



HAL
open science

Total belief theorem and conditional belief functions

Jean Dezert, Albena Tchamova, Deqiang Han

► **To cite this version:**

Jean Dezert, Albena Tchamova, Deqiang Han. Total belief theorem and conditional belief functions. 2020. <hal-02468001>

HAL Id: hal-02468001

<https://hal.science/hal-02468001v1>

Preprint submitted on 5 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Total Belief Theorem and Conditional Belief Functions

Jean Dezert,^{1,*} Albena Tchamova^{2,†} Deqiang Han^{3,‡}

¹ *ONERA - The French Aerospace Lab,*

Chemin de la Hunière, F-91761 Palaiseau, France.

² *Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria.*

³ *Institute of Integrated Automation,*

Xi'an Jiaotong University, Xi'an, Shaanxi, China 710049.

In this paper new theoretical results for reasoning with belief functions are obtained and discussed. After a judicious decomposition of the set of focal elements of a belief function, we establish the Total Belief Theorem (TBT) which is the direct generalization of the Total Probability Theorem when working in the framework of belief functions. The TBT is also generalized for dealing with different frames of discernments thanks to Cartesian product space. From TBT, we can derive and define formally the expressions of conditional belief functions which are consistent with the bounds of imprecise conditional probability. This work provides a direct establishment and solid justification of Fagin-Halpern belief conditioning formulas. The well-known Bayes' Theorem of Probability Theory is then generalized in the framework of belief functions and we illustrate it with an example at the end of this paper.

1 INTRODUCTION

In this paper, we present new theoretical results for reasoning with belief functions (BF) introduced by Shafer in¹, known as Dempster-Shafer Theory (DST) in the literature. The first result is the establishment of the Total Belief Theorem (TBT) which can be interpreted as a generalization of the Total Probability Theorem (TPT) for the belief functions framework. TBT is essential for formally establishing conditional belief functions in a constructive manner whose expressions are consistent with original Dempster's idea (through eq. (4.8) in²), rediscovered independently and popularized by Fagin-Halpern in^{3,4}. TBT also allows us to present a new formulation of Generalized Bayes' Theorem (GBT).

*email : jean.dezert@onera.fr

†email : tchamova@bas.bg

‡email : deqhan@gmail.com

Several methods have been proposed in the literature to address the belief conditioning problem. They essentially can be separated in two different approaches: 1) Shafer's belief conditioning method based on Dempster's rule of combination¹, and 2) the belief conditioning method consistent with imprecise probability calculus bounds^{2,3,5-8} based on the lower and upper probability interpretation of belief functions.

Although Shafer's belief functions offer an appealing mathematical framework for modeling epistemic uncertainty, their use and the validity of the results obtained in the applications are very controversial both for uncertain information fusion as well as for belief conditioning mainly due to Shafer's choice of Dempster's rule of combination as a pillar for combining evidences represented by belief functions and for conditioning. These well known problems of DST have already been reported and discussed by many experts in the fields over the last decades, see for example⁹⁻²³. That is why in this paper we focus on the second approach of belief conditioning based on the lower and upper probability interpretation of BF.

It is worth noting that Smets in nineties²⁵ did propose a preliminary version of GBT to generalize Bayes' Theorem (BT) to belief functions but Smets' GBT is based on conditional embedding, conjunctive merging and Shafer's conditioning which make it quite complicate to apply and whose results have been cast in doubt in²⁶. Here we propose a simpler and direct constructive manner to derive a new version of GBT without need of extra assumptions of some underlying ad-hoc principles as done by Smets. Of course, we prove that our TBT and GBT presented in this work are fully consistent with classical TPT and BT as soon as the belief functions are restricted to Bayesian belief functions (i.e. classical probability measures).

This paper is organized as follows. After a brief recall of basics of belief functions in Section 2 and Total Probability Theorem in Section 3, we present probability conditioning and Bayes' theorem in Section 4 followed by classical Shafer's and Fagin-Halpern's belief conditioning methods respectively in Sections 5 and 6. In Section 7, we present the decomposition of the set of focal elements of any basic belief assignment that allows us to establish formally the Total Belief Theorem and its generalization on Cartesian product space. The Section 8 presents and justifies the new belief conditioning formulas drawn from TBT which are fully consistent with Fagin-Halpern conditioning formulas. Section 9 presents the generalization of Bayes' theorem in the framework of belief functions obtained from TBT. We illustrate our new theoretical results with a quite simple GBT example in Section 10 to show how to make derivations of GBT and to prove that Shafer's conditioning results are inconsistent with GBT. Section 11 concludes this paper.

2 BASICS OF BELIEF FUNCTIONS

Belief functions (BF) have been introduced by Shafer in¹ to model epistemic uncertainty based on preliminary works done by Dempster^{2,27}. Shafer's Theory of Belief Functions is also referred as Dempster-Shafer Theory (DST) in the literature. We assume that the answer¹ of the problem under concern belongs to a known (or given) finite discrete frame of discernment (FoD) $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, with $n > 1$, and

¹i.e. the solution, or the decision to take.

where all elements of Θ are exhaustive and exclusive². The set of all subsets of Θ (including empty set \emptyset , and Θ) is the power-set of Θ denoted by 2^Θ . The number of elements (i.e. the cardinality) of 2^Θ is $2^{|\Theta|}$. A basic belief assignment (BBA) associated with a given source of evidence is defined as the mapping $m(\cdot) : 2^\Theta \rightarrow [0, 1]$ satisfying the conditions $m(\emptyset) = 0$ and $\sum_{A \in 2^\Theta} m(A) = 1$. The quantity $m(A)$ is called the mass of A committed by the source of evidence. Belief and plausibility functions are respectively defined by

$$Bel(A) = \sum_{\substack{X \in 2^\Theta \\ X \subseteq A}} m(X) \quad (1)$$

$$Pl(A) = \sum_{\substack{X \in 2^\Theta \\ X \cap A \neq \emptyset}} m(X) = 1 - Bel(\bar{A}). \quad (2)$$

where³ $\bar{A} \triangleq \Theta - \{A\} = \{X | X \in \Theta \text{ and } X \not\subseteq A\}$, i.e. \bar{A} is the complement of A in Θ . The notation \triangleq means *equal by definition*. The width $Pl(A) - Bel(A)$ of the belief interval $[Bel(A), Pl(A)]$ is usually called the *uncertainty on A* committed by the source of evidence, and will be denoted⁴ by $U(A^*)$. It represents in fact the imprecision on the probability of A granted by the source of evidence, which provides the BBA $m(\cdot)$.

A focal element X of a BBA $m(\cdot)$ is an element of 2^Θ such that $m(X) > 0$. Note that the empty set \emptyset is not a focal element of a BBA because $m(\emptyset) = 0$ (close-world assumption of Shafer's model for the FoD). The set of all focal elements of $m(\cdot)$ is denoted

$$\mathcal{F}_\Theta(m) = \{X \subseteq \Theta | m(X) > 0\} = \{X \in 2^\Theta | m(X) > 0\} \quad (3)$$

Because $m(\emptyset) = 0$, one always has $1 \leq |\mathcal{F}_\Theta(m)| \leq 2^{|\Theta|} - 1$. The set of focal elements of $m(\cdot)$ included in a subset A of Θ is denoted

$$\mathcal{F}_A(m) \triangleq \{X \subseteq A \subseteq \Theta | m(X) > 0\} = \{X \in \mathcal{F}_\Theta(m) | X \cap A = X\} \quad (4)$$

Note that if $A \subseteq B \subseteq \Theta$, then $\mathcal{F}_A(m) \subseteq \mathcal{F}_B(m)$, and one always has⁵ $\mathcal{F}_A(m) \cap \mathcal{F}_B(m) = \mathcal{F}_{A \cap B}(m)$ for any subsets A and B of Θ , but $\mathcal{F}_{A \cup B}(m) \neq \mathcal{F}_A(m) \cup \mathcal{F}_B(m)$ in general⁶.

By definition, all elements of 2^Θ not in $\mathcal{F}_\Theta(m)$ have a zero mass value, and there-

²This is so-called Shafer's model of FoD²⁸.

³Here the minus symbol denotes the set difference operator^{29,30}.

⁴In the literature it is usually denoted by $U(A)$. Here we use a new notation $U(A^*)$ which is not anecdotic. This new notation reveals its importance for the consistency of notations used in formulas we give in this paper.

⁵Proof: $\mathcal{F}_A(m) \cap \mathcal{F}_B(m) = \{X \in \mathcal{F}_\Theta(m) | (X \cap A) \cap (X \cap B) = X\} = \{X \in \mathcal{F}_\Theta(m) | X \cap (A \cap B) = X\} = \mathcal{F}_{A \cap B}(m)$.

⁶For example, consider the focal elements given in the example of section 10. One has $A_1 \cup \bar{B} = \{\theta_1, \theta_3, \theta_4, \theta_7\} \cup \{\theta_1, \theta_2, \theta_3\} = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_7\}$ and therefore $\mathcal{F}_{A_1 \cup \bar{B}} = \{X_2, X_4, X_5, X_7\}$, but $\mathcal{F}_{A_1} = \{X_4\}$ and $\mathcal{F}_{\bar{B}} = \{X_5\}$, so that $\mathcal{F}_{A_1} \cup \mathcal{F}_{\bar{B}} = \{X_4, X_5\} \neq \mathcal{F}_{A_1 \cup \bar{B}}$.

fore the definition of $Bel(A)$ and $Pl(A)$ given in (1)–(2) can also be expressed⁷

$$Bel(A) = \sum_{\substack{X \in \mathcal{F}_\Theta(m) \\ X \subseteq A}} m(X) = \sum_{X \in \mathcal{F}_A(m)} m(X) \quad (5)$$

$$Pl(A) = \sum_{\substack{X \in \mathcal{F}_\Theta(m) \\ X \cap A \neq \emptyset}} m(X) = 1 - Bel(\bar{A}). \quad (6)$$

The set of focal elements $\mathcal{F}_\Theta(m)$ of the BBA $m(\cdot)$ can always be partitioned as $\{\mathcal{F}_A(m), \mathcal{F}_{\bar{A}}(m), \mathcal{F}_{A^*}(m)\}$ where

$$\mathcal{F}_{A^*}(m) \triangleq \mathcal{F}_\Theta(m) - \mathcal{F}_A(m) - \mathcal{F}_{\bar{A}}(m) \quad (7)$$

$$= \{X \in \mathcal{F}_\Theta(m) \mid X \cap A \neq \emptyset \text{ and } X \cap \bar{A} \neq \emptyset\} \quad (8)$$

represents the set of focal elements of $m(\cdot)$ which are not subsets of A and not subsets of $\bar{A} = \Theta - \{A\}$.

The uncertainty $U(A^*)$ can also be expressed directly as

$$U(A^*) = \sum_{X \in \mathcal{F}_{A^*}(m)} m(X) \quad (9)$$

It is worth noting that $U(\bar{A}^*) = Pl(\bar{A}) - Bel(\bar{A}) = (1 - Bel(A)) - (1 - Pl(A)) = Pl(A) - Bel(A) = U(A^*)$, or equivalently

$$U(\bar{A}^*) = \sum_{X \in \mathcal{F}_{\bar{A}^*}(m)} m(X) \quad (10)$$

where $\mathcal{F}_{\bar{A}^*}(m) = \mathcal{F}_\Theta(m) - \mathcal{F}_{\bar{A}}(m) - \mathcal{F}_A(m) = \mathcal{F}_{A^*}(m)$.

When all elements of $\mathcal{F}_\Theta(m)$ are only singletons, $m(\cdot)$ is called a *Bayesian BBA*¹ and its corresponding $Bel(\cdot)$ and $Pl(\cdot)$ functions are homogeneous to a same (subjective) probability measure $P(\cdot)$.

The class of belief functions can be characterized without explicitly referencing to a BBA, see Shafers' theorem in¹ page 39, with its proof on page 51. More precisely, a mapping $Bel(\cdot) : 2^\Theta \mapsto [0, 1]$ is a belief function if and only if $Bel(\emptyset) = 0$, $Bel(\Theta) = 1$ and for every positive integer n and every collection A_1, \dots, A_n of subsets of Θ

$$Bel(A_1 \cup \dots \cup A_n) \geq \sum_{\substack{I \subseteq \{1, \dots, n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right) \quad (11)$$

⁷More precisely, we should write $Bel(A) = 0 + \sum_{X \in \mathcal{F}_A(m)} m(X)$ to get a well defined value even there is no $X \in \mathcal{F}_\Theta(m)$ such that $X \subseteq A$. For notation convenience, this zero additional term (as well other zero terms in formulas (9)–(10), (41), etc) will be omitted in the sequel being understood that a *sum of non existing terms is always equal to zero*.

