated data structures and algorithms. This model can be employed to compute both the FH and Dempster's conditional beliefs of an *arbitrary* proposition. This is exactly the challenge that Shafer refers to in Shafer (1990, p.348), viz., *"It remains to be seen how useful the fast Möbius transform will be in practice. It is clear, however, that it is not enough to make arbitrary belief function computations feasible."*

By reducing the number of operations being executed, the proposed approach takes significantly less computational and space complexity when compared with other approaches for conditional computation. As an example, our experiment results demonstrate that the average computational time taken to compute the conditional belief of an arbitrary proposition by the proposed approach is less than 2 (μs) for a FoD of size 10 and 0.7 (ms) for a FoD of size 20 ($\sim$1 million focal elements). This new model can also be utilized as a visualization tool for conditional computations and in analyzing characteristics of conditioning and updating operations. All software routines are available at ProFuSELab (2017). We believe that this computational model and the associated data structures constitute a significant step toward filling the void between what the DST framework can offer for reasoning under uncertainty and the practical implementation of DST strategies.

This paper is organized as follows: Section 2 provides a review of essential DST notions and computational tools. Our DS-Conditional-One computational model and our algorithms for efficient computation of DST conditionals appear next in Sections 3 and 4, respectively. The experimental results are provided next in Section 5. Finally, Section 6 offers some concluding remarks.

2. Preliminaries: DS Belief Theory

2.1 DST Basic Notions

In DS theory, the *frame of discernment (FoD)* refers to the set of all possible mutually exclusive and exhaustive propositions (Shafer, 1976). We consider the case where the FoD is finite and we denote it as $\Theta = \{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$. Proposition $\{\theta_i\}$, which is referred to as a *singleton,* represents the lowest level of discernible information. The power set of Θ, denoted by 2^Θ, form all the propositions of interest in DS theory. A proposition that is not a singleton is referred to as a *composite.* The set $A \setminus B$ denotes all singletons in $A \subseteq \Theta$ that are not included in $B \subseteq \Theta$, i.e., $A \setminus B = \{\theta_i \in \Theta \mid \theta_i \in A,\ \theta_i \notin B\}$. We use $\overline{A}$ to denote $\Theta \setminus A$ and $|A|$ to denote the cardinality of A. Note that $|\Theta| = n$.

In DS theory, the 'support' that is being strictly allocated to a proposition is captured via

Definition 1 (Basic Belief Assignment (BBA) or Masses) *The mapping* $m : 2^\Theta \mapsto [0,1]$ *is said to be a* basic belief assignment (BBA) *or a* mass assignment *if*

$$m(\emptyset) = 0 \textit{ and } \sum_{A \subseteq \Theta} m(A) = 1. \qquad \blacksquare$$

The mass of a composite proposition is free to move into its individual singletons, which allows one to model the notion of *ignorance.* Complete ignorance can be modeled via the *vacuous BBA,* viz., $m(\Theta) = 1$ and $m(A) = 0$, $\forall A \neq \Theta$. Propositions that possess nonzero mass are referred to as *focal elements;* the set of all focal elements in a FoD is referred to as its *core* $\mathfrak{F}$, i.e., $\mathfrak{F} = \{A \subseteq \Theta \mid m(A) > 0\}$. Note that $|\mathfrak{F}|$ is the number of focal elements. $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ is referred to as the *body of evidence (BoE).*

Definition 2 (Belief) *Given a BoE* $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$*, the* belief *and* plausibility *functions are the mappings* $Bl : 2^\Theta \mapsto [0,1]$ *and* $Pl : 2^\Theta \mapsto [0,1]$*, respectively, where*

$$Bl(A) = \sum_{B \subseteq A} m(B); \quad Pl(A) = \sum_{\substack{B \subseteq \Theta \\ B \cap A \neq \emptyset}} m(B). \qquad \blacksquare$$

The *belief* assigned to a proposition takes into account the support for all of its subsets. It is easy to see that, $Pl(A) = 1 - Bl(\overline{A}) \geq Bl(A)$, $\forall A \subseteq \Theta$. So, the *plausibility* measures the extent to which a proposition is plausible, i.e., the amount of belief not strictly supporting the complement of the proposition. Propositions that possess nonzero belief are denoted by $\widehat{\mathfrak{F}}$, i.e., $\widehat{\mathfrak{F}} = \{A \subseteq \Theta \mid Bl(A) > 0\}$.

