Extend Transferable Belief Models with Probabilistic Priors

Chunlai Zhou
Computer Science Department
School of Information
Renmin University
Beijing, CHINA 100872
chunlai.zhou@gmail.com

Yuan Feng
Centre for Quantum Computation and Intelligent Systems
Faculty of Engineering and Information technology
University of Technology, Sydney
Broadway, NSW 2007 AUSTRALIA
Yuan.Feng@uts.edu.au

Abstract

In this paper, we extend Smets' *transferable belief model* (TBM) with *probabilistic* priors. Our first motivation for the extension is about evidential reasoning when the underlying prior knowledge base is Bayesian. We extend standard Dempster models with prior probabilities to represent beliefs and distinguish between two types of induced mass functions on an extended Dempster model: one for believing and the other essentially for decision-making. There is a natural correspondence between these two mass functions. In the extended model, we propose two conditioning rules for evidential reasoning with probabilistic knowledge base. Our second motivation is about the partial dissociation of betting at the pignistic level from believing at the credal level in TBM. In our *extended* TBM, we coordinate these two levels by employing the extended Dempster model to represent beliefs at the credal level. Pignistic probabilities are derived not from the induced mass function for believing but from the one for decision-making in the model and hence need not rely on the choice of frame of discernment. Moreover, we show that the above two proposed conditionings and marginalization (or coarsening) are consistent with pignistic transformation in the extended TBM.

1 INTRODUCTION

Reasoning about uncertainty is a fundamental issue for Artificial Intelligence [HALPERN, 2005]. Numerous approaches have been proposed, including the Dempster-Shafer theory of belief functions [SHAFER, 1976] (also called the theory of evidence or simply DS theory). Ever since the pioneering works by Dempster and Shafer, the theory of belief functions has become a powerful formalism in Artificial Intelligence for knowledge representation and decision-making.

The *transferable belief model* (TBM) is a model developed to justify the use of belief functions (including Dempster's rule of combination) to model someone's beliefs [SMETS AND KENNES, 1994]. A TBM $M = \langle(\Omega, m), Betp\rangle$ is a two-level mental model which distinguishes between two aspects of beliefs on a frame of discernment Ω, beliefs for weighted opinions, and beliefs for decision making. The two levels are: the credal level, where beliefs are entertained and represented by a mass function m, and the pignistic level, where beliefs are used to make decisions and quantified as a probability distribution $Betp_m$, which is derived from mass function m by the so-called *pignistic transformation* (usually denoted by $Betp$). The justification for the use of pignistic probabilities is usually linked to "rational" behavior exhibited by an ideal agent involved in some betting or decision contexts. But those probabilities do not represent the agent's beliefs; they are only the functions needed to derive the best decision. Let's consider a motivating example in TBM [SMETS AND KENNES, 1994].

Example 1.1 (Betting under total ignorance) Consider a guard in a huge power plant. On the emergency panel, alarms A_1 and A_2 are both on. The guard never heard about these two alarms. He takes the instruction book and discovers that A_1 is on iff circuit C is in state C_1 or C_2 and that alarm A_2 is on iff circuit D is in state D_1, D_2 or D_3. He never heard about these C and D circuits. There, his beliefs on the C circuit will be characterized by a "vacuous" belief function bel^{Ω_C} on space $\Omega_C = \{C_1, C_2\}$, i.e., a belief function whose mass function satisfies $m^{\Omega_C}(\Omega_C) = 1$ (this particular belief function is the one that represents the state of total ignorance). By the application of pignistic transformation, his pignistic probabilities will be given by

$$Betp^{\Omega_C}(C_1) = Betp^{\Omega_C}(C_2) = \tfrac{1}{2}.$$

Similarly, for the D circuit, the guard's belief bel^{Ω_D} on the space $\Omega_D = \{D_1, D_2, D_3\}$ will be vacuous, i.e., its corresponding mass function $m^{\Omega_D}(\Omega_D) = 1$, and the pignistic probabilities are

$$Betp^{\Omega_D}(D_1) = Betp^{\Omega_D}(D_2) = Betp^{\Omega_D}(D_3) = \tfrac{1}{3}.$$