There is a one-to-one relationship between a BBA $m(\cdot)$ and its corresponding belief function $Bel(\cdot)$. The BBA $m(\cdot)$ that produces a given belief function is unique and is obtained for any $A \subseteq \Theta$ by the following Möbius inverse formula (see¹, p.39)

$$m(A) = \sum_{B \subseteq A \subseteq \Theta} (-1)^{|A-B|} Bel(B) \quad (12)$$

In DST framework, Shafer¹ did propose to combine $s \geq 2$ distinct sources of evidence represented by BBAs $m_1(\cdot), \dots, m_s(\cdot)$ over the same FoD with Dempster's rule (i.e. the normalized conjunctive rule). Discussions on the justification of Dempster's rule with examples can be found in²¹⁻²³.

3 TOTAL PROBABILITY THEOREM (TPT)

We recall briefly the Total Probability Theorem because we will present its extension in Belief function framework. In probability theory, the elements θ_i of the space Θ are experimental outcomes. The subsets of Θ are called events and the event $\{\theta_i\}$ consisting of the single element θ_i is an elementary event. The space Θ is called the *sure event* and the empty set \emptyset is the *impossible event*. We assign to each event A a number $P(A)$ in $[0, 1]$, called the probability of A , which satisfies the three Kolmogorov's conditions: 1) $P(\emptyset) = 0$; 2) $P(\Theta) = 1$; and 3) if $A \cap B = \{\emptyset\}$, then $P(A \cup B) = P(A) + P(B)$. These conditions are the axioms of the theory of probability^{30,31}. The fundamental Theorem of the probability theory is the following Total Probability Theorem (TPT), also called a the law of total probability, see³¹ and Theorem 1B of³².

Total Probability Theorem (TPT): Consider an event B and any partition⁸ $\{A_i, i = 1, \dots, k\}$ of the space Θ , then

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_k) \quad (13)$$

4 CONDITIONAL PROBABILITY AND BAYES' FORMULA

Starting from TPT formula (13) and assuming $P(B) > 0$, we get for any $i \in \{1, \dots, k\}$ after dividing each side of (13) by $P(B)$ and rearranging terms the equality

$$\frac{P(A_i \cap B)}{P(B)} = 1 - \sum_{\substack{j=1, \dots, k \\ j \neq i}} \frac{P(A_j \cap B)}{P(B)} = 1 - \frac{P(\bar{A}_i \cap B)}{P(B)} \quad (14)$$

which allows us to define the conditional probability $P(A_i|B)$ by⁹

$$P(A_i|B) \triangleq P(A_i \cap B)/P(B) \quad (15)$$

Similarly, by considering an event A_i of Θ and the partition $\{B, \bar{B}\}$ of Θ , the TPT formula $P(A_i) = P(A_i \cap B) + P(A_i \cap \bar{B})$ applies, and by dividing it by $P(A_i)$ (assuming $P(A_i) > 0$), we get

$$P(A_i \cap B)/P(A_i) = 1 - P(A_i \cap \bar{B})/P(A_i) \quad (16)$$

⁸A partition of a set Θ is a collection of mutually exclusive subsets of Θ whose union equals Θ .

⁹In probability theory, the notation $P(A_i, B) \equiv P(A_i \cap B)$ is also used to represent the probability of the joint occurrence (intersection) of events A_i and B .

which allows us to define also the conditional probability $P(B|A_i)$ by

$$P(B|A_i) \triangleq P(A_i \cap B)/P(A_i) \quad (17)$$

From (15) and (17), one deduces the equality

$$P(A_i \cap B) = P(A_i|B)P(B) = P(B|A_i)P(A_i) \quad (18)$$

From (18) and assuming $P(B) > 0$ we get $P(A_i|B) = P(B|A_i)P(A_i)/P(B)$, and assuming $P(A_i) > 0$ we get $P(B|A_i) = P(A_i|B)P(B)/P(A_i)$.

Using TPT formula (13) and noting that $P(A_i \cap B) = P(B|A_i)P(A_i)$, we get

$$P(B) = \sum_{i=1}^k P(B|A_i)P(A_i) \quad (19)$$

Substituting (19) in $P(A_i|B) = P(B|A_i)P(A_i)/P(B)$, we get the well-known Bayes' Theorem formula (BTF)

$$P(A_i|B) = P(B|A_i)P(A_i) / \sum_{i=1}^k P(B|A_i)P(A_i) \quad (20)$$

It can be easily verified that the conditional probability defined by (15) verifies the three axioms of the Theory of probability³¹: 1) $P(\emptyset|B) = 0$, 2) $P(\Theta|B) = 1$ and 3) if $A_1 \cap A_2 = \emptyset$, then $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$.

In the previous presentation, A_i ($i = 1, \dots, k$) and B are events (subsets) of the same space Θ . How to proceed to compute $P(A_i|B)$ if the events A_i ($i = 1, \dots, k$) and B are subsets of different spaces, say if $A_i \subseteq \Theta_1 = \{x_1, \dots, x_m\} = \{x_p, p = 1, 2, \dots, m\}$ ($i = 1, \dots, k$), and if $B \subseteq \Theta_2 = \{y_1, \dots, y_n\} = \{y_q, q = 1, 2, \dots, n\}$ with $\Theta_1 \neq \Theta_2$? Such situation corresponds to a so-called combined experiment³¹. In fact, one can prove that similar conditioning formulas can also be established. For this, we need to work with the Cartesian product space $\Theta \triangleq \Theta_1 \times \Theta_2$ whose elementary elements are all the ordered pairs (x_p, y_q) with $x_p \in \Theta_1$ and $y_q \in \Theta_2$. The two experiments are viewed as a single combined one whose outcomes are pairs (x_p, y_q) . In this space $\Theta = \Theta_1 \times \Theta_2$, x_p is not an elementary element but a subset of n elements of Θ , i.e. $\{x_p\} = \{(x_p, y_1), \dots, (x_p, y_n)\}$. Similarly, y_q is not an elementary element but a subset of m elements of Θ , i.e. $\{y_q\} = \{(x_1, y_q), \dots, (x_m, y_q)\}$. If $A_i \subseteq \Theta_1$ and $B \subseteq \Theta_2$, then $A_i \times B = \{(x_p, y_q) | x_p \in A_i; y_q \in B\} \subseteq \Theta$. If one forms $A_i \times \Theta_2$ and $\Theta_1 \times B$ one sees that $A_i \times B = (A_i \times \Theta_2) \cap (\Theta_1 \times B) = (\Theta_1 \times B) \cap (A_i \times \Theta_2)$. Because the event $A_i \times \Theta_2$ occurs in the combined experiment if the event A_i of the experiment 1 occurs no matter what the outcome of experiment 2, one has $P(A_i \times \Theta_2) = P_1(A_i)$ where $P_1(A_i)$ is the probability of event A_i in the experiment 1. Similarly, the event $\Theta_1 \times B$ occurs if B occurs in experiment 2 no matter what the outcome of experiment 1, so that $P(\Theta_1 \times B) = P_2(B)$ where $P_2(B)$ is the probability of event B in the experiment 2. One considers a partition $\{A_1, A_2, \dots, A_k\}$ of Θ_1 and a subset (event) $B \subseteq \Theta_2$. Based on set theory and property of Cartesian product, one has

$$\begin{aligned} \Theta_1 \times B &= (\Theta_1 \times B) \cap (\Theta_1 \times \Theta_2) = (\Theta_1 \times B) \cap ((A_1 \cup A_2 \cup \dots \cup A_k) \times \Theta_2) \\ &= (\Theta_1 \times B) \cap ((A_1 \times \Theta_2) \cup \dots \cup (A_k \times \Theta_2)) = \bigcup_i ((\Theta_1 \times B) \cap (A_i \times \Theta_2)) \end{aligned}$$

The elements $A_i \times \Theta_2$, $i = 1, \dots, k$ being disjoint¹⁰, one has the following TPT formula

$$\begin{aligned} P(\Theta_1 \times B) &= P(\bigcup_i ((\Theta_1 \times B) \cap (A_i \times \Theta_2))) \\ &= P((\Theta_1 \times B) \cap (A_1 \times \Theta_2)) + \dots + P((\Theta_1 \times B) \cap (A_k \times \Theta_2)) \quad (21) \end{aligned}$$

¹⁰because A_i are disjoint since $\{A_1, \dots, A_k\}$ is a partition of Θ_1 .

After dividing each side of formula (21) by $P(\Theta_1 \times B)$ (assumed positive) and rearranging terms, we get

$$\frac{P((A_i \times \Theta_2) \cap (\Theta_1 \times B))}{P(\Theta_1 \times B)} = 1 - \sum_{\substack{j=1, \dots, k \\ j \neq i}} \frac{P((A_j \times \Theta_2) \cap (\Theta_1 \times B))}{P(\Theta_1 \times B)} \quad (22)$$

Formula (22) suggests naturally to define the conditional probability $P(A_i \times \Theta_2 | \Theta_1 \times B)$ by

$$P(A_i \times \Theta_2 | \Theta_1 \times B) \triangleq P(A_i \times B) / P(\Theta_1 \times B) \quad (23)$$

Using same reasoning as before and working on Cartesian product space $\Theta = \Theta_1 \times \Theta_2$, one can also prove¹¹ that if $P(A_i \times \Theta_2) > 0$ one can define

$$P(\Theta_1 \times B | A_i \times \Theta_2) = P(A_i \times B) / P(A_i \times \Theta_2) \quad (24)$$

From (23) and (24), one deduces the equality

$$P(A_i \times \Theta_2 | \Theta_1 \times B) P(\Theta_1 \times B) = P(\Theta_1 \times B | A_i \times \Theta_2) P(A_i \times \Theta_2) \quad (25)$$

From equality (25) and assuming $P(\Theta_1 \times B) > 0$, we get

$$P(A_i \times \Theta_2 | \Theta_1 \times B) = P(\Theta_1 \times B | A_i \times \Theta_2) P(A_i \times \Theta_2) / P(\Theta_1 \times B) \quad (26)$$

From equality (25) and assuming $P(A_i \times \Theta_2) > 0$ we get

$$P(\Theta_1 \times B | A_i \times \Theta_2) = P(A_i \times \Theta_2 | \Theta_1 \times B) P(\Theta_1 \times B) / P(A_i \times \Theta_2) \quad (27)$$

Using TPT formula (21) and formula(24), we get $P(\Theta_1 \times B) = \sum_{i=1}^k P(\Theta_1 \times B | A_i \times \Theta_2) P(A_i \times \Theta_2)$. Putting this expression in (26), we obtain the Bayes' Theorem formula (BTF) when $A \subseteq \Theta_1$ and $B \subseteq \Theta_2$ and $\Theta_1 \neq \Theta_2$, which is written as

$$P(A_i \times \Theta_2 | \Theta_1 \times B) = \frac{P(\Theta_1 \times B | A_i \times \Theta_2) P(A_i \times \Theta_2)}{\sum_{i=1}^k P(\Theta_1 \times B | A_i \times \Theta_2) P(A_i \times \Theta_2)} \quad (28)$$

For notation convenience (and notation abuse), we can use classical formulas when working with different sets of experimental outcomes Θ_1 and Θ_2 with keeping in mind that in this case A_i must be understood as $A_i \times \Theta_2$ and B as $\Theta_1 \times B$.

5 SHAFER'S CONDITIONING

In the belief functions framework, Shafer did propose formulas to calculate conditional belief functions $Bel(A|B)$ and $Pl(A|B)$. Shafer's formulas have been obtained from the conditional BBA $m(\cdot|B)$ obtained from Dempster's rule of combination of the original BBA $m(\cdot)$ with the BBA $m_B(B) = 1$ focused on B under the condition that $Bel(\bar{B}) < 1$, or equivalently¹² under the condition that $Pl(B) > 0$. Shafer's conditioning formulas for belief and plausibility functions were established by Shafer in Theorem 3.6 p. 66 of¹. For $A, B \subseteq \Theta$ with $Pl(B) > 0$, $Bel(A|B)$ and $Pl(A|B)$ are given by

$$Bel(A|B) = (Bel(A \cup \bar{B}) - Bel(\bar{B})) / (1 - Bel(\bar{B})) \quad (29)$$

$$Pl(A|B) = Pl(A \cap B) / Pl(B) \quad (30)$$

¹¹The proof is left to the reader due to space limitation restraint.