Given a valid belief function $Bl : 2^\Theta \mapsto [0,1]$, one may generate the corresponding BBA $m : 2^\Theta \mapsto [0,1]$ via the Möbius transform (Shafer, 1976)

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} Bl(B),\ \forall A \subseteq \Theta. \tag{1}$$

The following notation will be useful for our work:

$$\mathcal{S}(A; B) = \sum_{\substack{\emptyset \neq C \subseteq A; \\ \emptyset \neq D \subseteq B}} m(C \cup D). \tag{2}$$

So, $\mathcal{S}(A; B)$ denotes the sum of all masses of propositions that 'straddle' both $A \subseteq \Theta$ and $B \subseteq \Theta$.

The following result is of critical importance for our work.

Proposition 3 *Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$. For $B \subseteq \Theta$, consider the mappings $\Gamma_A : 2^\Theta \mapsto [0,1]$ and $\Pi_A : 2^\Theta \mapsto [0,1]$, where*

$$\Gamma_A(B) = \sum_{\emptyset \neq X \subseteq \overline{A}} m((A \cap B) \cup X); \;\; \Pi_A(B) = \sum_{Y \subseteq (A \cap B)} \Gamma_A(Y).$$

Then the following are true:

(i) *$\Gamma_A(A \cap B) = \Gamma_A(B)$ and $\Pi_A(A \cap B) = \Pi_A(B)$. So, w.l.o.g., we assume that $B \subseteq A$.*

(ii) *$\Gamma_A(\emptyset) = Bl(\overline{A})$.*

Proof These follow by direct substitution. ■

2.2 Fagin-Halpern (FH) Conditional

FH conditional can be considered the most natural generalization of the probabilistic conditional notion because of its close connection with the inner and outer conditional probability measures in probability theory (Fagin and Halpern, 1990).

Definition 4 (Fagin-Halpern (FH) Conditional) *(Fagin and Halpern, 1990) Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \in \widehat{\mathfrak{F}}$. The* conditional belief *$Bl(B|A)$ of B given the conditioning event A is*

$$Bl(B|A) = \frac{Bl(A \cap B)}{Bl(A \cap B) + Pl(A \cap \overline{B})}.$$ ■

The conditional plausibility $Pl(B|A)$ of B given A is computed as $Pl(B|A) = 1 - Bl(\overline{B}|A)$. Of course, once the conditional beliefs of all the propositions are computed, one may obtain the corresponding conditional BBA via a Möbius transform of the type in (1).

Suppose the BoE $\{\Theta, \mathfrak{F}, m(\cdot)\}$ is being conditioned w.r.t. the proposition $A \in \widehat{\mathfrak{F}}$. The propositions that retain a nonzero mass after conditioning are referred to as the *conditional focal elements;* the set of all such conditional focal elements is referred to as the *conditional core* $\mathfrak{F}_A$, i.e., $\mathfrak{F}_A = \{B \subseteq A \in \widehat{\mathfrak{F}} \mid m(B|A) > 0\}$.

In our work, we will exploit several previous results related to the conditional core (Kulasekere et al., 2004; Wickramarathne et al., 2013). Of particular importance is the following result:

Lemma 5 *(Kulasekere et al., 2004) Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \in \widehat{\mathfrak{F}}$. Then,*

(i) *$m(B|A) = 0$ whenever $\overline{A} \cap B \neq \emptyset$, and*

(ii) *$Bl(B|A)$ can be expressed as*

$$Bl(B|A) = \frac{Bl(A \cap B)}{Pl(A) - \mathcal{S}(\overline{A}; A \cap B)}, \; B \subseteq A.$$ ■

Note that, **(i)** states that FH conditioning annuls those propositions that 'straddle' the conditioning proposition A and its complement $\overline{A}$. So, w.l.o.g., for FH conditioning, one may consider only those propositions $B \subseteq A$.

For our work, we will need the following alternate expression for the FH conditional:

Proposition 6 *Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \in \widehat{\mathfrak{F}}$. Then, we may express $Bl(B|A)$ as*

$$Bl(B|A) = \frac{Bl(A \cap B)}{1 - Bl(\overline{A}) - \mathcal{S}(\overline{A}; A \cap B)}, \; B \subseteq \Theta.$$ □

Proof This follows directly from Lemma 5**(ii)** by using the fact that $Bl(A) = 1 - Pl(\overline{A})$. ■

2.3 Dempster's Conditional

Dempster's conditional is perhaps the most widely employed DST conditional notion.

Definition 7 (Dempster's Conditional) *(Shafer, 1976) Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$ s.t. $Bl(\overline{A}) \neq 1$, or equivalently, $Pl(A) \neq 0$. The* conditional belief $Bl(B\|A)$ *of B given the conditioning event A is*

$$Bl(B\|A) = \frac{Bl(\overline{A} \cup B) - Bl(\overline{A})}{1 - Bl(\overline{A})}. \qquad \blacksquare$$

One may compute the corresponding conditional mass $m(B\|A)$ and $Pl(B\|A)$ from $Bl(B\|A)$. Similarly to FH conditioning, Dempster's conditioning also annuls masses of all those propositions that 'straddle' the conditioning proposition A and its complement $\overline{A}$. So, w.l.o.g., for Dempster's conditioning, one may consider only those propositions $B \subseteq A$.