Now by reading the next page on the manual, the guard discovers that circuits C and D are so made that whenever C is in the state C_1, circuit D is in state D_1 and vice versa. So he learns that C_1 and D_1 are equivalent (given what the guard knows) and that C_2 and (D_2 or D_3) are equivalent. In the TBM, this information does not modify his belief about which circuit is broken. Within the transferable belief model, the only requirement is that equivalent propositions should receive equal beliefs (it is satisfied as $bel^{\Omega_C}(C_1) = bel^{\Omega_D}(D_1) = 0$). Pignistic probabilities depend not only on these beliefs but also on *the structure of the betting frame*. In contrast, according to Bayesian approach, equivalent propositions should receive identical beliefs and therefore identical probabilities. However, $Betp^{\Omega_C}(C_1) = \frac{1}{2}$ and $Betp^{\Omega_D}(D_1) = \frac{1}{3}$ although $bel^{\Omega_C}(C_1) = bel^{\Omega_D}(D_1) = 0$.

The fact that the TBM can cope easily with such states of ignorance results from the partial dissociation between the credal and the pignistic levels. But this kind of separation between betting from believing makes the TBM vulnerable to Dutch books in decision-making [SNOW, 1998].

In this paper, we extend Smets' TBM with a probabilistic prior to coordinate reasoning at the credal and pignistic levels. Our first motivation is about evidential reasoning when the underlying prior knowledge base is Bayesian. In order to incorporate the influence of the Bayesian knowledge base, we extend standard Dempster models, which are used for representing belief functions, with probabilistic priors. For an extended Dempster model M with a prior probability pr, there are two induced mass functions. The first one m_D is derived in the standard way from the Dempster part D of M without the prior probability and hence complies with the well-known DS theory, especially with Dempster's rule of combination. The second m_M is induced by combining m_D with the prior probability pr. Conversely, m_D can be obtained from m_M by removing the influence of pr. So, there is a natural correspondence between m_D and m_M. However, these two mass functions are essentially different: m_D measures the belief update and m_M absolute belief or weighted opinion. We propose a new combination rule for the mass functions m_M's which incorporate prior probabilities. The new combination rule is shown to be parallel to Dempster's rule for the mass functions m_D's without the influence of prior probabilities. According to the new combination rule, we provide two *prediction-style* conditioning rules: one for *certain* conditioning knowledge and the other for *uncertain* knowledge.

Our second motivation is to coordinate reasoning at the credal and pignistic levels. We extend Smets' TBM by employing an extended Dempter model M to represent beliefs at the credal level and provide a corresponding generalized pignistic transformation $Betp$ for this extended TBM. We prove that the above two new conditioning rules in M are consistent with this pignistic transformation. In our extended TBM, since beliefs are represented by the induced mass function m_D of the Dempster part of M, they are insensitive to the choice of frame. Pignistic probabilities are derived not from the induced mass function m_D of the Dempster part of M but from the induced mass function m_M, which have incorporated the prior probability pr. We show by transforming the prior probability that pignistic probabilities obtained in this way need not rely on the choice of frame of discernment.

2 BASIC DEFINITIONS AND NOTIONS

Let Ω be a frame of discernment and $\mathcal{A} = 2^{\Omega}$ be the Boolean algebra of events. A *mass function* (or *mass assignment*) is a mapping $m : \mathcal{A} \to [0,1]$ satisfying $\sum_{A\in\mathcal{A}} m(A) = 1$. A mass function m is called *normal* if $m(\emptyset) = 0$. Without further notice, all mass functions in this paper are assumed to be normal. A set is called *focal* if $m(A) > 0$. A mass function m is called *categorical* if it has only one focal set. A *belief function* is a function $bel : \mathcal{A} \to [0,1]$ satisfying the following conditions:

1. $bel(\emptyset) = 0$, $bel(\Omega) = 1$; and
2. $bel(\bigcup_{i=1}^{n} A_i) \geq \sum_{\emptyset\neq I\subseteq\{1,\cdots,n\}} (-1)^{|I|+1} bel(\cap_{i\in I} A_i)$ where $A_i \in \mathcal{A}$ for all $i \in \{1,\cdots,n\}$.