¹²Indeed, if $Bel(\bar{B}) < 1$ then $Pl(B) = 1 - Bel(\bar{B})$ is greater than zero.

The expression (29) of $Bel(A|B)$ is equivalent to

$$Bel(A|B) = (Pl(B) - Pl(B \cap \bar{A}))/Pl(B) \quad (31)$$

because one has always (from definition of belief functions) $Pl(B) = 1 - Bel(\bar{B})$, and the numerator of (29) can be written as

$$\begin{aligned} Bel(A \cup \bar{B}) - Bel(\bar{B}) &= (1 - Bel(\bar{B})) - (1 - Bel(A \cup \bar{B})) \\ &= Pl(B) - Pl(\overline{A \cup \bar{B}}) = Pl(B) - Pl(B \cap \bar{A}) \end{aligned}$$

Using (29)–(30) and taking $A = \emptyset$, we get $Bel(\emptyset|B) = Pl(\emptyset|B) = 0$, and taking $A = \Theta$ we get $Bel(\Theta|B) = Pl(\Theta|B) = 1$. Also in taking $B = \Theta$ we get $Bel(A|\Theta) = Bel(A)$ and $Pl(A|\Theta) = Pl(A)$. Note that taking $B = A$ in (30)–(31), we obtain $Bel(A|A) = Pl(A|A) = 1$ which fits with the common sense.

In reversing the roles played by A and B and switching the notations in previous expressions, the following formulas also hold (assuming $Pl(A) > 0$)

$$Bel(B|A) = (Pl(A) - Pl(A \cap \bar{B}))/Pl(A) \quad (32)$$

$$Pl(B|A) = Pl(B \cap A)Pl(A) \quad (33)$$

From (30) and (33), one deduces $Pl(A \cap B) = Pl(A|B)Pl(B) = Pl(B|A)Pl(A)$. Hence, the following formula applies for conditional plausibilities when $Pl(B) > 0$

$$Pl(A|B) = Pl(B|A)Pl(A)/Pl(B) \quad (34)$$

Note that this formula for conditional plausibilities is similar to the expression for conditional probabilities given in (15) when replacing plausibilities by probabilities.

The main drawback of Shafer's conditioning is its incompatibility with probability calculus when working with imprecise probabilities. More precisely, the bounds of belief interval defined by $[Bel(A|B), Pl(A|B)]$ obtained by (29)–(30) are in general¹³ incompatible with lower and upper bounds of the conditional probability $P(A|B)$. This problem makes Shafer's conditioning very disputable and cast serious doubts on pertinence (validity) of Shafer's conditioning results when used in the applications, which is a direct consequence of the validity of Dempster's rule reported in^{3,9–23,33,34}. Shafer's conditioning problem has already been reported and addressed by several authors^{3,6,7,14,24} in the past with some examples. To easily show this incompatibility of Shafer's conditioning with probability calculus we present briefly the famous Ellsberg's urn example³⁵.

Example 1 (Ellsberg's urn): We consider an urn with red (R) balls, black (B) and yellow (Y) balls. The a priori information one has on the repartition of the balls in the urn is the following: 1/3 of balls are red balls and 2/3 of balls are black and yellow balls. We don't know precisely the percentage of black balls, nor the percentage of yellow balls. So the a priori information about the chance to pick a ball in the urn can be represented by a (parametric) probability mass function $P(\cdot)$ with $P(R) = 1/3$, $P(B) = 2/3 - x$, $P(Y) = x$, where x is an unknown number/parameter in $[0, 2/3]$, $P(R)$ is the probability to pick at random a red ball in the urn, $P(B)$ is the probability to pick at random a black ball in the urn, and $P(Y)$ is the probability to pick at random a yellow ball in the urn. Of course because x is unknown but bounded, $P(B)$ and $P(Y)$ are unknown but their bounds are known. In fact, this problem can be seen as a problem of imprecise probabilities where $P(R) \in [1/3, 1/3]$, $P(B) \in [0, 2/3]$, $P(Y) \in [0, 2/3]$ and

¹³but if the BBA $m(\cdot)$ is Bayesian.

with the constraint $P(R) + P(B) + P(Y) = 1$. Now let's suppose that someone picks a ball at random in the urn and tell us that the color of the ball is not black, i.e. the event $\bar{B} = R \cup Y$ has occurred. How do we must revise (update) our prior probabilities with this new information? The correct answer to this question is obtained by computing the conditional probabilities $P(R|\bar{B})$, $P(B|\bar{B})$ and $P(Y|\bar{B})$ and by analyzing their bounds. This is done as follows using the fact that $P(\bar{B}) = P(R \cup Y) = P(R) + P(Y) - P(R \cap Y) = P(R) + P(Y) = (1/3) + x$. Indeed, $P(R \cap Y) = 0$ because the events R and Y are mutually exclusive. So, we get

$$\begin{aligned} P(R|\bar{B}) &= P(R \cap (R \cup Y))/P(R \cup Y) = P(R)/((1/3) + x) = (1/3)/((1/3) + x) \\ P(B|\bar{B}) &= P(B \cap (R \cup Y))/P(R \cup Y) = P(\emptyset)/((1/3) + x) = 0/((1/3) + x) \\ P(Y|\bar{B}) &= P(Y \cap (R \cup Y))/P(R \cup Y) = P(Y)/((1/3) + x) = x/((1/3) + x) \end{aligned}$$

If $x = 0$, then $P(R|\bar{B}) = 1$ and $P(Y|\bar{B}) = 0$. If $x = 2/3$, then $P(R|\bar{B}) = 1/3$ and $P(Y|\bar{B}) = 2/3$. Therefore after conditioning by $\bar{B} = R \cup Y$ we get as bounds of conditional probabilities values the following intervals $P(R|\bar{B}) \in [1/3, 1]$, $P(B|\bar{B}) \in [0, 0]$, $P(Y|\bar{B}) \in [0, 2/3]$ with the constraint $P(R|\bar{B}) + P(B|\bar{B}) + P(Y|\bar{B}) = 1$.

Let's examine what we get using Shafer's conditioning approach. For this, the problem is modeled directly in the belief function framework using the *a priori* BBA $m(\cdot)$ defined on the FoD $\Theta = \{R, B, Y\}$ with $m(R) = 1/3$, $m(B \cup Y) = 2/3$ which corresponds to the following a priori belief intervals $[Bel(R), Pl(R)] = [1/3, 1/3]$, $[Bel(B), Pl(B)] = [0, 2/3]$, $[Bel(Y), Pl(Y)] = [0, 2/3]$.

With Shafer's conditioning formulas and noting that $Pl(R) = 1/3$, $Pl(B) = 2/3$, $Pl(Y) = 2/3$, and $Pl(R \cup Y) = 1$, we get incompatible results with the real bounds of conditional probabilities because

$$\begin{aligned} [Bel(R|\bar{B}), Pl(R|\bar{B})] &= [1/3, 1/3] \text{ (by Shafer)} \neq [1/3, 1] \text{ (correct bounds)} \\ [Bel(B|\bar{B}), Pl(B|\bar{B})] &= [0, 0] \text{ (by Shafer)} = [0, 0] \text{ (correct bounds)} \\ [Bel(Y|\bar{B}), Pl(Y|\bar{B})] &= [2/3, 2/3] \text{ (by Shafer)} \neq [0, 2/3] \text{ (correct bounds)} \end{aligned}$$

To overcome this problem, Fagin and Halpern did propose a more efficient conditioning approach which is, by construction, always consistent with conditional probability bounds. It is presented in the next section.

6 FAGIN-HALPERN CONDITIONING

Fagin and Halpern (FH) proposed in^{3,4} to define the conditional belief as the lower envelope (i.e. the infimum) of a family of conditional probability functions to make belief conditioning consistent with imprecise conditional probability calculus. Assuming $Bel(B) > 0$, Fagin and Halpern proposed the following conditional formulas (FH formulas for short)

$$Bel(A|B) = Bel(A \cap B)/(Bel(A \cap B) + Pl(\bar{A} \cap B)) \quad (35)$$

$$Pl(A|B) = Pl(A \cap B)/(Pl(A \cap B) + Bel(\bar{A} \cap B)) \quad (36)$$

Fagin and Halpern did prove in³ with long derivations and great effort that the conditional belief $Bel(A|B)$ given by (35) satisfies also the three conditions for defining a true belief function according to Shafer's theorem in¹, p. 39. Therefore, the formula (35) is also a good candidate and serious alternative for conditioning belief functions. However, it is quite mysterious how Fagin and Halpern did obtain (construct) these close-form expressions. According to the authors, these expressions were rather established from a very good intuition. A better justification has been given by Sundberg and Wagner in⁷ (p. 268) but it is still not so clear in our opinion. In

this paper, we justify clearly and directly the establishment of FH formulas from the simple and direct consequence of the Total Belief Theorem (TBT) which is one of the main contributions of our work. From FH conditioning formulas (35)-(36), we can verify that the common sense results are also obtained, that is $Bel(\emptyset|B) = Pl(\emptyset|B) = 0$, $Bel(\Theta|B) = Pl(\Theta|B) = 1$, $Bel(A|\Theta) = Bel(A)$, $Pl(A|\Theta) = Pl(A)$, and $Bel(A|A) = Pl(A|A) = 1$.

FH conditioning formulas are consistent with Bayes conditioning formulas when the underlying BBA $m(\cdot)$ is Bayesian. Indeed if $m(\cdot)$ is Bayesian, then $Pl(A \cap B) = Bel(A \cap B) = P(A \cap B)$, $Pl(\bar{A} \cap B) = Bel(\bar{A} \cap B) = P(\bar{A} \cap B)$ and $Pl(\bar{B} \cap A) = Bel(\bar{B} \cap A) = P(\bar{B} \cap A)$ so that the FH formulas become equivalent to $Bel(A|B) = P(A \cap B)/(P(A \cap B) + P(\bar{A} \cap B))$ and $Pl(A|B) = P(A \cap B)/(P(A \cap B) + P(\bar{A} \cap B))$. Thanks to total probability theorem (TPT) formula (13), the denominator involved in these formula is $P(A \cap B) + P(\bar{A} \cap B) = P(B)$, therefore $Bel(A|B) = Pl(A|B) = P(A \cap B)/P(B) = P(A|B)$.

Similarly, one can also easily verify that $Bel(B|A) = Pl(B|A) = P(A \cap B)/P(A) = P(B|A)$. The advantage of FH conditioning is its complete compatibility with the conditional probability calculus⁷. Let us show what provides FH conditioning in the previous Ellsberg's urn example.

Ellsberg's urn example revisited: Let's see the result obtained by formulas (62) and (64) for Ellsberg's urn example. Applying formulas (62) and (64) with the conditioning event $\bar{B} = R \cup Y$ we obtain

$$\begin{aligned} Bel(R|\bar{B}) &= \frac{Bel(R \cap (R \cup Y))}{Bel(R \cap (R \cup Y)) + Pl((B \cup Y) \cap (R \cup Y))} = \frac{1/3}{(1/3) + (2/3)} = 1/3 \\ Pl(R|\bar{B}) &= \frac{Pl(R \cap (R \cup Y))}{Bel((B \cup Y) \cap (R \cup Y)) + Pl(R \cap (R \cup Y))} = \frac{1/3}{0 + (1/3)} = 1 \\ Bel(B|\bar{B}) &= \frac{Bel(B \cap (R \cup Y))}{Bel(B \cap (R \cup Y)) + Pl((R \cup Y) \cap (R \cup Y))} = \frac{0}{0 + 1} = 0 \\ Pl(B|\bar{B}) &= \frac{Pl(B \cap (R \cup Y))}{Bel((R \cup Y) \cap (R \cup Y)) + Pl(B \cap (R \cup Y))} = \frac{0}{(1/3) + 0} = 0 \\ Bel(Y|\bar{B}) &= \frac{Bel(Y \cap (R \cup Y))}{Bel(Y \cap (R \cup Y)) + Pl((R \cup B) \cap (R \cup Y))} = \frac{0}{0 + (1/3)} = 0 \\ Pl(Y|\bar{B}) &= \frac{Pl(Y \cap (R \cup Y))}{Bel((R \cup B) \cap (R \cup Y)) + Pl(Y \cap (R \cup Y))} = \frac{2/3}{(1/3) + (2/3)} = 2/3 \end{aligned}$$

Hence with FH conditioning formulas, we get the correct conditional probability bounds

$$[Bel(R|\bar{B}), Pl(R|\bar{B})] = [1/3, 1] \text{ (by Fagin-Halpern)} = [1/3, 1] \text{ (correct bounds)}$$

$$[Bel(B|\bar{B}), Pl(B|\bar{B})] = [0, 0] \text{ (by Fagin-Halpern)} = [0, 0] \text{ (correct bounds)}$$

$$[Bel(Y|\bar{B}), Pl(Y|\bar{B})] = [0, 2/3] \text{ (by Fagin-Halpern)} = [0, 2/3] \text{ (correct bounds)}$$

We can also verify that $Bel(\emptyset|\bar{B}) = 0$, $Bel(R \cup B|\bar{B}) = 1/3$, $Bel(R \cup Y|\bar{B}) = 1$, $Bel(B \cup Y|\bar{B}) = 0$ and $Bel(R \cup B \cup Y|\bar{B}) = 1$. Applying Möbius inverse formula (12) with this conditional belief function $Bel(\cdot|\bar{B})$, we get the conditional mass of belief given by $m(R|\bar{B}) = 1/3$ and $m(R \cup Y|\bar{B}) = 2/3$ and all other mass values are equal to zero, whereas with Shafer's approach based on Dempster's rule of combination we get $m(R|\bar{B}) = 1/3$ and $m(Y|\bar{B}) = 2/3$. We see the difference between Shafer's and FH conditioning approaches. With Shafer's conditioning approach, because $(B \cup Y) \cap (R \cup Y) \neq \emptyset$ the mass $m(B \cup Y) = 2/3$ is entirely transferred (optimistically) to the most specific focal element Y included in $\bar{B} = R \cup Y$. With the FH conditioning method the mass $m(B \cup Y) = 2/3$ is entirely transferred (pessimistically, or cautiously) to the least specific focal element $R \cup Y$ included in $\bar{B} = R \cup Y$.