For our work, we will need the following alternate expression for the Dempster's conditional:

Proposition 8 *Consider the BoE $\mathcal{E} = \{\Theta, \mathfrak{F}, m(\cdot)\}$ and $A \subseteq \Theta$ s.t. $Bl(\overline{A}) \neq 1$. Then, $Bl(B\|A)$ can be expressed as*

$$Bl(B\|A) = \frac{Bl(A \cap B) + \mathcal{S}(\overline{A}; A \cap B)}{1 - Bl(\overline{A})}, \; B \subseteq \Theta. \qquad \square$$

Proof This follows directly from Definition 7 by using the fact that $Bl(\overline{A} \cup B) = Bl(\overline{A} \cup (A \cap B)) = Bl(\overline{A}) + Bl(A \cap B) + \mathcal{S}(\overline{A}; A \cap B)$. $\blacksquare$

Propositions 6 and 8 highlight an important fact: the three quantities $Bl(\overline{A})$, $Bl(A \cap B)$, and $\mathcal{S}(\overline{A}; A \cap B)$ fully determine both FH and Dempster's conditionals $Bl(B|A)$ and $Bl(B\|A)$, respectively. It is this fact that we exploit for computing the conditionals of an arbitrary proposition.

2.4 The REGAP Property

The work in Polpitiya et al. (2016) proposes new data structures — *DS-Vector, DS-Matrix and DS-Tree* — and computationally efficient algorithms for computing the basic DST operations of belief and plausibility. For this purpose, the authors utilize what is referred to as the *REGAP (REcursive Generation of and Access to Propositions)* property.

To be more specific, consider the FoD $\Theta = \{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$. Suppose we desire to determine the belief potential $Bl(A)$ associated with $A = \{\theta_{k_0}, \theta_{k_1}, \ldots, \theta_{k_{|A|-1}}\} \subseteq \Theta$. Then, $REGAP(A)$ recursively generates all the $2^{|A|} - 1$ propositions whose masses are required to compute $Bl(A)$, viz., all subsets of A (including A itself). It is implemented in the following manner: Start with $\{\emptyset\}$. First insert the singleton $\{\theta_{k_0}\} \in A$. Only one proposition is associated with this singleton, viz., $\{\emptyset\} \cup \{\theta_{k_0}\} = \{\theta_{k_0}\}$ itself. Next insert another singleton $\{\theta_{k_1}\} \in A$. The new propositions that are associated with this singleton are $\{\emptyset\} \cup \{\theta_{k_1}\} = \{\theta_{k_1}\}$ and $\{\theta_{k_0}\} \cup \{\theta_{k_1}\} = \{\theta_{k_0}, \theta_{k_1}\}$. Inserting the next singleton $\{\theta_{k_2}\} \in A$ brings the new propositions $\{\emptyset\} \cup \{\theta_{k_2}\} = \{\theta_{k_2}\}$, $\{\theta_{k_0}\} \cup \{\theta_{k_2}\} = \{\theta_{k_0}, \theta_{k_2}\}$, $\{\theta_{k_1}\} \cup \{\theta_{k_2}\} = \{\theta_{k_1}, \theta_{k_2}\}$, and $\{\theta_{k_0}, \theta_{k_1}\} \cup \{\theta_{k_2}\} = \{\theta_{k_0}, \theta_{k_1}, \theta_{k_2}\}$. In essence, when a new singleton is added, new propositions associated with it can be recursively generated by adding the new singleton to each existing proposition. Of course, all propositions of interest within the FoD Θ can be generated by $REGAP(\Theta)$, i.e., when $A = \Theta$.

The propositions recursively generated via the REGAP property can be represented as a vector, DS-Vector, a matrix, DS-Matrix, or a tree, DS-Tree, and utilized to capture a BoE. We will utilize this REGAP property and the DS-Matrix structure in this work too.

3. DS-Conditional-One Computational Model

DS-Conditional-One is a computational model that enables one to compute the FH and Dempster's conditional beliefs of an *arbitrary* proposition. DS-Conditional-One model facilitates the representation, access, and efficient computation of the quantities that are needed to compute these conditionals (see Propositions 6 and 8).