A mapping $f : \mathcal{A} \to [0,1]$ is a belief function if and only if its Möbius transform is a mass function [SHAFER, 1976]. In other words, if $m : \mathcal{A} \to [0,1]$ is a mass function, then it determines a belief function $bel : \mathcal{A} \to [0,1]$ as follows: $bel(A) = \sum_{B\subseteq A} m(B)$ for all $A \in \mathcal{A}$. Moreover, given a belief function bel, we can obtain its corresponding mass function m as follows: $m(A) = \sum_{B\subseteq A} (-1)^{|A\setminus B|} bel(B)$ for all $A \in \mathcal{A}$. Intuitively, for a subset event A, $m(A)$ measures the belief that an agent commits *exactly* to A, not the total belief $bel(A)$ that an agent commits to A. The corresponding *plausibility function* $pl : 2^{\Omega} \to [0,1]$ is dual to bel in the sense that $pl(A) = 1 - bel(\overline{A})$ for all $A \subseteq \Omega$. If m_1 and m_2 are two mass functions on Ω induced by two *independent* evidential sources, the combined mass function is calculated according to *Dempster's rule of combination*: for any $C \subseteq \Omega$,

$$(m_1 \oplus m_2)(C) = \frac{\sum_{A\cap B=C} m_1(A)m_2(B)}{\sum_{A\cap B\neq\emptyset} m_1(A)m_2(B)} \qquad (1)$$

When an event E is observed, then the conditional mass function of m is obtained according to *Dempster conditioning*: for any $C \subseteq \Omega$,

$$m(C \mid E) = \frac{\sum_{B\cap E=C} m(B)}{pl(E)} \qquad (2)$$

A transferable belief model $M = \langle(\Omega, m), Betp\rangle$ [SMETS AND KENNES, 1994] is a two-level mental model: the credal level where beliefs are represented by a mass function m, and the pignistic level where decisions are made by maximizing expected utility. Hence we must build a probability distribution to compute these expectations. This probability distribution is *based on* the agent's beliefs, but should not be understood as representing the agent's beliefs. It is just a probability distribution *derived from* the mass function through *pignistic transformation* $Betp$. The pignistic transformation for the above mass function m is given by

$$Betp_m(\{\omega\}) := \sum_{\omega \in B \subseteq \Omega} \frac{1}{|B|} m(B) \text{ for any } \omega \in \Omega.$$

Note that $Betp_m$ is a probability distribution on Ω and is called a *pignistic probability distribution*. When the context is clear, we usually use m to denote the belief model M.

In order to show the sensitivity of pignistic transformation to the choice of frames of discernment, we need to set up a setting in terms of *refinements and coarsenings* of frames of discernment. The idea that one frame Ω of discernment is obtained from another frame Θ by splitting some or all of the elements of Θ may be represented mathematically by specifying, for each $\theta \in \Theta$, the subset $\omega(\{\theta\})$ of Ω consisting of those possibilities into which θ has been split. Such a mapping ω is called a *refining*. Whenever $\omega : 2^\Theta \to 2^\Omega$ is a refining, we call Ω a *refinement* of Θ and Θ a *coarsening* of Ω. In this paper, we are particularly interested in the case when Θ is the set of equivalence classes with respect to some partition Π of Ω. So the mapping $\omega(\{\Pi(w)\}) = \Pi(w)$ for each $w \in \Omega$ is a refinement and Θ is a coarsening of Ω where $\Pi(w)$ is the equivalence class of w. We denote this special coarsening Θ of Ω as Ω/Π. On the other hand, Ω/Π may be regarded as a subalgebra $\mathcal{B}$ of the powerset of Ω with the set of *atoms* forming the partition Π of Ω. In the following sections, we won't distinguish between Ω/Π and $\langle\Omega, \mathcal{B}\rangle$. For each $A \subseteq \Omega$, we define $\mathbf{B}(A) := \bigcap\{B \in \mathcal{B} : A \subseteq B\}$. In other words, $\mathbf{B}(A)$ is the least element of $\mathcal{B}$ that contains A as a subset and hence is called the *upper approximation* of A in $\mathcal{B}$. For example, $\Pi = \{\{w_1\}, \{w_2, w_3\}, \{w_4, w_5, w_6\}\}$ is a partition of $\Omega = \{w_1, w_2, w_3, w_4, w_5, w_6\}$. Then the associated subalgebra $\mathcal{B}$ consists of the sets $\bigcup_{B \subseteq \Pi} B$ with the atoms $\{w_1\}, \{w_2, w_3\}$, and $\{w_4, w_5, w_6\}$ in $\mathcal{B}$. If $A = \{w_1, w_3, w_5\}$, then $\mathbf{B}(A) = \Omega$.