7 TOTAL BELIEF THEOREM (TBT)

In this section, we extend TPT theorem to belief and plausibility functions and we establish the Total Belief Theorem (TBT). Before this, we need to explain how the set of focal elements of a given BBA $m(\cdot)$ must be decomposed because it is the basis of the establishment of TBT.

7.1 Decomposition of the set of focal elements $\mathcal{F}_\Theta(m)$

Let us consider a FoD $\Theta = \{\theta_1, \dots, \theta_{|\Theta|}\}$ with $|\Theta| > 1$ elements, and a BBA $m(\cdot)$ defined on 2^Θ with a given set of focal elements $\mathcal{F}_\Theta(m)$. Consider any partition $\{A_1, A_2, \dots, A_k\}$ of the FoD Θ , then one can always decompose $\mathcal{F}_\Theta(m)$ as the union of following subsets

$$\mathcal{F}_\Theta(m) = \mathcal{F}_{A_1}(m) \cup \dots \cup \mathcal{F}_{A_k}(m) \cup \mathcal{F}_{A^*}(m) \quad (37)$$

where $\mathcal{F}_{A_i}(m)$ ($i = 1, \dots, k$) is the set of focal elements of $m(\cdot)$ included in A_i , and $\mathcal{F}_{A^*}(m)$ is the set of focal elements of $m(\cdot)$ which are not included in A_i , $i = 1, \dots, k$. We use the notation A^* for representing the entity characterized by the focal set $\mathcal{F}_{A^*}(m)$ mathematically defined by

$$\mathcal{F}_{A^*}(m) \triangleq \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \dots - \mathcal{F}_{A_k}(m) \quad (38)$$

The entity A^* has in general no explicit form and it is used only for notation convenience to make presentation of formulas more concise in the sequel. Because A_i for $i = 1, \dots, k$ are mutually exclusive (disjoint), the sets $\mathcal{F}_{A_i}(m)$ are also mutually exclusive and therefore $\cap_{i=1, \dots, k} (\mathcal{F}_\Theta(m) - \mathcal{F}_{A_i}(m)) = \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \dots - \mathcal{F}_{A_k}(m)$ because all possible intersections of focal sets including $\mathcal{F}_{A_i}(m) \cap \mathcal{F}_{A_j}(m)$ for $j \neq i$ equal the empty set. Hence $\mathcal{F}_{A^*}(m)$ can also be expressed as

$$\mathcal{F}_{A^*}(m) = \cap_{i=1, \dots, k} \bar{\mathcal{F}}_{A_i}(m) \quad (39)$$

where $\bar{\mathcal{F}}_{A_i}(m) \triangleq \mathcal{F}_\Theta(m) - \mathcal{F}_{A_i}(m) = \mathcal{F}_{\bar{A}_i}(m) + \mathcal{F}_{A_i^*}(m)$ represents the set of focal elements of $m(\cdot)$ which are not subsets of A_i .

Example 2: Consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and a BBA $m(\cdot)$ defined on 2^Θ , with set of focal elements $\mathcal{F}_\Theta(m) = \{X_1, X_2, \dots, X_8\}$ with $X_1 = \theta_1$, $X_2 = \theta_1 \cup \theta_2$, $X_3 = \theta_2 \cup \theta_3$, $X_4 = \theta_3 \cup \theta_4$, $X_5 = \theta_4$. Consider the partition $\{A_1, A_2, A_3\}$ of Θ with $A_1 = \{\theta_1, \theta_2\}$, $A_2 = \{\theta_3, \theta_4\}$ and $A_3 = \{\theta_5\}$. In this example, one has

$$\begin{aligned} \mathcal{F}_{A_1}(m) &= \{X_1, X_2\} = \{\theta_1, \theta_1 \cup \theta_2\} \\ \mathcal{F}_{A_2}(m) &= \{X_4, X_5\} = \{\theta_3 \cup \theta_4, \theta_4\} \\ \mathcal{F}_{A_3}(m) &= \{X_8\} = \{\theta_5\} \\ \mathcal{F}_{A^*}(m) &= \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \mathcal{F}_{A_2}(m) - \mathcal{F}_{A_3}(m) = \{X_3, X_6, X_7\} \\ &= \{\theta_2 \cup \theta_3, \theta_4 \cup \theta_5, \theta_1 \cup \theta_3 \cup \theta_5\} \end{aligned}$$

One sees that

$$\begin{aligned} \bar{\mathcal{F}}_{A_1}(m) &= \mathcal{F}_\Theta(m) - \{X_1, X_2\} = \{X_3, X_4, X_5, X_6, X_7, X_8\} \\ \bar{\mathcal{F}}_{A_2}(m) &= \mathcal{F}_\Theta(m) - \{X_4, X_5\} = \{X_1, X_2, X_3, X_6, X_7, X_8\} \\ \bar{\mathcal{F}}_{A_3}(m) &= \mathcal{F}_\Theta(m) - \{X_8\} = \{X_1, X_2, X_3, X_4, X_5, X_6, X_7\} \end{aligned}$$

and applying (39), we get

$$\bar{\mathcal{F}}_{A_1}(m) \cap \bar{\mathcal{F}}_{A_2}(m) \cap \bar{\mathcal{F}}_{A_3}(m) = \{X_3, X_6, X_7\} = \mathcal{F}_{A^*}(m)$$

Example 3: Consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and a BBA $m(\cdot)$ defined on 2^Θ , with the degenerate set of focal elements with only one focal element as follows $\mathcal{F}_\Theta(m) = \{X_1 = \Theta\}$ corresponding to the vacuous BBA. Consider the partition $\{A_1, A_2, A_3\}$ of Θ where $A_1 \triangleq \{x_3, x_5\}$, $A_2 \triangleq \{x_2\}$ and $A_3 \triangleq \{x_1, x_4\}$. Then, we get $\mathcal{F}_{A_1}(m) = \emptyset$, $\mathcal{F}_{A_2}(m) = \emptyset$, $\mathcal{F}_{A_3}(m) = \emptyset$ and $\mathcal{F}_{A^*}(m) = \{X_1\} - \emptyset - \emptyset - \emptyset = \Theta$. Note that, $\bar{\mathcal{F}}_{A_1}(m) = \bar{\mathcal{F}}_{A_2}(m) = \bar{\mathcal{F}}_{A_3}(m) = \Theta$, and therefore $\bar{\mathcal{F}}_{A_1}(m) \cap \bar{\mathcal{F}}_{A_2}(m) \cap \bar{\mathcal{F}}_{A_3}(m) = \Theta = \mathcal{F}_{A^*}(m)$, and of course $\mathcal{F}_\Theta(m) = \mathcal{F}_{A_1}(m) \cup \mathcal{F}_{A_2}(m) \cup \mathcal{F}_{A_3}(m) \cup \mathcal{F}_{A^*}(m) = \emptyset \cup \emptyset \cup \emptyset \cup \Theta = \Theta$.

7.2 Total Belief Theorem (TBT)

Based on the previous decomposition of the set of focal elements $\mathcal{F}_\Theta(m)$ according to any given partition $\{A_1, \dots, A_k\}$ of the FoD Θ , the following Total Belief Theorem (TBT) is established.

Total Belief Theorem (TBT): Let's consider a frame of discernment Θ with $|\Theta| \geq 2$ elements and a BBA $m(\cdot)$ defined on 2^Θ with the set of focal elements $\mathcal{F}_\Theta(m)$. For any chosen partition $\{A_1, \dots, A_k\}$ of Θ and for any $B \subseteq \Theta$, one has

$$Bel(B) = \sum_{i=1, \dots, k} Bel(A_i \cap B) + U(A^* \cap B) \quad (40)$$

where $\mathcal{F}_{A^*}(m) \triangleq \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \dots - \mathcal{F}_{A_k}(m)$ and

$$U(A^* \cap B) \triangleq \sum_{X \in \mathcal{F}_{A^*}(m) | X \in \mathcal{F}_B(m)} m(X). \quad (41)$$

Proof of TBT: See appendix.

A^* is a shorthand notation for the entity associated to the set of focal elements $\mathcal{F}_{A^*}(m)$ of the BBA $m(\cdot)$ involved in the summation (41) of $U(A^* \cap B)$. From the formula (41), one sees that $U(A^* \cap B) \in [0, 1]$. Note that if $B = \Theta$ and if the FoD Θ is simply partitioned as $\{A \triangleq A_1, \bar{A} \triangleq A_2\}$, then $U(A^* \cap B) = U(A^* \cap \Theta) = U(A^*) = Pl(A) - Bel(A) = Pl(\bar{A}) - Bel(\bar{A})$.

If one applies TBT with $B = \Theta$, we get for any chosen partition $\{A_1, \dots, A_k\}$ of Θ

$$\sum_{i=1, \dots, k} Bel(A_i) + U(A^*) = 1 \quad (42)$$

where $U(A^*) \triangleq \sum_{X \in \mathcal{F}_{A^*}(m)} m(X)$. This equality corresponds to TPT if $U(A^*) = 0$ (i.e. there is no uncertainty on the value of probabilities of $A_i, i = 1, \dots, k$).

Corollary of TBT: If $m(\cdot)$ is Bayesian, then TBT is consistent with the Total Probability Theorem (TPT).

Proof: See appendix.

From TBT one can establish the following (not so elegant) Total Plausibility Theorem (TPIT).

Total Plausibility Theorem (TPIT): For any BBA $m(\cdot) : 2^\Theta \mapsto [0, 1]$, and for any partition $\{A_1, \dots, A_k\}$ of Θ , one has for any $B \subseteq \Theta$

$$Pl(B) = \sum_{i=1, \dots, k} Pl(\bar{A}_i \cup B) + 1 - k - U(A^* \cap \bar{B}) \quad (43)$$

Proof: See appendix.

Example 4: Consider the FoD $\Theta = \{\theta_1, \dots, \theta_7\}$ and the set of focal elements $\mathcal{F}_\Theta(m) = \{X_1, X_2, \dots, X_9\}$ of a BBA $m(\cdot)$ defined over 2^Θ given in Table I.