Henceforth, we will denote the conditioning proposition A, its complement $\overline{A}$, and the conditioned proposition B as $\{a_0, a_1, \ldots, a_{|A|-1}\}$, $\{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$, and $\{b_0, b_1, \ldots, b_{|B|-1}\}$, respectively. Here, $\Theta = \{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$ denotes the FoD and $a_i, \alpha_j, b_k \in \Theta$. When dealing with FH and Dempster's conditioning, it is implicitly assumed that $A \in \widehat{\mathfrak{F}}$ and $Bl(\overline{A}) \neq 1$, respectively.

Furthermore, we will represent singletons of the conditioning event $A = \{a_0, a_1, \ldots, a_{|A|-1}\}$ as *column singletons* and singletons of the complement of conditioning event $\overline{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$ as *row singletons* in a DS-Matrix. See Fig. 1.

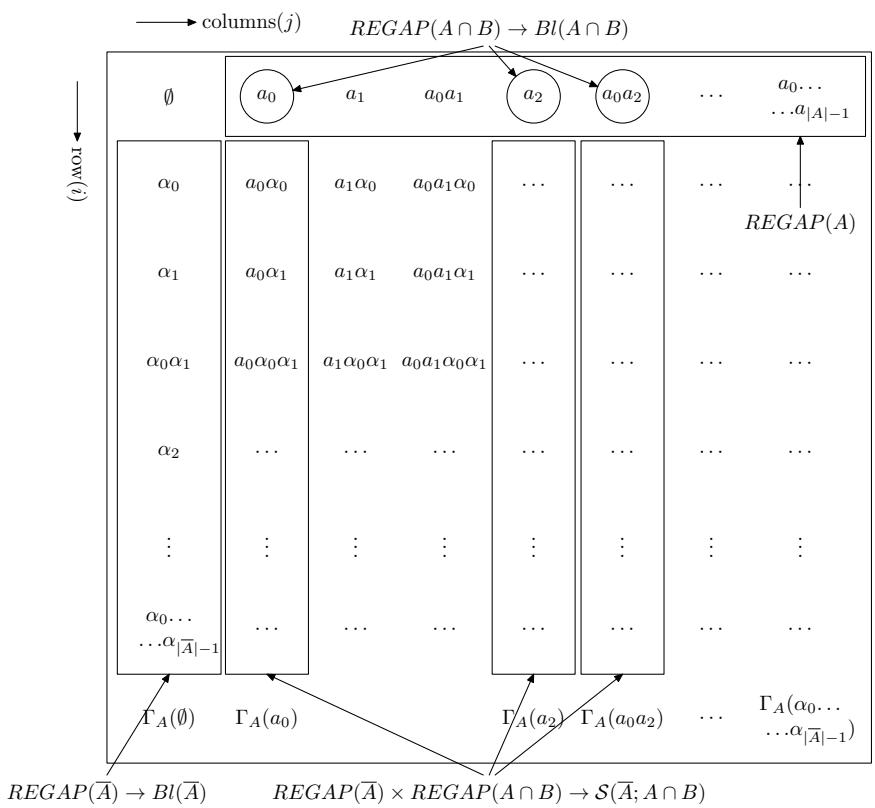

Figure 1: **DS-Conditional-One model.** Quantities related to $Bl(B|A)$ computation when $A = \{a_0, a_1, \ldots, a_{|A|-1}\}$ and $\overline{A} = \{\alpha_0, \alpha_1, \ldots, \alpha_{|\overline{A}|-1}\}$, and $B = \{a_0, a_2\} \subseteq A$.

The proposed DS-Conditional-One computational model allows direct identification of $REGAP(A)$, $REGAP(\overline{A})$, $REGAP(A \cap B)$, $(REGAP(\overline{A}) \times REGAP(A \cap B))$, $(REGAP(A) \times REGAP(\overline{A}))$, and $\Gamma_A(C)$, $\forall C \subseteq B$. Among these, the following three quantities are required to compute both FH and Dempster's conditional beliefs (see Propositions 6 and 8): **(a)** $REGAP(A \cap B)$: Use this to compute $Bl(A \cap B)$ (see Algorithm 1). **(b)** $REGAP(\overline{A})$: Use this to compute $Bl(\overline{A})$ (see Algorithm 2). **(c)** $(REGAP(\overline{A}) \times REGAP(A \cap B))$, the Cartesian product of $REGAP(\overline{A})$ and $REGAP(A \cap B)$: Use this to compute $\mathcal{S}(\overline{A}; A \cap B)$ (see Algorithm 3).

Fig. 1 depicts these quantities for $A = \{a_0, a_1, \ldots, a_{|A|-1}\}$ and $B = \{a_0, a_2\} \subseteq A$

In the algorithms to follow, we use a lookup table named $power$ to enhance the computational efficiency. It contains 2 to the power of singleton indexes in increasing order and it is implemented using a dynamic array that replaces run-time computation of 2 to the power values with a simpler array indexing operation. $power[i]$, the i-th entry of the $power$ table, refers to 2^i. $index[]$ is a dynamic array which keeps the indexes of subset propositions of $A \cap B$.