Let $\langle\Omega, \mathcal{B}\rangle$ be a coarsening of Ω where $\mathcal{B}$ is a subalgebra of the powerset 2^Ω with its atoms forming a partition of Ω. Each element B of $\mathcal{B}$ is a disjoint union of some atoms in $\mathcal{B}$. Suppose that $bel : 2^\Omega \to [0, 1]$ is a belief function on Ω with m as its corresponding mass function. Then *the derived mass function $m_\mathcal{B}$ on the coarsening $\langle\Omega, \mathcal{B}\rangle$* can be obtained through the formula: for any $B \in \mathcal{B}$, $m_\mathcal{B}(B) = \sum_{\mathbf{B}(A)=B, A\subseteq\Omega} m(A)$. Let $bel_\mathcal{B}$ denote the corresponding belief function. It is easy to check that, for any $B \in \mathcal{B}$, $bel_\mathcal{B}(B) = bel(B)$. Intuitively, $bel_\mathcal{B}$ is the derived belief function on the coarsening frame of discernment with less distinctions. The beliefs in the same propositions in these two different frames with different distinctions should be the same as each other. In this sense, believing in terms of belief functions is insensitive to the choice of frame of discernment.

3 EXTENDED DEMPSTER MODELS

In order to motivate our work of extending Smets' transferable belief models with probabilistic priors, we first represent belief functions through Dempster models.

3.1 EXTENDED DEMPSTER MODELS

Definition 3.1 A *Dempster model* is a tuple $\langle(U, Pr), \Gamma, \Omega\rangle$ where (U, Pr) is a probability space and Γ is a multivalued mapping from U to Ω, i.e., a mapping from U to 2^Ω, the powerset of Ω. $\lhd$

The multivalued mapping Γ is essentially a random subset on Ω, and it induces a mass function m on Ω: $m(A) := Pr(\Gamma^{-1}(A))$ for any $A \subseteq \Omega$. We have the corresponding belief function $Bel(A) = \sum_{B\subseteq A} Pr(\Gamma^{-1}(B))$. Conversely, any mass function on Ω can be represented as the induced mass function of some Dempster model. Before we extend Dempster models with probabilistic priors on Ω, we use the well-known three prisoner paradox to show the necessity of the probabilistic priors.

Example 3.2 (The Three Prisoners Paradox [HALPERN, 2005]) Of three prisoners a, b and c, only one of them is to be executed but a does not know which one. He therefore says to the jailer, "Since either b and c is certainly going to be declared innocent, you will give me no information about my chances if you give me the name of one man, either b or c, who is going to be freed." Accepting this argument, the jailer truthfully replies,"b will be freed." Thereupon a feels sad because of the Bayesian conditioning on $U := \{a, b, c\}$: before the jailer replied, his own chances of being executed was one-third, but afterwards there are only two people, himself and c, who could be the one being executed, and so his chances of execution increases and is one-half.