Focal element X	BBA $m(X)$
$X_1 = \theta_2 \cup \theta_3 \cup \theta_4 \cup \theta_5 \cup \theta_7$	$m(X_1) = 0.01$
$X_2 = \theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4$	$m(X_2) = 0.02$
$X_3 = \theta_3 \cup \theta_5 \cup \theta_6$	$m(X_3) = 0.03$
$X_4 = \theta_4 \cup \theta_7$	$m(X_4) = 0.04$
$X_5 = \theta_2$	$m(X_5) = 0.20$
$X_6 = \theta_6 \cup \theta_7$	$m(X_6) = 0.30$
$X_7 = \theta_2 \cup \theta_3 \cup \theta_7$	$m(X_7) = 0.20$
$X_8 = \theta_1 \cup \theta_4 \cup \theta_6$	$m(X_8) = 0.15$
$X_9 = \theta_6$	$m(X_9) = 0.05$

Let's consider the partition $\{A_1, A_2, A_3\}$ of Θ with $A_1 \triangleq \theta_1 \cup \theta_3 \cup \theta_4 \cup \theta_7$, $A_2 \triangleq \theta_2 \cup \theta_5$ and $A_3 \triangleq \theta_6$ and the subset $B = \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7$ of Θ having positive belief $Bel(B) = m(X_4) + m(X_6) + m(X_9) = 0.39$. Table II summarizes the belief values of different subsets of Θ which are needed in the derivations to apply TBT.

Table II: Belief and plausibility values used for the derivations.

Subsets of Θ	$Bel(\cdot)$
$B = \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7$	$Bel(B) = 0.39$
$A_1 = \theta_1 \cup \theta_3 \cup \theta_4 \cup \theta_7$	$Bel(A_1) = 0.04$
$A_2 = \theta_2 \cup \theta_5$	$Bel(A_2) = 0.20$
$A_3 = \theta_6$	$Bel(A_3) = 0.05$
$A_1 \cap B = \theta_4 \cup \theta_7$	$Bel(A_1 \cap B) = 0.04$
$A_2 \cap B = \theta_5$	$Bel(A_2 \cap B) = 0$
$A_3 \cap B = \theta_6$	$Bel(A_3 \cap B) = 0.05$

In this example, one has

$$\begin{aligned}
\mathcal{F}_B(m) &= \{X_4, X_6, X_9\} \text{ and } \mathcal{F}_{\bar{B}}(m) = \{X_5\} \\
\mathcal{F}_{A_1}(m) &= \{X_4\} \text{ and } \mathcal{F}_{\bar{A}_1}(m) = \{X_5, X_9\} \\
\mathcal{F}_{A_2}(m) &= \{X_5\} \text{ and } \mathcal{F}_{\bar{A}_2}(m) = \{X_4, X_6, X_8, X_9\} \\
\mathcal{F}_{A_3}(m) &= \{X_9\} \text{ and } \mathcal{F}_{\bar{A}_3}(m) = \{X_1, X_2, X_4, X_5, X_7\} \\
\mathcal{F}_{A^*}(m) &= \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \mathcal{F}_{A_2}(m) - \mathcal{F}_{A_3}(m) = \{X_1, X_2, X_3, X_6, X_7, X_8\}
\end{aligned}$$

Therefore,

$$U(A^* \cap B) = \sum_{X \in \mathcal{F}_{A^*}(m) | X \in \mathcal{F}_B(m)} m(X) = m(X_6) = 0.30$$

In applying TBT formula (40), one can easily verify that

$$\begin{aligned}
Bel(B) &= Bel(B \cap A_1) + Bel(B \cap A_2) + Bel(B \cap A_3) + U(A^* \cap B) \\
&= 0.04 + 0 + 0.05 + 0.30 = 0.39
\end{aligned}$$

7.3 Special case : A partition with only two elements

If we consider any simple partition $\{A, \bar{A}\}$ of the FoD Θ and any B subset of Θ , then the TBT and TPIT formulas (40) and (43) reduce to¹⁴

$$Bel(B) = Bel(A \cap B) + Bel(\bar{A} \cap B) + U(A^* \cap B) \quad (44)$$

$$Pl(B) = Pl(\bar{A} \cup B) + Pl(A \cup B) - 1 - U(A^* \cap \bar{B}) \quad (45)$$

Remark: If the BBA $m(\cdot)$ is Bayesian then $U(A^* \cap B) = 0$. Therefore the previous formulas reduce to

$$Bel(B) = Bel(A \cap B) + Bel(\bar{A} \cap B) \quad (46)$$

$$Pl(B) = Pl(\bar{A} \cup B) + Pl(A \cup B) - 1 \quad (47)$$

$m(\cdot)$ being a Bayesian BBA, $Bel(\cdot)$ and $Pl(\cdot)$ are homogeneous to a same (possibly subjective) probability measure $P(\cdot)$. Therefore, the previous equalities can be rewritten as

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) \quad (48)$$

$$P(B) = P(\bar{A} \cup B) + P(A \cup B) - 1 \quad (49)$$

The formula (48) is valid because $\{A, \bar{A}\}$ is a partition of Θ and because of TPT theorem. The formula (49) is nothing but a dual form of TPT formula. It is also valid because

$$\begin{aligned} P(\bar{A} \cup B) + P(A \cup B) - 1 &= P(\bar{A}) + P(B) - P(\bar{A} \cap B) + P(A) + P(B) - P(A \cap B) - 1 \\ &= (P(\bar{A}) + P(A) - 1) + 2P(B) - (P(\bar{A} \cap B) + P(A \cap B)) \\ &= 0 + 2P(B) - P(B) = P(B) \end{aligned}$$

7.4 Generalization of TBT

Previously, the TBT formula was established when the partition $\{A_1, \dots, A_k\}$ was related to a given FoD Θ and B was a subset of the same FoD Θ . We can generalize TBT in considering $\{A_i, \dots, A_k\}$ as any partition of a FoD $\Theta_1 = \{x_1, \dots, x_m\} = \{x_p, p = 1, 2, \dots, m\}$, and B as being a subset of another FoD $\Theta_2 = \{y_1, \dots, y_n\} = \{y_q, q = 1, 2, \dots, n\}$ with $\Theta_1 \neq \Theta_2$. For this, we need to work within the Cartesian product space $\Theta \triangleq \Theta_1 \times \Theta_2$. In the space $\Theta = \Theta_1 \times \Theta_2$, x_p is not an elementary element but a subset of n elements of Θ , i.e. $\{x_p\} = \{(x_p, y_1), \dots, (x_p, y_n)\}$. Similarly, y_q is not an elementary element but a subset of m elements of Θ , i.e. $\{y_q\} = \{(x_1, y_q), \dots, (x_m, y_q)\}$. If $A_i \subseteq \Theta_1$ and $B \subseteq \Theta_2$, then $A_i \times B = \{(x_p, y_q) | x_p \in A_i; y_q \in B\} \subseteq \Theta$. Because $\{A_1, \dots, A_k\}$ is a partition of Θ , then $\{A_1 \times \Theta_2, \dots, A_k \times \Theta_2\}$ defines a partition of $\Theta = \Theta_1 \times \Theta_2$. Because $\Theta_1 \times B = \bigcup_{i=1, \dots, k} ((\Theta_1 \times B) \cap (A_i \times \Theta_2))$, we can apply TBT in the Cartesian space Θ . More precisely,

$$\begin{aligned} Bel(\Theta_1 \times B) &= Bel(\bigcup_i ((\Theta_1 \times B) \cap (A_i \times \Theta_2))) \\ &= Bel((\Theta_1 \times B) \cap (A_1 \times \Theta_2)) + \dots + Bel((\Theta_1 \times B) \cap (A_k \times \Theta_2)) \\ &\quad + U((A^* \times \Theta_2) \cap (\Theta_1 \times B)) \end{aligned}$$

where the quantity $U((A^* \times \Theta_2) \cap (\Theta_1 \times B))$ is now defined by

$$U((A^* \times \Theta_2) \cap (\Theta_1 \times B)) \triangleq \sum_{X \in \mathcal{F}_{A^* \times \Theta_2}(m) | X \in \mathcal{F}_{\Theta_1 \times B}(m)} m(X) \quad (50)$$

¹⁴Take $k = 2$, and set $A \triangleq A_1$ and $\bar{A} \triangleq A_2$ in (40) and (43).

The previous TBT formula when working in the Cartesian space $\Theta = \Theta_1 \times \Theta_2$ can be written more concisely as

$$Bel(\Theta_1 \times B) = \sum_{i=1, \dots, k} Bel(A_i \times B) + U(A^* \times B) \quad (51)$$

because $(\Theta_1 \times B) \cap (A_i \times \Theta_2) = (A_i \times \Theta_2) \cap (\Theta_1 \times B) = A_i \times B$, and by notation convention $U(A^* \times B) = U((A^* \times \Theta_2) \cap (\Theta_1 \times B))$.

Note that the formula (51) can be used if and only if one knows the joint BBA $m(\cdot)$ (or equivalently the joint belief) defined over the powerset of the Cartesian space $\Theta = \Theta_1 \times \Theta_2$.

8 CONDITIONAL BELIEF FUNCTIONS BASED ON TBT

In this section we show how FaginHalpern belief conditioning formulas can be established directly from TBT. This result is important because it provides a solid construction of FH formulas and it justifies its use for applications where belief conditioning is necessary. For deriving FH formulas from TBT we consider a partition $\{A_i, \bar{A}_i\}$ of the FoD Θ and a subset B of Θ . Using TBT, one has

$$Bel(B) = Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) + U(A^* \cap B) \quad (52)$$

where $\mathcal{F}_{A^*}(m) \triangleq \mathcal{F}_\Theta(m) - \mathcal{F}_{A_i}(m) - \mathcal{F}_{\bar{A}_i}(m)$ and

$$U(A^* \cap B) \triangleq \sum_{X \in \mathcal{F}_{A^*}(m) | X \in \mathcal{F}_B(m)} m(X). \quad (53)$$

Hence

$$Bel(B) - U(A^* \cap B) = Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) \quad (54)$$

At this stage, one may be tempted to divide right and left side of previous equality by $Bel(B) - U(A^* \cap B)$ (assuming its positiveness) to get

$$1 = \frac{Bel(A_i \cap B)}{Bel(B) - U(A^* \cap B)} + \frac{Bel(\bar{A}_i \cap B)}{Bel(B) - U(A^* \cap B)}$$

which would suggest to define $Bel(A_i|B)$ by taking

$$Bel(A_i|B) = Bel(A_i \cap B) / (Bel(B) - U(A^* \cap B)) \quad (55)$$

Unfortunately, it can be seen from Ellsberg's urn example that the conditional belief defined by (55) is inconsistent with bounds of imprecise conditional probabilities. Therefore, we need to go one step beyond in the calculus for defining consistent conditional belief and plausibility functions. Because by definition $U((\bar{A}_i \cap B)^*) \triangleq Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B)$, we have

$$Bel(\bar{A}_i \cap B) = Pl(\bar{A}_i \cap B) - U((\bar{A}_i \cap B)^*) \quad (56)$$

Putting this expression of $Bel(\bar{A}_i \cap B)$ into (54) and rearranging terms, we get

$$Bel(B) + \Delta(U) = Bel(A_i \cap B) + Pl(\bar{A}_i \cap B) \quad (57)$$

with $\Delta(U) \triangleq U((\bar{A}_i \cap B)^*) - U(A^* \cap B)$ and $\Delta(U) \in [0, 1]$ (see proof in appendix).

Assuming $Bel(B) > 0$, and dividing each side of (57) by $Bel(B) + \Delta(U)$, we get

$$1 = \frac{Bel(A_i \cap B)}{Bel(B) + \Delta(U)} + \frac{Pl(\bar{A}_i \cap B)}{Bel(B) + \Delta(U)} \quad (58)$$

or equivalently

$$\frac{Bel(A_i \cap B)}{Bel(B) + \Delta(U)} = 1 - \frac{Pl(\bar{A}_i \cap B)}{Bel(B) + \Delta(U)} \quad (59)$$

Because the general relationship $Bel(X) = 1 - Pl(\bar{X})$ between the belief and the plausibility must always be satisfied for any $X \subseteq \Theta$, the equality (59) allows to define the conditional belief $Bel(A_i|B)$ and $Pl(\bar{A}_i|B)$ by taking

$$Bel(A_i|B) \triangleq \frac{Bel(A_i \cap B)}{Bel(B) + \Delta(U)} \quad (60)$$

$$Pl(\bar{A}_i|B) \triangleq \frac{Pl(\bar{A}_i \cap B)}{Bel(B) + \Delta(U)} \quad (61)$$

Using equality (57), the previous conditioning formulas can be rewritten equivalently as

$$Bel(A_i|B) = \frac{Bel(A_i \cap B)}{Bel(A_i \cap B) + Pl(\bar{A}_i \cap B)} \quad (62)$$

$$Pl(\bar{A}_i|B) = \frac{Pl(\bar{A}_i \cap B)}{Bel(A_i \cap B) + Pl(\bar{A}_i \cap B)} \quad (63)$$

In replacing \bar{A}_i by A_i in notations of formulas (61)–(63) we get¹⁵ the conditional plausibility $Pl(A_i|B)$ as

$$Pl(A_i|B) \triangleq \frac{Pl(A_i \cap B)}{Bel(B) + U((A_i \cap B)^*) - U(A^* \cap B)} = \frac{Pl(A_i \cap B)}{Bel(\bar{A}_i \cap B) + Pl(A_i \cap B)} \quad (64)$$

Formulas (62) and (64) coincide with Fagin-Halpern formulas⁴ which were originally proposed from essentially a very good intuition. In this work, we have derived Fagin-Halpern formulas only from TBT using the proper decomposition of the set of focal elements of the a priori BBA. Note that the definition of $Bel(A_i|B)$ given in (60) satisfies the conditions $Bel(\emptyset|B) = 0$, $Bel(\Theta|B) = 1$, and $Bel(A_i|B) \in [0, 1]$. To prove that $Bel(A_i|B)$ defined by (62) is a belief function one must prove that it is also an n -monotone ($n \geq 2$) Choquet's capacity³⁶ on the finite set Θ , or equivalently that the following inequality holds for any $B \subseteq \Theta$ with $Bel(B) > 0$ and for any collection A_1, \dots, A_n of subsets of Θ

$$Bel(A_1 \cup \dots \cup A_n|B) \geq \sum_{\substack{I \subseteq \{1, \dots, n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Bel(\bigcap_{i \in I} A_i|B)$$

The proof of this inequality is complicate. However, three very different proofs have already been given by Fagin and Halpern³, Jaffray⁶, and Sundberg and Wagner⁷, the latter one being the clearest of fashion.