Algorithm 1 Compute $Bl(A \cap B)$ (with complexity $\mathcal{O}(2^{|A \cap B|})$)

```
1:  procedure BLB(Singletons A, Singletons B, DS-Matrix BBA)
2:      belief ← 0
3:      count ← 0
4:      for each a_i in A ∩ B do
5:          index[count] ← power[i]
6:          temp ← count
7:          count ← count + 1
8:          for j ← 0, temp − 1 do
9:              index[count] ← index[j] + power[i]
10:             count ← count + 1
11:         end for
12:     end for
13:     for i ← 0, power[|A ∩ B|] − 2 do
14:         belief ← belief + BBA[0][index[i]]
15:     end for
16:     Return belief
17: end procedure
```

Time Complexity of Algorithm 1. This computes $Bl(A \cap B)$ in $\mathcal{O}(2^{|A \cap B|})$ complexity. *Line #1:* The algorithm inputs are the conditioning event A, conditioned event B, and the DS-Matrix BBA. *Lines #4-12:* The outer loop is executed $|A \cap B|$ times. *Lines #8-11:* The inner loop is executed $temp - 1$ times. It can be shown that for $\ell = 0, 1, 2, \ldots, |A \cap B| - 1$, $temp = (2^\ell - 1)$. *Lines #5 and #9* are constant time operations. Thus, the computational complexity of *lines #4-12* is given by

$$\sum_{\ell=0}^{|A \cap B|-1} (1 + temp) = \sum_{\ell=0}^{|A \cap B|-1} 2^\ell = 2^{|A \cap B|} - 1 = \mathcal{O}(2^{|A \cap B|}). \tag{3}$$

Lines #13-15: The required number of iterations is $2^{|A \cap B|} - 1$ and the complexity of this segment is $\mathcal{O}(2^{|A \cap B|})$. *Line #16:* The algorithm output is $Bl(A \cap B)$.

Algorithm 2 Compute $Bl(\overline{A})$ (with complexity $\mathcal{O}(2^{|\overline{A}|})$)

1: **procedure** BLCOMP(Singletons $\overline{A}$, DS-Matrix BBA)
2: $belief \leftarrow 0$
3: **for** $i \leftarrow 1, power[|\overline{A}|] - 1$ **do**
4: $belief \leftarrow belief + BBA[i][0]$
5: **end for**
6: Return $belief$
7: **end procedure**

Time Complexity of Algorithm 2. This computes $Bl(\overline{A})$ in $\mathcal{O}(2^{|\overline{A}|})$ complexity. *Line #1:* The algorithm inputs are the complement of conditioning event $\overline{A}$ and the DS-Matrix BBA. *Lines #3-5:* The required number of iterations is $2^{|\overline{A}|} - 1$ and the computational complexity of this segment is $\mathcal{O}(2^{|\overline{A}|})$. *Line #6:* The algorithm output is the belief potential $Bl(\overline{A})$.

Algorithm 3 Compute $\mathcal{S}(\overline{A}; A \cap B)$ (with complexity $\mathcal{O}(2^{|\overline{A}|+|A\cap B|})$)

1: **procedure** STRAD(Singletons $\overline{A}$, Singletons A, Singletons B, DS-Matrix BBA)
2: $belief \leftarrow 0$
3: $count \leftarrow 0$
4: **for** each a_i in $A \cap B$ **do**
5: $index[count] \leftarrow power[i]$
6: $temp \leftarrow count$
7: $count \leftarrow count + 1$
8: **for** $j \leftarrow 0, temp - 1$ **do**
9: $index[count] \leftarrow index[j] + power[i]$
10: $count \leftarrow count + 1$
11: **end for**
12: **end for**
13: **for** $i \leftarrow 1, power[|\overline{A}|] - 1$ **do**
14: **for** $j \leftarrow 0, power[|A \cap B|] - 2$ **do**
15: $belief \leftarrow belief + BBA[i][index[j]]$
16: **end for**
17: **end for**
18: Return $belief$
19: **end procedure**