Is a justified in believing that his chances of being executed have increased? Now we formulate this problem in the framework of a Dempter model. Consider the set of all possible outcomes: $\Omega := \{(a, b), (a, c), (b, c), (c, b)\}$ where, for example, (a, b) means that a is to be executed and the jailer says that b will be freed. Suppose that at first a assumes that the initial decision as to who will be executed is made at random but assumes *nothing* about how the jailer will act except that he will tell the truth. Let the random choice of who will be executed be represented by the probability space (U, Pr) where Pr is the uniform distribution on U. A multivalued mapping $\Gamma : U \to 2^{\Omega}$ for delineating the possible outcomes when a, b or c is to be executed is given by: $\Gamma(a) = \{(a,b),(a,c)\}, \Gamma(b) = \{(b,c)\}, \Gamma(c) = \{(c,b)\}$. So the induced mass function m at the credal level is given by: $m(\{(a,b),(a,c)\}) = m(\{(b,c)\}) = m(\{(c,b)\}) = \frac{1}{3}$. Let E_a denote the event that a will be executed and J_b the event that the jailer says that b will be freed. Then $E_a = \{(a,b),(a,c)\}$ and $J_b = \{(a,b),(c,b)\}$. According to Dempster's rule of conditionalization, we get that $Bel(E_a \mid J_b) = Pl(E_a \mid J_b) = \frac{1}{2}$. So Dempster's conditioning provides the same answer as that by the above a's conditioning on the "naive" space U according to Bayesian rule [GRÜNWALD AND HALPERN, 2003]. By applying Smets' pignistic transformation, we obtain its probability distribution at the pignistic level: $Betp_m(a,b) = Betp_m(a,c) = 1/6$ and $Betp_m(b,c) = Betp_m(c,b) = 1/3$.

More generally, we may assume that the jailer will tell the truth and a's knowledge about the jailer's preference over his possible choices is formulated by a probabilistic prior on Ω, which is *independent* of the assumption that the executed prisoner is chosen at random. Now we extend standard Dempster models by incorporating this kind of probabilities and express the induced beliefs at the credal level.

Definition 3.3 *An extended Dempster-model* $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$ is a Dempster model $\langle (U, Pr), \Gamma, \Omega\rangle$ plus a prior probability pr on Ω where pr is independent of Γ with respect to Pr. $\lhd$

Now we explain this independence through a representation result of extended Dempster models.

Lemma 3.4 *Every extended Dempster model* $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$ *can be represented as a standard Dempster model* $\langle (U', Pr'), \Gamma', \Omega\rangle$ *with an additional mapping* γ' *from* U' *to* Ω *for some probability space* (U', Pr') *and some multivalued mapping* Γ' *from* U' *to* Ω.

Proof. For a given extended Dempster model $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$, we define a new probability space (U', Pr'), which is essentially the Cartesian product of (U, Pr) and (Ω, pr), as follows:

- $U' = U \times \Omega$;
- $Pr'(x, y) = Pr(x)pr(y)$ for any $(x, y) \in U'$.

Further we define a multivalued mapping $\Gamma' : U' \to 2^{\Omega}$ and a mapping $\gamma' : U' \to \Omega$ as follows:

- $\Gamma'(x, y) = \Gamma(x)$,
- $\gamma'(x, y) = y$ for any $(x, y) \in U'$.

It is easy to check that $Pr'((\Gamma')^{-1}(A)) = Pr(\Gamma^{-1}(A))$, and $Pr'((\gamma')^{-1}(A)) = pr(A)$ for any $A \subseteq \Omega$. QED

So, in the following sections of this paper, we won't distinguish these two forms of extended Dempster models and will sometimes write an extended Dempster model as $M = \langle (U, Pr), \Gamma, \gamma, \Omega\rangle$ where $\langle (U, Pr), \Gamma, \Omega\rangle$ is a standard Dempster model and γ is a mapping from U to Ω. In M, the prior probability pr is obtained by $pr(A) = Pr(\{u \in \Omega : \gamma(u) \in A\})$. In this paper, $\Gamma = A$ is shorthand for the event $\{u \in U : \Gamma(u) = A\}$, $\gamma \in A$ for $\{u \in U : \gamma(u) \in A\}$ and $\gamma \in \Gamma$ denotes $\{u \in U : \gamma(u) \in \Gamma(u)\}$. In M, the independence of the prior probability pr of the multivalued mapping Γ with respect to Pr means the independence of γ and Γ: for any subsets A and B of Ω,

$$Pr((\Gamma = A) \cap (\gamma \in B)) = Pr(\Gamma = A)Pr(\gamma \in B).$$

Just as in a Dempster model, we associate each extended Dempster model $M = \langle (U, Pr), \Gamma, \gamma, \Omega\rangle$ with a mapping $m_M : 2^{\Omega} \to [0, 1]$ which incorporates the mapping γ as follows:

$$m_M(A) := Pr(\Gamma = A \mid \gamma \in \Gamma) \qquad (3)$$

It is easy to see that, since Γ and γ are independent with respect to Pr, $Pr(\gamma \in \Gamma) = \sum_{A \subseteq \Omega} Pr((\gamma \in A) \cap (\Gamma = A)) = \sum_{A \subseteq \Omega} Pr(\gamma \in A)Pr(\Gamma = A)$. And $Pr(\gamma \in \Gamma)$ is used to measure the degree of *consistency* of the evidence represented by Γ with the prior represented by γ. It follows that

$$\begin{aligned}
\sum_{A \subseteq \Omega} m_M(A) &= \sum_{A \subseteq \Omega} Pr(\Gamma = A \mid \gamma \in \Gamma) \\
&= \sum_{A \subseteq \Omega} \frac{Pr((\Gamma = A) \cap (\gamma \in \Gamma))}{Pr(\gamma \in \Gamma)} \\
&= \sum_{A \subseteq \Omega} \frac{Pr((\Gamma = A) \cap (\gamma \in A))}{Pr(\gamma \in \Gamma)} \\
&= \sum_{A \subseteq \Omega} \frac{Pr(\Gamma = A)Pr(\gamma \in A)}{Pr(\gamma \in \Gamma)} \\
&= 1
\end{aligned}$$

So such a defined mapping m_M is actually a mass function on Ω and is called *the induced mass function* of M.

Next we show that extended Dempster models are as expressive as standard Dempster models in the sense that any mass function m on Ω can be represented as the induced mass function m_M of some extended Dempster model M. We prove a lemma which implies this expressiveness result.

Lemma 3.5 *For any mass function* m *and probability distribution* pr *on* Ω*, there is an extended Dempster model* $M = \langle (U, Pr), \Gamma, \gamma, \Omega\rangle$ *such that*

1. $m_M(A) = m(A)$ *for each* $A \subseteq \Omega$ *where* m_M *is the induced mass function of* M*;*

2. $pr(A) = Pr(\gamma^{-1}(A))$ *for any* $A \subseteq \Omega$*.*

Proof. Given a mass function m and a probability function pr on Ω, we define a mapping $m_D : 2^\Omega \to [0,1]$ as follows: for any $A \subseteq \Omega$,

$$m_D(A) = \frac{\frac{m(A)}{pr(A)}}{\sum_{A\subseteq\Omega} \frac{m(A)}{pr(A)}} \quad (4)$$

Since $\sum_{A\subseteq\Omega} m_D(A) = 1$, m_D is a mass function on Ω. It follows that there is a standard Dempster model $\langle (U_D, Pr_D), \Gamma_D, \Omega\rangle$ such that $m_D(A) = Pr_D(\Gamma_D^{-1}(A))$ for any $A \subseteq \Omega$. From the proof of Lemma 3.4, we know that the extended Dempster model $\langle (U_D, Pr_D), \Gamma_D, (\Omega, pr)\rangle$ with the prior probability pr can be represented as a Dempster model $\langle (U, Pr), \Gamma, \Omega\rangle$ with γ as a mapping from U to Ω. For this equivalent representation $M := \langle (U, Pr), \Gamma, \gamma, \Omega\rangle$ of the extended model, we have that

- $Pr(\Gamma = A) = Pr(\Gamma^{-1}(A)) = Pr_D(\Gamma_D^{-1}(A)) = m_D(A)$;
- $Pr(\gamma \in A) = Pr(\gamma^{-1}(A)) = pr(A)$.

It follows that

$$\begin{aligned} Pr(\gamma \in \Gamma) &= \sum_{A\subseteq\Omega} Pr(\Gamma = A) Pr(\gamma \in A) \\ &= \sum_{A\subseteq\Omega} m_D(A) pr(A) \\ &= \sum_{A\subseteq\Omega} \frac{\frac{m(A)}{pr(A)}}{\sum_{A\subseteq\Omega}\frac{m(A)}{pr(A)}} pr(A) \\ &= \frac{\sum_{A\subseteq\Omega} m(A)}{\sum_{A\subseteq\Omega}\frac{m(A)}{pr(A)}} \\ &= \frac{1}{\sum_{A\subseteq\Omega}\frac{m(A)}{pr(A)}} \end{aligned}$$