9 GENERALIZATION OF BAYES' THEOREM

In this section and thanks to the previous results, we generalize Bayes' Theorem (BT) in the framework of belief functions. Assuming $Bel(B) > 0$, we have shown that Fagin-Halpern expression of $Bel(A_i|B)$ given by

$$Bel(A_i|B) = \frac{Bel(A_i \cap B)}{Bel(A_i \cap B) + Pl(\bar{A}_i \cap B)} \quad (65)$$

¹⁵It is worth to note that one has always $U(A^*) = U(A_i^*) = U(\bar{A}_i^*)$ when partitioning Θ as $\{A_i, \bar{A}_i\}$ because $U(A_i^*) = Pl(A_i) - Bel(A_i) = (1 - Bel(\bar{A}_i)) - (1 - Pl(\bar{A}_i)) = Pl(\bar{A}_i) - Bel(\bar{A}_i) = U(\bar{A}_i^*)$.

is equal to the formula (60), i.e.

$$Bel(A_i|B) = \frac{Bel(A_i \cap B)}{Bel(B) + U((\bar{A}_i \cap B)^*) - U(A^* \cap B)} \quad (66)$$

In replacing $Bel(B)$ by the expression (40) of TBT we get

$$Bel(A_i|B) = \frac{Bel(A_i \cap B)}{\sum_{i=1, \dots, k} Bel(A_i \cap B) + U((\bar{A}_i \cap B)^*)} \quad (67)$$

Assuming $Bel(A_i) > 0$, Fagin-Halpern expression of $Bel(B|A_i)$ given by

$$Bel(B|A_i) = \frac{Bel(B \cap A_i)}{Bel(B \cap A_i) + Pl(\bar{B} \cap A_i)} \quad (68)$$

is equal to

$$Bel(B|A_i) = \frac{Bel(B \cap A_i)}{Bel(A_i) + U((\bar{B} \cap A_i)^*) - U(B^* \cap A_i)} \quad (69)$$

where

$$U((\bar{B} \cap A_i)^*) \triangleq Pl(\bar{B} \cap A_i) - Bel(\bar{B} \cap A_i) = \sum_{X \in \mathcal{F}_{(\bar{B} \cap A_i)^*}(m)} m(X) \quad (70)$$

with $\mathcal{F}_{(\bar{B} \cap A_i)^*}(m) = \mathcal{F}_\Theta(m) - \mathcal{F}_{\bar{B} \cap A_i}(m) - \mathcal{F}_{B \cup \bar{A}_i}(m)$, and where

$$U(B^* \cap A_i) \triangleq \sum_{X \in \mathcal{F}_{B^* \cap A_i}(m)} m(X) \quad (71)$$

with $\mathcal{F}_{B^* \cap A_i}(m) = \mathcal{F}_\Theta(m) - \mathcal{F}_B(m) - \mathcal{F}_{\bar{B}}(m)$.

From (69), one obtains

$$Bel(A_i \cap B) = Bel(B|A_i)[Bel(A_i) + U((\bar{B} \cap A_i)^*) - U(B^* \cap A_i)]$$

By replacing the above expression of $Bel(A_i \cap B)$ into (67), we obtain

$$Bel(A_i|B) = \frac{Bel(B|A_i)q(A_i, B)}{\sum_{i=1}^k Bel(B|A_i)q(A_i, B) + U((\bar{A}_i \cap B)^*)} \quad (72)$$

where the factor $q(A_i, B)$ introduced here for notation conciseness is defined by

$$q(A_i, B) \triangleq Bel(A_i) + U((\bar{B} \cap A_i)^*) - U(B^* \cap A_i) \quad (73)$$

This result allows us to establish the following Generalized Bayes' Theorem (GBT).

Generalized Bayes' Theorem (GBT): For any partition $\{A_i, i = 1, \dots, k\}$ of a FoD Θ , any belief function $Bel(\cdot) : 2^\Theta \mapsto [0, 1]$, and any subset B of Θ with $Bel(B) > 0$, then one has

$$Bel(A_i|B) = \frac{Bel(B|A_i)q(A_i, B)}{\sum_{i=1}^k Bel(B|A_i)q(A_i, B) + U((\bar{A}_i \cap B)^*)} \quad (74)$$

where $U((\bar{A}_i \cap B)^*) \triangleq \sum_{X \in \mathcal{F}_{(\bar{A}_i \cap B)^*}(m)} m(X) = Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B)$, and where $q(A_i, B) = Bel(A_i) + U((\bar{B} \cap A_i)^*) - U(B^* \cap A_i)$.

Lemma: GBT reduces to Bayes' Theorem if $Bel(\cdot) : 2^\Theta \mapsto [0, 1]$ is a Bayesian belief function.

Proof: See appendix.

When $A_i \subseteq \Theta_1$ and $B \subseteq \Theta_2$ with $\Theta_1 \neq \Theta_2$, we must work in the Cartesian product space $\Theta = \Theta_1 \times \Theta_2$ and the GBT formula is similar to (74) in replacing A_i by $A_i \times \Theta_2$, B by $\Theta_1 \times B$, and where

$$\begin{aligned} U((\bar{A}_i \cap B)^*) &\triangleq \sum_{X \in \mathcal{F}_{((\bar{A}_i \times \Theta_2) \cap (\Theta_1 \times B))^*}^{(m)}} m(X) \\ &= Pl((\bar{A}_i \times \Theta_2) \cap (\Theta_1 \times B)) - Bel((\bar{A}_i \times \Theta_2) \cap (\Theta_1 \times B)) \end{aligned} \quad (75)$$

and where the factor $q(A_i, B)$ must be replaced by

$$q(A_i \times \Theta_2, \Theta_1 \times B) \triangleq Bel(A_i \times \Theta_2) + U((\bar{B} \cap A_i)^*) - U(B^* \cap A_i) \quad (76)$$

with

$$\begin{aligned} U((\bar{B} \cap A_i)^*) &\triangleq \sum_{X \in \mathcal{F}_{(\Theta_1 \times \bar{B}) \cap (A_i \times \Theta_2)^*}^{(m)}} m(X) \\ &= Pl((\Theta_1 \times \bar{B}) \cap (A_i \times \Theta_2)) - Bel((\Theta_1 \times \bar{B}) \cap (A_i \times \Theta_2)) \end{aligned} \quad (77)$$

$$U(B^* \cap A_i) \triangleq \sum_{X \in \mathcal{F}_{\Theta_1 \times B^*}^{(m)} | X \in \mathcal{F}_{A_i \times \Theta_2}^{(m)}} m(X) \quad (78)$$

and $\mathcal{F}_{\Theta_1 \times B^*}(m) = \mathcal{F}_{\Theta_1 \times \Theta_2}(m) - \mathcal{F}_{\Theta_1 \times B}(m) - \mathcal{F}_{\Theta_1 \times \bar{B}}(m)$.

In the formulas (75)–(78), X is an elementary element of the Cartesian space $\Theta = \Theta_1 \times \Theta_2$, and $m(X)$ is the (joint) BBA value of X defined on the power set of Cartesian product space.

The application of GBT formula when working with $A_i \subseteq \Theta_1$ and $B \subseteq \Theta_2$ with $\Theta_1 \neq \Theta_2$ is not easy in general because it requires the knowledge of joint BBA $m(\cdot)$ defined over $2^{\Theta_1 \times \Theta_2}$ which is rarely known in practice. If the joint BBA $m(\cdot)$ can be expressed (or approximated) as a function of two marginal BBAs $m^{\Theta_1}(\cdot)$ and $m^{\Theta_2}(\cdot)$ (assumed to be known) defined respectively over Θ_1 and Θ_2 , then GBT formula should become tractable.

10 ILLUSTRATIVE EXAMPLE OF GBT

In this section, we provide a complete quite simple illustrative example to show how belief conditioning formulas work and how to apply GBT.

Let us consider the FoD $\Theta = \{\theta_i, i = 1, \dots, 7\}$ and the set of focal elements $\mathcal{F}_\Theta(m) = \{X_1, X_2, \dots, X_9\}$ of a BBA $m(\cdot)$ defined over 2^Θ given in Table III. Let's consider the partition $\{A_1, A_2, A_3\}$ of Θ with $A_1 = \theta_1 \cup \theta_3 \cup \theta_4 \cup \theta_7$, $A_2 = \theta_2 \cup \theta_5$ and $A_3 = \theta_6$, and let consider the subset $B = \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7$ of Θ having positive belief $Bel(B) = m(X_4) + m(X_6) + m(X_9) = 0.39$. Table IV summarizes the belief and plausibility values of different subsets of Θ which are needed in the derivations.

Table III: Focal elements and their masses.

Focal element X	BBA $m(X)$
$X_1 = \theta_2 \cup \theta_3 \cup \theta_4 \cup \theta_5 \cup \theta_7$	$m(X_1) = 0.01$
$X_2 = \theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4$	$m(X_2) = 0.02$
$X_3 = \theta_3 \cup \theta_5 \cup \theta_6$	$m(X_3) = 0.03$
$X_4 = \theta_4 \cup \theta_7$	$m(X_4) = 0.04$
$X_5 = \theta_2$	$m(X_5) = 0.20$
$X_6 = \theta_6 \cup \theta_7$	$m(X_6) = 0.30$
$X_7 = \theta_2 \cup \theta_3 \cup \theta_7$	$m(X_7) = 0.20$
$X_8 = \theta_1 \cup \theta_4 \cup \theta_6$	$m(X_8) = 0.15$
$X_9 = \theta_6$	$m(X_9) = 0.05$

Table IV: Belief and plausibility values used for the derivations.