Time Complexity of Algorithm 3. This computes $\mathcal{S}(\overline{A}; A \cap B)$ in $\mathcal{O}(2^{|\overline{A}|+|A\cap B|})$ complexity. *Line #1:* The algorithm inputs are the complement of conditioning event $\overline{A}$, the conditioning and conditioned propositions A and B, respectively, and the DS-Matrix BBA. *Lines #4-12:* Subset propositions of $A \cap B$ are generated via $REGAP(A \cap B)$. Computational complexity of this segment is $\mathcal{O}(2^{|A\cap B|})$, which can be obtained from equation 3. *Lines #13-17:* The outer loop is executed $(2^{|\overline{A}|} - 1)$ times. *Lines #14-16:* The inner loop is executed $(2^{|A\cap B|} - 1)$ times. Complexity of an access operation is $\mathcal{O}(1)$. Thus, the computational complexity of *lines #13-17* is $(2^{|\overline{A}|} - 1)\,(2^{|A\cap B|} - 1) = \mathcal{O}(2^{|\overline{A}|+|A\cap B|})$. *Line #18:* The algorithm output is $\mathcal{S}(\overline{A}; A \cap B)$.

Space Complexity of Algorithms 1, 2, and 3. The matrix in Fig. 1 is of size $2^{|A|} \times 2^{|\overline{A}|}$. Hence, the space complexity associated with each algorithm above is $\mathcal{O}(2^{|\Theta|})$.

Note that, in the DS-Conditional-One model, $REGAP(A)$ captures all propositions that may contribute to the conditional core $\mathfrak{F}_A$, and $REGAP(\overline{A})$ and $(REGAP(A) \times REGAP(\overline{A}))$, the Cartesian product of $REGAP(A)$ and $REGAP(\overline{A})$, capture all propositions whose masses are annulled (as identified by Lemma 5 (Kulasekere et al., 2004)). See Fig 1.

4. Efficient Computation of DST Conditionals

4.1 Computation of the FH Conditional Belief of an Arbitrary Proposition

To compute the FH conditional belief of an arbitrary proposition B, one can now use the expression in Proposition 6, where $Bl(A \cap B)$, $Bl(\overline{A})$ and $\mathcal{S}(\overline{A}; A \cap B)$ are obtained via Algorithms 1, 2, and 3, respectively. Thus the computational complexity of this computation remains as $\mathcal{O}(2^{|\overline{A}|+|A \cap B|})$.

As an example, to compute $Bl(B|A)$, where $B = \{a_0, a_2\}$, we may proceed as follows: **(a)** $REGAP(A \cap B)$ captures the propositions that contribute to $Bl(A \cap B)$. Use Algorithm 1 to compute this. **(b)** $REGAP(\overline{A})$ captures the propositions that contribute to $Bl(\overline{A})$. Use Algorithm 2 to compute this. Note that $Bl(\overline{A})$ is represented by $\Gamma_A(\emptyset)$ in Fig. 1. **(c)** The Cartesian product $(REGAP(\overline{A}) \times REGAP(A \cap B))$ captures the propositions that contribute to $\mathcal{S}(\overline{A}; A \cap B)$. Use Algorithm 3 to compute this. $\mathcal{S}(\overline{A}; A \cap B) = \Gamma_A(\{a_0\}) + \Gamma_A(\{a_2\}) + \Gamma_A(\{a_0, a_2\})$.

Then, $Bl(B|A)$ for $B = \{a_0, a_2\}$ is computed as

$$Bl(B|A) = \frac{Bl(A \cap B)}{1 - \Gamma_A(\{\emptyset\}) - \Gamma_A(\{a_0\}) - \Gamma_A(\{a_2\}) - \Gamma_A(\{a_0, a_2\})}. \tag{4}$$

4.2 Computation of the Dempster's Conditional Belief of an Arbitrary Proposition

To compute the Dempster's conditional belief of an arbitrary proposition B, one can use the expression in Proposition 8, where $Bl(A \cap B)$, $Bl(\overline{A})$ and $\mathcal{S}(\overline{A}; A \cap B)$ are obtained via Algorithms 1, 2, and 3, respectively. Thus the computational complexity is $\mathcal{O}(2^{|\overline{A}|+|A \cap B|})$.

Consider the same example as before, viz., $B = \{a_0, a_2\}$. Then, we may compute $Bl(B\|A)$ as

$$Bl(B\|A) = \frac{Bl(A \cap B) + \Gamma_A(\{a_0\}) + \Gamma_A(\{a_2\}) + \Gamma_A(\{a_0, a_2\})}{1 - \Gamma_A(\{\emptyset\})}. \tag{5}$$

Computation of the Dempster's Conditional Mass Using Specialization Matrix. It is noteworthy that Klawonn and Smets (1992) and Smets (2002) have proposed a matrix calculus based algorithm for *direct* computation of Dempster's conditional *masses.* It employs a $2^{|\Theta|} \times 2^{|\Theta|}$-sized stochastic matrix $\mathfrak{S}_A$ (with each entry '0' or '1') referred to as the conditioning specialization matrix and a $2^{|\Theta|} \times 1$-sized vector $m(\cdot)$ containing the BoE's focal elements. Then $m(\cdot\|A) = \mathfrak{S}_A \cdot m(\cdot)$ yields Dempster's conditioning masses *without normalization.* The computational and space complexity of the specialization matrix multiplication is $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$, a prohibitive burden even for modest FoD sizes.