So we have that the induced mass function m_M:

$$\begin{aligned} m_M(A) &= Pr(\Gamma = A \mid \gamma \in \Gamma) \\ &= \frac{Pr(\gamma \in A) Pr(\Gamma = A)}{Pr(\gamma \in \Gamma)} \\ &= m(A). \end{aligned}$$

QED

From the above proof, we know that, for any extended Dempster model $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$, there are two *induced* mass functions on Ω: the induced mass function $m_D(A)(= Pr(\Gamma = A))$ in the part $D := \langle (U, Pr), \Gamma, \Omega\rangle$, which is actually a standard Dempster model, and the induced mass function $m_M(A)(= Pr(\Gamma = A \mid \gamma \in \Gamma))$ of M. m_D measures the *belief update* and is called *basic certainty value*, while m_M measures *absolute belief.* This distinction is crucial to our following extension of Smets' transferable belief models with probabilistic priors. In our extended belief models, we use mass functions m_D for believing and mass functions m_M for decision-making. Mass functions for believing are based on the theory of evidence while mass functions for decision-making are essentially Bayesian and hence consistent with pignistic transformation. Basic certainty values are used in the probabilistic interpretation of CF in MYCIN [HECKERMAN, 1985]. For a given extended Dempster model $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$, there is a one-to-one correspondence (see Eqs.(3) and (4)) between the induced mass function m_M of M and the induced m_D in the standard-Dempster-model part $D = \langle (U, Pr), \Gamma, \Omega\rangle$. Assume that pr is given. The induced mass function m_M can be expressed in terms of m_D as follows: $m_M(A) = \frac{pr(A)m_D(A)}{\sum_{A\subseteq\Omega} pr(A)m_D(A)}$. We denote this expression as $m_M = m_D \circ pr$. Moreover, m_D can be expressed in terms of m_M: $m_D(A) = \frac{m_M(A)/pr(A)}{\sum_{A\subseteq\Omega} m_M(A)/pr(A)}$, which is denoted as $m_D = m_M/pr$. From the proof of Lemma 3.5, we know that the two operations $\circ$ and $/$ are reverse to each other in the sense that $(m_D \circ pr)/pr = m_D$ and $(m_M/pr) \circ pr = m_M$.

Let $M_1 = \langle (U_1, Pr_1), \Gamma_1, (\Omega, pr)\rangle$ and $M_2 = \langle (U_2, Pr_2), \Gamma_2, (\Omega, pr)\rangle$ be two extended Dempster models representing two *independent* bodies of evidence on the same probability space (Ω, pr). Let m_{D_1} and m_{D_2} be the two induced mass functions for belief updates in the standard-Dempster-model parts $D_1 = \langle (U_1, Pr_1), \Gamma_1, \Omega\rangle$ and $D_2 = \langle (U_2, Pr_2), \Gamma_2, \Omega\rangle$, respectively. As in Dempster models, m_{D_1} and m_{D_2} are combined according to the well-known *Dempster's rule*: for any $C \subseteq \Omega$,

$$(m_{D_1} \oplus_D m_{D_2})(C) = \frac{\sum_{A_1\cap A_2 = C} m_{D_1}(A_1) m_{D_2}(A_2)}{K_D} \quad (5)$$

where $K_D = \sum_{A_1\cap A_2\neq\emptyset} m_{D_1}(A_1)\cdot m_{D_2}(A_2)$ is the normalization factor. So the combination $(m_{D_1} \oplus_D m_{D_2})$ also measures belief update for the same probability space (Ω, pr). Let m_{M_1} and m_{M_2} denote the two induced mass functions for *absolute belief* on the extended Dempster models M_1 and M_2, respectively. Now we provide a *new* combination rule for the extended Dempster models as follows: for any $C \subseteq \Omega$,

$$(m_{M_1} \oplus_M m_{M_2})(C) = \frac{\sum_{A_1\cap A_2 = C} \frac{pr(A_1\cap A_2)}{pr(A_1)Pr(A_2)} m_{M_1}(A_1) m_{M_2}(A_2)}{K_M} \quad (6)$$

where

$$K_M := \sum_{A_1\cap A_2\neq\emptyset} \frac{pr(A_1\cap A_2)}{pr(A_1)Pr(A_2)} m_{M_1}(A_1) m_{M_2}(A_2)$$

is the normalization factor. The following proposition says that the new combination $\oplus_M$ of mass functions for absolute beliefs is consistent with the Dempster combination $\oplus_D$ of their corresponding mass functions for belief updates.