Subsets of Θ	$Bel(\cdot)$	$Pl(\cdot)$
$B = \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7$	$Bel(B) = 0.39$	$Pl(B) = 0.80$
$A_1 = \theta_1 \cup \theta_3 \cup \theta_4 \cup \theta_7$	$Bel(A_1) = 0.04$	$Pl(A_1) = 0.75$
$A_2 = \theta_2 \cup \theta_5$	$Bel(A_2) = 0.20$	$Pl(A_2) = 0.46$
$A_3 = \theta_6$	$Bel(A_3) = 0.05$	$Pl(A_3) = 0.53$
$A_1 \cap B = \theta_4 \cup \theta_7$	$Bel(A_1 \cap B) = 0.04$	$Pl(A_1 \cap B) = 0.72$
$A_2 \cap B = \theta_5$	$Bel(A_2 \cap B) = 0$	$Pl(A_2 \cap B) = 0.04$
$A_3 \cap B = \theta_6$	$Bel(A_3 \cap B) = 0.05$	$Pl(A_3 \cap B) = 0.53$
$\bar{A}_1 \cap B = \theta_5 \cup \theta_6$	$Bel(\bar{A}_1 \cap B) = 0.05$	$Pl(\bar{A}_1 \cap B) = 0.54$
$\bar{A}_2 \cap B = \theta_4 \cup \theta_6 \cup \theta_7$	$Bel(\bar{A}_2 \cap B) = 0.39$	$Pl(\bar{A}_2 \cap B) = 0.80$
$\bar{A}_3 \cap B = \theta_4 \cup \theta_5 \cup \theta_7$	$Bel(\bar{A}_3 \cap B) = 0.04$	$Pl(\bar{A}_3 \cap B) = 0.75$
$A_1 \cap \bar{B} = \theta_1 \cup \theta_3$	$Bel(A_1 \cap \bar{B}) = 0$	$Pl(A_1 \cap \bar{B}) = 0.41$
$A_2 \cap \bar{B} = \theta_2$	$Bel(A_2 \cap \bar{B}) = 0.20$	$Pl(A_2 \cap \bar{B}) = 0.43$
$A_3 \cap \bar{B} = \emptyset$	$Bel(A_3 \cap \bar{B}) = 0$	$Pl(A_3 \cap \bar{B}) = 0$

In this example, one has

$$\begin{aligned}
\mathcal{F}_B(m) &= \{X_4, X_6, X_9\} \text{ and } \mathcal{F}_{\bar{B}}(m) = \{X_5\} \\
\mathcal{F}_{B^*}(m) &= \mathcal{F}_\Theta(m) - \mathcal{F}_B(m) - \mathcal{F}_{\bar{B}}(m) = \{X_1, X_2, X_3, X_7, X_8\} \\
\mathcal{F}_{A_1}(m) &= \{X_4\} \text{ and } \mathcal{F}_{\bar{A}_1}(m) = \{X_5, X_9\} \\
\mathcal{F}_{A_2}(m) &= \{X_5\} \text{ and } \mathcal{F}_{\bar{A}_2}(m) = \{X_4, X_6, X_8, X_9\} \\
\mathcal{F}_{A_3}(m) &= \{X_9\} \text{ and } \mathcal{F}_{\bar{A}_3}(m) = \{X_1, X_2, X_4, X_5, X_7\} \\
\mathcal{F}_{A^*}(m) &= \mathcal{F}_\Theta(m) - \mathcal{F}_{A_1}(m) - \mathcal{F}_{A_2}(m) - \mathcal{F}_{A_3}(m) = \{X_1, X_2, X_3, X_6, X_7, X_8\}
\end{aligned}$$

• **Results with Fagin-Halpern conditioning formulas**

Using Fagin-Halpern conditioning formulas (62) and (68) and the fact that $Pl(A_i|B) = 1 - Bel(\bar{A}_i|B)$ and $Pl(B|A_i) = 1 - Bel(\bar{B}|A_i)$, we obtain in this example the conditional belief and plausibility values given in Tables V–VI

Table V: $Bel(A_i|B)$ and $Pl(A_i|B)$ with Fagin-Halpern conditioning.

Subsets of Θ	$Bel(A_i B)$	$Pl(A_i B)$
A_1	$Bel(A_1 B) \approx 0.0690$	$Pl(A_1 B) \approx 0.9351$
A_2	$Bel(A_2 B) = 0$	$Pl(A_2 B) \approx 0.0930$
A_3	$Bel(A_3 B) \approx 0.0625$	$Pl(A_3 B) \approx 0.9298$

Table VI: $Bel(B|A_i)$ and $Pl(B|A_i)$ with Fagin-Halpern conditioning.

Subsets of Θ	$Bel(B A_i)$	$Pl(B A_i)$
A_1	$Bel(B A_1) \approx 0.0889$	$Pl(B A_1) = 1$
A_2	$Bel(B A_2) = 0$	$Pl(B A_2) \approx 0.1667$
A_3	$Bel(B A_3) = 1$	$Pl(B A_3) = 1$

To apply and verify GBT on this example, one needs to compute $Bel(A_i)$, $U((\bar{B} \cap A_i)^*)$ and $U(B^* \cap A_i)$ to calculate $q(A_i, B)$ factors and also $U((\bar{A}_i \cap B)^*)$ because they enter in GBT formula (74). These values are listed in Table VII for convenience.

Table VII: Values of $q(A_i, B)$ and $U((\bar{A}_i \cap B)^*)$ for GBT formula.

Subsets of Θ	$q(A_i, B)$	$U((\bar{A}_i \cap B)^*)$
A_1	0.45	0.49
A_2	0.43	0.41
A_3	0.05	0.71

The value $q(A_1, B) = 0.45$ appearing in Table VII has been calculated as follows

$$q(A_1, B) \triangleq Bel(A_1) + U((\bar{B} \cap A_1)^*) - U(B^* \cap A_1) = 0.45$$

because

$$\begin{aligned} Bel(A_1) &= 0.04 \\ U((\bar{B} \cap A_1)^*) &= Pl(\bar{B} \cap A_1) - Bel(\bar{B} \cap A_1) = 0.41 \\ U(B^* \cap A_1) &= \sum_{X \in \mathcal{F}_{A_1}(m) | X \in \mathcal{F}_{B^*}(m)} m(X) = 0 \end{aligned}$$

The value $U((\bar{A}_1 \cap B)^*) = 0.49$ appearing in Table VII is calculated as follows

$$U((\bar{A}_1 \cap B)^*) = Pl(\bar{A}_1 \cap B) - Bel(\bar{A}_1 \cap B) = 0.54 - 0.05 = 0.49$$

Other values of Table VII are calculated similarly.

One verifies that GBT formula (74) works because we retrieve correct values obtained with FH formula, given in Table V. Indeed, one has

$$\begin{aligned} Bel(A_1|B) &= \frac{Bel(B|A_1)q(A_1, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_1 \cap B)^*)} \\ &\approx \frac{0.0889 \cdot 0.45}{(0.0889 \cdot 0.45) + (0 \cdot 0.43) + (1 \cdot 0.05) + 0.49} \approx 0.0690 \end{aligned}$$

$$\begin{aligned} Bel(A_2|B) &= \frac{Bel(B|A_2)q(A_2, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_2 \cap B)^*)} \\ &\approx \frac{0 \cdot 0.43}{(0.0889 \cdot 0.45) + (0 \cdot 0.43) + (1 \cdot 0.05) + 0.41} = 0 \end{aligned}$$

$$\begin{aligned} Bel(A_3|B) &= \frac{Bel(B|A_3)q(A_3, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_3 \cap B)^*)} \\ &\approx \frac{1 \cdot 0.05}{(0.0889 \cdot 0.45) + (0 \cdot 0.43) + (1 \cdot 0.05) + 0.71} \approx 0.0625 \end{aligned}$$

• **Results with Shafer's conditioning formulas**

Using Shafer's conditioning formulas (30) and (31), we obtain in this example the conditional belief and plausibility values given in Table VIII and IX.

Table VIII: $Bel(A_i|B)$ and $Pl(A_i|B)$ with Shafer's conditioning.

Subsets of Θ	$Bel(A_i B)$	$Pl(A_i B)$
A_1	$Bel(A_1 B) = 0.3250$	$Pl(A_1 B) = 0.9000$
A_2	$Bel(A_2 B) = 0$	$Pl(A_2 B) = 0.0500$
A_3	$Bel(A_3 B) = 0.0625$	$Pl(A_3 B) = 0.6625$

Table IX: $Bel(B|A_i)$ and $Pl(B|A_i)$ with Shafer's conditioning.

Subsets of Θ	$Bel(B A_i)$	$Pl(B A_i)$
A_1	$Bel(B A_1) \approx 0.4533$	$Pl(B A_1) \approx 0.9600$
A_2	$Bel(B A_2) \approx 0.0652$	$Pl(B A_2) \approx 0.0870$
A_3	$Bel(B A_3) = 1$	$Pl(B A_3) = 1$

As shown in the previous Ellsberg's urn example, one knows that Shafer's belief conditioning formulas are inconsistent with lower and upper bounds of imprecise conditional probabilities, and with this example one shows that Shafer's belief conditioning is also incompatible with GBT formula (74). We emphasize that GBT has been established by a constructive manner from TBT

using a direct and relatively simple calculus¹⁶ without need of rule of combination of basic belief assignments. When using Shafer's belief conditioning formulas, one sees that the conditional values are not coherent since they do not verify GBT because we obtain in this example

$$\begin{aligned}
Bel(A_1|B) &= 0.3250 \text{ (from the results in Table VIII using eq. (31))} \\
&\neq \frac{Bel(B|A_1)q(A_1, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_1 \cap B)^*)} \\
&\approx \frac{0.4533 \cdot 0.45}{(0.4533 \cdot 0.45) + (0.0652 \cdot 0.43) + (1 \cdot 0.05) + 0.49} \approx 0.2642 \\
\\
Bel(A_2|B) &= 0 \text{ (from the results in Table VIII using eq. (31))} \\
&\neq \frac{Bel(B|A_2)q(A_2, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_2 \cap B)^*)} \\
&\approx \frac{0.0652 \cdot 0.43}{(0.4533 \cdot 0.45) + (0.0652 \cdot 0.43) + (1 \cdot 0.05) + 0.41} \approx 0.0405 \\
\\
Bel(A_3|B) &= 0.0625 \text{ (from the results in Table VIII using eq. (31))} \\
&\neq \frac{Bel(B|A_3)q(A_3, B)}{\sum_{i=1}^3 Bel(B|A_i)q(A_i, B) + U((\bar{A}_3 \cap B)^*)} \\
&\approx \frac{1 \cdot 0.05}{(0.4533 \cdot 0.45) + (0.0652 \cdot 0.43) + (1 \cdot 0.05) + 0.71} \approx 0.0504
\end{aligned}$$

Ellsberg's urn example and this example show clearly that Dempster's rule of combination used by Shafer to establish his belief and conditioning formulas does not provide coherent and satisfactory results since they are inconsistent with lower and upper bounds of imprecise conditional probabilities and they do not satisfy GBT also.

11 CONCLUSION

In this paper new important results for reasoning with belief functions were obtained and discussed. The Total Belief Theorem (TBT) was established from a simple decomposition of the set of focal elements of any basic belief assignment. TBT is a generalization of Total Probability Theorem for belief functions, and based on it we are able to derive conditional belief and conditional plausibility functions that coincide with Fagin-Halpern conditioning formulas and which are coherent with lower and upper bounds of imprecise conditional probability. Hence, this work provides a solid justification of the establishment of formulas presented by Fagin and Halpern. The TBT has been generalized for dealing with different frame of discernments as well thanks to the Cartesian product space. Also as a direct consequence of TBT, we have presented a generalization of the well-known Bayes' Theorem for the framework of belief functions called the Generalized Bayesian Theorem (GBT). We have proved that TBT and GBT reduce to TPT and BT respectively as soon as we work with Bayesian belief function because in this case the Bayesian belief function is homogeneous to a probability measure. On the base of Ellsberg's urn an example and an illustrative example we have shown that Dempster's rule of combination used by Shafer to establish his belief and conditioning formulas does not provide coherent and satisfactory results because they are inconsistent with lower and upper bounds of imprecise conditional probabilities and because they do not satisfy GBT also. These new theoretical results should (we hope) reconcile the Bayesian reasoning practioners with evidential reasoning practioners and bring new foundations for reasoning with uncertainty thanks to belief functions.

¹⁶assuming $Bel(B)$ and $Bel(A_i)$ being positive to have well defined expressions as it is for this example.

12 ACKNOWLEDGEMENT

The reported work is partially supported by the National Natural Science Foundation of China (No. 61573275, No. 61671370), Postdoctoral Science Research Foundation of Shaanxi Province (No. 2016BSHEDZZ46), Postdoctoral Science Foundation of China (No. 2016M592790), and Fundamental Research Funds for the Central Universities (No. xjj2016066)

APPENDIX

Proof of TBT

From the basic definition of $Bel(B)$ one has for any $B \subseteq \Theta$, $Bel(B) = \sum_{X \in \mathcal{F}_\Theta(m) | X \subseteq B} m(X)$. Because the set of focal elements $\mathcal{F}_\Theta(m)$ can always be decomposed as the union $\mathcal{F}_\Theta(m) = \mathcal{F}_{A_1}(m) \cup \dots \cup \mathcal{F}_{A_k}(m) \cup \mathcal{F}_{A^*}(m)$, then one can always decompose the previous sum as follows

$$\begin{aligned}
 Bel(B) &= \sum_{X \in \mathcal{F}_\Theta(m) | X \subseteq B} m(X) \\
 &= \sum_{X \in \mathcal{F}_{A_1}(m) | X \subseteq \mathcal{F}_B(m)} m(X) + \dots \\
 &\quad + \sum_{X \in \mathcal{F}_{A_k}(m) | X \subseteq \mathcal{F}_B(m)} m(X) \\
 &\quad + \sum_{X \in \mathcal{F}_{A^*}(m) | X \subseteq \mathcal{F}_B(m)} m(X) \\
 &= Bel(A_1 \cap B) + \dots + Bel(A_k \cap B) + \sum_{X \in \mathcal{F}_{A^*}(m) | X \subseteq \mathcal{F}_B(m)} m(X) \\
 &= \sum_{i=1, \dots, k} Bel(A_i \cap B) + U(A^* \cap B)
 \end{aligned}$$

where $U(A^* \cap B) \triangleq \sum_{X \in \mathcal{F}_{A^*}(m) | X \subseteq \mathcal{F}_B(m)} m(X)$, which completes the proof of TBT.