5. Experiments

Recall that Algorithms 1, 2, and 3 yield all the parameters (viz., $Bl(A \cap B)$, $Bl(\overline{A})$, and $\mathcal{S}(\overline{A}; A \cap B)$) required for both FH and Dempster's conditional belief computations. Once these quantities are

computed, computational times for both conditional belief computations are similar because they require constant time (see Propositions 6 and 8).

For a given FoD size, we selected a random set of focal elements, with randomly selected mass values, and conducted 10,000 conditional computations for randomly chosen propositions A and $B \subseteq A$. Table 1 lists the average computational times taken by the DS-Conditional-One model and the specialization matrix based method in Klawonn and Smets (1992) and Smets (2002).

With the DS-Conditional-One model (which applies to *both* FH and Dempster's conditionals), we use a 'brute force' approach to compute all the conditional beliefs (i.e., compute the conditional belief of every proposition); we then use the FMT to get the conditional masses for all the propositions (Shafer, 1976; Fagin and Halpern, 1990). The specialization matrix based method (which applies to the Dempster's conditional *only*) yields the conditional masses of *all* propositions, but the time taken already far exceeds what the DS-Conditional-One model takes (even including the FMT). So we did not compute the conditional beliefs with the specialization matrix based method (which would have required the FMT).

All conditional computations for an *arbitrary* proposition were done on an iMac running Mac OS X 10.12.3 (with 2.9GHz Intel Core i5 processor and 8GB of 1600MHz DDR3 RAM). Conditional computations for *all* propositions were done on the same iMac for smaller FoDs and on a supercomputer (`http://ccs.miami.edu/pegasus`) for larger FoDs (<u>underlined</u> in Table 1). The complete C++ library is available at ProFuSELab (2017).

Method →		**DS-Conditional-One Model**			**Specialization Matrix**
Conditional →		**FH or Dempster's**			**Dempster's**
FoD		$Bl(B\|A)$ or $Bl(B\|A)$	$Bl(B\|A)$ or $Bl(B\|A)$	$m(B\|A)$ or $m(B\|A)$	$m(B\|A)$
$\|\Theta\|$	**Max.** $\|\mathfrak{F}\|$	**(Arbitrary)**	**(All)**	**(All)**	**(All)**
2	3	0.0005	0.0011	0.0016	0.0011
4	15	0.0005	0.0038	0.0050	0.0063
6	63	0.0006	0.0128	0.0170	0.0696
8	255	0.0009	0.0517	0.0679	1.0154
10	1,023	0.0017	0.2428	0.3090	<u>93.1590</u>
12	4,095	0.0040	1.3528	1.6186	<u>1485.6300</u>
14	16,383	0.0120	<u>18.4885</u>	<u>22.4995</u>	<u>25051.8200</u>
16	65,535	0.0405	<u>146.1480</u>	<u>151.9600</u>	<u>***</u>
18	262,143	0.1516	<u>1,087.2800</u>	<u>1,113.5300</u>	<u>***</u>
20	1,048,575	0.6011	<u>8,485.4500</u>	<u>8,862.9800</u>	<u>***</u>

Table 1: DS-Conditional-One model versus specialization matrix based method. Average computational times (ms). (*** denotes computations not completed within a feasible time).

The significant speed advantage offered by the proposed computational model over the specialization matrix based approach is evident from Table 1. For larger FoDs, the computational burden associated with the specialization matrix based approach becomes prohibitive because of its space complexity of $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$. For example, an FoD of size 20 would need 128 ($= 2^{20} \times 2^{20}/8$) GB of memory to represent the specialization matrix, if each matrix entry occupies only 1 bit.

With increasing FoD size, the computational time requirement of the DS-Conditional-One model is significantly less compared to what the specialization matrix based approach requires.

6. Concluding Remarks

This paper provides a general framework for computation of DST conditionals. The DS-Conditional-One model that we propose can also serve as a tool for visualization and further analysis of the conditional computation process. We believe that the algorithms we have developed constitute a significant step forward in harnessing the strengths of DST methods in practical applications.

The efficiency of these algorithms is mainly because of the significantly reduced number of operations that are executed. Computational complexity associated with conditional belief computation of an arbitrary proposition is $\mathcal{O}(2^{|\overline{A}|+|A\cap B|})$. This is a significant improvement over the $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$ complexity associated with the specialization matrix based approach. The DS-Conditional-One model also provides a significant advantage in terms of memory usage: it requires a $\mathcal{O}(2^{|\Theta|})$ space complexity versus $\mathcal{O}(2^{|\Theta|} \times 2^{|\Theta|})$ for the specialization matrix based approach.