Proof of the corollary of TBT

If $m(\cdot)$ is Bayesian then any focal element X of $\mathcal{F}_\Theta(m)$ is a singleton of 2^Θ which either belongs to A_i , or to \bar{A}_i (but it cannot belong to both). Therefore, $\mathcal{F}_\Theta(m) = \mathcal{F}_{A_1}(m) \cup \dots \cup \mathcal{F}_{A_k}(m)$ and $\mathcal{F}_{A^*}(m) = \emptyset$. TBT formula applies with¹⁷ $U(A^* \cap B) = \sum_{X \in \mathcal{F}_{A^*}(m) | X \subseteq \mathcal{F}_B(m)} m(X) = \sum_{X \in \emptyset | X \subseteq \mathcal{F}_B(m)} m(X) = 0$ and thanks to TBT one has in this case for any partition $\{A_1, \dots, A_k\}$ of Θ and any subset B of Θ the following equality satisfied

$$Bel(B) = \sum_{i=1, \dots, k} Bel(A_i \cap B) \tag{79}$$

When $m(\cdot)$ is Bayesian, its corresponding belief function $Bel(\cdot)$ is homogeneous to a probability measure $P(\cdot)$ ¹, and therefore the previous equality is consistent with TPT formula (13), which completes the proof of the corollary of the TBT.

¹⁷We recall that if a summation has no term then its value is set to zero.

Proof of TPIT

From equality $Pl(B) = 1 - Bel(\bar{B})$ and TBT, one has

$$\begin{aligned}
Pl(B) &= 1 - Bel(\bar{B}) \\
&= 1 - \sum_{i=1, \dots, k} Bel(A_i \cap \bar{B}) - U(A^* \cap \bar{B}) \\
&= 1 - \sum_{i=1, \dots, k} (Bel(A_i \cap \bar{B}) + 1 - 1) - U(A^* \cap \bar{B}) \\
&= 1 - \sum_{i=1, \dots, k} (-1 + Bel(A_i \cap \bar{B}) + 1) - U(A^* \cap \bar{B}) \\
&= 1 - \sum_{i=1, \dots, k} (-(1 - Bel(A_i \cap \bar{B})) + 1) - U(A^* \cap \bar{B}) \\
&= 1 + \sum_{i=1, \dots, k} Pl(\overline{A_i \cap \bar{B}}) - k - U(A^* \cap \bar{B}) \\
&= \sum_{i=1, \dots, k} Pl(\bar{A}_i \cup B) + 1 - k - U(A^* \cap \bar{B})
\end{aligned}$$

which completes the proof of TPIT.

Proof that $\Delta(U) \in [0, 1]$

$$\begin{aligned}
\Delta(U) &\triangleq U((\bar{A}_i \cap B)^*) - U(A^* \cap B) \\
&= [Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B)] - [Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) - Bel(B)] \\
&= Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B) + Bel(B) - Bel(A_i \cap B) - Bel(\bar{A}_i \cap B)
\end{aligned}$$

To prove that $\Delta(U) \geq 0$, one must prove equivalently that

$$Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B) + Bel(B) \geq Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) \quad (80)$$

Using TBT, one has

$$Bel(B) = Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) + U(A^* \cap B)$$

Replacing expression of $Bel(B)$ in inequality (80), one must verify if the following equality is satisfied

$$\begin{aligned}
Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B) + Bel(A_i \cap B) + Bel(\bar{A}_i \cap B) + U(A^* \cap B) \\
\geq Bel(A_i \cap B) + Bel(\bar{A}_i \cap B)
\end{aligned}$$

After simplification, we have to prove that the following inequality holds

$$Pl(\bar{A}_i \cap B) + U(A^* \cap B) \geq Bel(\bar{A}_i \cap B)$$

Because $Pl(\bar{A}_i \cap B) = Bel(\bar{A}_i \cap B) + U((\bar{A}_i \cap B)^*)$, one has to verify if the following inequality holds

$$Bel(\bar{A}_i \cap B) + U((\bar{A}_i \cap B)^*) + U(A^* \cap B) \geq Bel(\bar{A}_i \cap B)$$

After simplification (omitting both $Bel(\bar{A}_i \cap B)$ in left and right side of the previous inequality), one has to prove that the inequality below is satisfied to prove that $\Delta(U) \geq 0$

$$U((\bar{A}_i \cap B)^*) + U(A^* \cap B) \geq 0$$

Because $U((\bar{A}_i \cap B)^*) \in [0, 1]$ and $U(A^* \cap B) \in [0, 1]$, the previous inequality always holds which proves that $U((\bar{A}_i \cap B)^*) - U(A^* \cap B) \geq 0$. Moreover because $U(A^* \cap B) \in [0, 1]$, then $-U(A^* \cap B) \in [-1, 0]$. Because $U((\bar{A}_i \cap B)^*) \in [0, 1]$, one deduces that $U((\bar{A}_i \cap B)^*) - U(A^* \cap B) \leq 1$. This completes the proof.

Proof of Lemma

If $Bel(\cdot) : 2^\Theta \mapsto [0, 1]$ is a Bayesian belief function, then all focal elements of its corresponding BBA $m(\cdot)$ are singletons of 2^Θ . In this case $Bel(\cdot)$ and $Pl(\cdot)$ functions coincide and therefore one has $U((\bar{A}_i \cap B)^*) = Pl(\bar{A}_i \cap B) - Bel(\bar{A}_i \cap B) = 0$ and $U((\bar{B} \cap A_i)^*) = Pl(\bar{B} \cap A_i) - Bel(\bar{B} \cap A_i) = 0$. Any focal element (singleton) of $m(\cdot)$ is either a subset of B or a subset of \bar{B} of the FoD Θ . Therefore, $\mathcal{F}_{B^*}(m) = \emptyset$, which implies $U(B^* \cap A_i) = 0$, so that $q(A_i, B) = Bel(A_i)$. The GBT formula (74) with in this case $q(A_i, B) = Bel(A_i)$ and $U((\bar{A}_i \cap B)^*) = 0$ reduces to formula

$$Bel(A_i|B) = \frac{Bel(B|A_i)Bel(A_i)}{\sum_{i=1}^k Bel(B|A_i)Bel(A_i)}$$

which coincides with formula (20) because $Bel(\cdot)$ (being a Bayesian belief function) is homogeneous to a probability measure $P(\cdot)$. This completes the proof that GBT formula is consistent with Bayesian Theorem formula when the Belief function is Bayesian.

References

1. Shafer G. A Mathematical Theory of Evidence, Princeton Univ. Press, 1976.
2. Dempster A. Upper and lower probabilities induced by a multivalued mapping, Ann. of Math. Stat., 1967; 38: 325-339.
3. Fagin R, Halpern JY. A new approach to updating beliefs, in Proc. Conf. on Uncertainty in Artificial Intelligence (UAI), 1991; 317-325.
4. Halpern JY, Fagin R. Two views of belief: belief as generalized probability and belief as evidence, Artificial Intelligence, 1992; 54: 275-317.
5. De Campos L, Lamata T, Moral S. The concept of conditional fuzzy measure, Int. J. of Intel. Systems, 1990; 5: 237-246.
6. Jaffray JY. Bayesian updating and belief functions, IEEE Trans on SMC, Sept.-Oct. 1992; 22(5): 1144-1152.
7. Sunberg C, Wagner C. Generalized finite differences and Bayesian conditioning of Choquet capacities, Advances in Applied Mathematics, 1992, 13: 262-272.
8. Walley P. Measures of uncertainty in expert systems, Artificial Intelligence, 1996; 83: 1-58.
9. Aitchison J. Discussion on Professor Dempster's paper, Journal of the Royal Statistical Society, Serie B, 1968; 30: 234-237.
10. Diaconis P. Review of "A Mathematical Theory of Evidence", Journal of the American Statistical Society, 1978; 73 (363): 677-678.

11. Diaconis P, Zabell SL. Updating subjective probability, *Journal of the American Statistical Society*, 1982; 77(380): 822-830.
12. Zadeh L. A Mathematical Theory of Evidence (book review), *AI Magazine*, 1984; 5(3): 81-83.
13. Lemmer J. Confidence factors, empiricism and the Dempster-Shafer theory of evidence, *Proc. of 1st Conf. on UAI*, 1985; 160-176.
14. Black PK. Is Shafer general Bayes?, *Proc. of 3rd Workshop on Uncertainty in Artificial Intelligence*, Seattle, WA, USA, 1987; 2-9.
15. Hunter D. Dempster-Shafer vs. probabilistic logic, *Third AAAI Uncertainty in Artificial Intelligence Workshop*, 1987; 22-29.
16. Voorbraak F. On the justification of Dempster's rule of combination, *Utrecht Univ., Netherlands, Logic Group Preprint Series*, 1988; No. 42.
17. Wang P. A defect in Dempster-Shafer theory, in *Proc. of 10th Conf. on Uncertainty in AI*, 1994; 560-566.
18. Gelman A. The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, *American Statistician*, 2006; 60(2): 146-150.
19. Pearl J. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, 1988.
20. Pearl J. Reasoning with belief functions: a critical assessment, *Technical Report R-136*, UCLA, 1989.
21. Dezert J, Wang P, Tchamova A. On the validity of Dempster-Shafer theory, *Proc. of Fusion 2012*, Singapore, July 9–12, 2012.
22. Tchamova A, Dezert J. On the behavior of Dempster's Rule of combination and the foundations of Dempster-Shafer theory, *6th IEEE Int. Conf. on Int. Syst.*, Sofia, Bulgaria, Sept. 6–8, 2012.
23. Dezert J, Tchamova A. On the validity of Dempster's fusion rule and its interpretation as a generalization of Bayesian fusion rule, *Int. J. of Intell. Syst.*, March 2014; 29(3): 223-252.
24. Heendani, J.N., Premaratne, K., Murthi, M.N., Uscinski, J., Scheutz, M. A generalization of Bayesian inference in the Dempster-Shafer belief theoretic framework, *Fusion 2016 Proc.*, Heidelberg, July 2016.
25. Smets P. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem, *Int. J. of Approximate Reasoning*, 1993; 9(1): 1-35.
26. Dubois D, Denœux T. Conditioning in Dempster-Shafer Theory: Prediction vs. Revision, in *Proc. of Belief 2012 Int. Conf*, Compiègne, France, May 2012; (*Belief Functions: Theory & Appl.*, AISC 164, Springer, 385-392).
27. Dempster A. A generalization of Bayesian inference, *J. of Royal Sta. Soc.*, 1968; B30: 205-247.
28. Smarandache F, Dezert J.(Editors) *Advances and applications of DSMT for information fusion*, American Research Press, 2004–2015; Vols. 1–4.
<http://www.onera.fr/staff/jean-dezert?page=2>
29. Halmos PR. *Naive Set Theory*, Springer-Verlag New York, 1974.
30. Li XR. *Probability, random signals, and statistics*, CRC Press, 1999.
31. Papoulis A. *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984; 37-38.

32. Grimmett G, Welsh D. Probability: An Introduction, Oxford Science Publications, 1986.
33. Dezert J, Tchamova A, Han D, Tacnet JM. Why Dempster's rule doesn't behave as Bayes rule with informative priors, Proc. of 2013 IEEE Int. Symp. on INnovations in Intelligent SysTems and Application (INISTA 2013), Albena, Bulgaria, June 19–21, 2013.
34. Dezert J, Tchamova A, Han D, Tacnet JM. Why Dempster's fusion rule is not a generalization of Bayes fusion rule, Fusion 2013 Proc., Istanbul, July 9-12, 2013.
35. Ellsberg D. Risk, ambiguity, and the Savage axioms, Quart. J. Econ. 1961; 75: 643-669.
36. Choquet G. Theory of capacities, Ann. Inst. Fourier, 1953; 5: 131-295.