Another advantage of the proposed approach is that it can be utilized for either the FH conditional or Dempster's conditional belief computations. An outcome of this research is a conditional computation library (in C++) which is available at ProFuSELab (2017). We expect that this library will be useful for practical application of DST methods.

Our current research work is focused on conditional computations on potentially dynamic FoDs (where the singletons may have to be removed or new singletons may have to be appended as operations are carried out). This would be of immense value for enhanced resource utilization. It also appears possible to further enhance the algorithms that we have developed via parallel computing optimizations because of the underlying matrix structure.

Acknowledgments

This work is based on research supported by the U.S. Office of Naval Research (ONR) via grant #N00014-10-1-0140 and the U.S. National Science Foundation (NSF) via grant #1343430.

References

T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley, Chichester, UK, May 2014.

A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.*, 38(2):325–339, Apr 1967.

A. P. Dempster. A generalization of Bayesian inference. *J. R. Stat. Soc. Ser. B*, 30(2):205–247, 1968.

T. Denœux. 40 years of Dempster–Shafer theory. *Int. J. Approx. Reason.*, 79(C):1–6, Dec 2016.

R. Fagin and J. Y. Halpern. A new approach to updating beliefs. In *Proc. 6th Conf. Uncertainty in Artificial Intelligence (UAI)*, pages 347–374, Cambridge, MA, Jul 1990.

R. Kennes. Computational aspects of the Möbius transformation of graphs. *IEEE Trans. Syst. Man. Cybern.*, 22(2):201–223, Mar./Apr. 1992.

F. Klawonn and P. Smets. The dynamic of belief in the transferable belief model and specialization-generalization matrices. In *Proc. 8th Conf. Uncertainty in Artificial Intelligence (UAI)*, pages 130–137, Stanford, CA, Jul 1992.

E. C. Kulasekere, K. Premaratne, D. A. Dewasurendra, M.-L. Shyu, and P. H. Bauer. Conditioning and updating evidence. *Int. J. Approx. Reason.*, 36(1):75–108, 2004.

H. T. Nguyen and P. Smets. On dynamics of cautious belief and conditional objects. *Int. J. Approx. Reason.*, 8(2):89–104, Feb 1993.

L. G. Polpitiya, K. Premaratne, M. N. Murthi, and D. Sarkar. A framework for efficient computation of belief theoretic operations. In *Proc. 19th Int. Conf. Information Fusion (FUSION)*, pages 1570–1577, Heidelberg, Germany, Jul 2016.

K. Premaratne, M. N. Murthi, J. Zhang, M. Scheutz, and P. H. Bauer. A Dempster-Shafer theoretic conditional approach to evidence updating for fusion of hard and soft data. In *Proc. 12th Int. Conf. Information Fusion (FUSION)*, pages 2122–2129, Seattle, WA, Jul 2009.

ProFuSELab. Conditional Computation Library, 2017. URL `https://profuselab.github.io/Conditional-Computation-Library/`.

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

G. Shafer. Perspectives on the theory and practice of belief functions. *Int. J. Approx. Reason.*, 4 (5-6):323–362, Oct 1990.

SIPTA. Software tools for imprecise probabilities, 2017. URL `http://www.sipta.org/index.php?id=sfw`.

P. Smets. Practical uses of belief functions. In *Proc. 15th Conf. Uncertainty in Artificial Intelligence (UAI)*, pages 612–621, Stockholm, Sweden, Jul 1999.

P. Smets. The application of the matrix calculus to belief functions. *Int. J. Approx. Reason.*, 31(1): 1–30, 2002.

H. M. Thoma. *Factorization of belief functions*. PhD thesis, Dept. Stat., Harvard Univ., Cambridge, MA, 1989.

T. L. Wickramarathne, K. Premaratne, M. N. Murthi, M. Scheutz, S. Kübler, and M. Pravia. Belief theoretic methods for soft and hard data fusion. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2388–2391, Prague, Czech Republic, May 2011.

T. L. Wickramarathne, K. Premaratne, and M. N. Murthi. Toward efficient computation of the Dempster–Shafer belief theoretic conditionals. *IEEE Trans. Cybern.*, 43(2):712–724, Apr 2013.

H. Xu and P. Smets. Reasoning in evidential networks with conditional belief functions. *Int. J. Approx. Reason.*, 14(2):155–185, 1996.

R. R. Yager and L. Liu, editors. *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Springer-Verlag, Berlin Heidelberg, 2008.