Proposition 3.6 *The combination $m_{M_1} \oplus_M m_{M_2}$ of m_{M_1} and m_{M_2} for absolute belief satisfies the following property:*

$$m_{M_1} \oplus_M m_{M_2} = (m_{D_1} \oplus_D m_{D_2}) \circ pr. \tag{7}$$

Proof. For any $A \subseteq \Omega$,

$$\begin{aligned} & \sum_{A\subseteq\Omega}[pr(A)\sum_{A_1\cap A_2=A} m_{D_1}(A_1)m_{D_2}(A_2)] \\ = & \sum_{A\subseteq\Omega}[pr(A)\sum_{A_1\cap A_2=A} \frac{\frac{m_{M_1}(A_1)}{pr(A_1)}}{K_1} \frac{\frac{m_{M_2}(A_2)}{pr(A_2)}}{K_2}] \\ = & \sum_{A\subseteq\Omega}[\sum_{A_1\cap A_2=A} \frac{pr(A)}{pr(A_1)pr(A_2)} \frac{m_{M_1}(A_1)}{K_1} \frac{m_{M_2}(A_2)}{K_2}] \end{aligned}$$

where $K_1 = \sum_{A\subseteq\Omega} \frac{m_{M_1}(A)}{pr(A)}$ and $K_2 = \sum_{A\subseteq\Omega} \frac{m_{M_2}(A)}{pr(A)}$. The first equality comes from Eq.(3) and the second from Eq.(4). So we have

$$\begin{aligned} & ((m_{D_1} \oplus_D m_{D_2}) \circ pr)(A) \\ = & \frac{(m_{D_1} \oplus_D m_{D_2})(A)pr(A)}{\sum_{A\subseteq\Omega}(m_{D_1} \oplus_D m_{D_2})(A)pr(A)} \\ = & \frac{pr(A)\sum_{A_1\cap A_2=A} m_{D_1}(A_1)m_{D_2}(A_2)}{\sum_{A\subseteq\Omega}[pr(A)\sum_{A_1\cap A_2=A} m_{D_1}(A_1)m_{D_2}(A_2)]} \\ = & \frac{\sum_{A_1\cap A_2=A} \frac{pr(A)}{pr(A_1)pr(A_2)} \frac{m_{M_1}(A_1)}{K_1} \frac{m_{M_2}(A_2)}{K_2}}{\sum_{A\subseteq\Omega}[\sum_{A_1\cap A_2=A} \frac{pr(A)}{pr(A_1)pr(A_2)} \frac{m_{M_1}(A_1)}{K_1} \frac{m_{M_2}(A_2)}{K_2}]} \\ = & \frac{\sum_{A_1\cap A_2=A} \frac{pr(A)}{pr(A_1)pr(A_2)} m_{M_1}(A_1)m_{M_2}(A_2)}{\sum_{A\subseteq\Omega}[\sum_{A_1\cap A_2=A} \frac{pr(A)}{pr(A_1)pr(A_2)} m_{M_1}(A_1)m_{M_2}(A_2)]} \\ = & (m_{M_1} \oplus_M m_{M_2})(A) \end{aligned}$$

QED

3.2 TWO CONDITIONING RULES

There are two types of conditioning in Bayesian probability theory [DUBOIS AND DENOEUX, 2012]. The first one is known as *revision*. Given a probability function Pr (which usually is a subjective probability), one learns a hard evidence in terms of a sure event C. The problem is to determine the new subjective probability measure Pr', such that $Pr'(C) = 1$, according to some minimal change principle. The other one is called *prediction*. When dealing with prediction, we have at our disposal a model of uncertainty in the form of a probability measure Pr issued from a representative set of statistical data. Moreover, given the knowledge C on the current state of the world, we combine this knowledge with the belief model Pr and predict some property A of the current world with its associated degree of belief $Pr(A|C)$. For belief functions, however, these two types of conditioning are *essentially* different and the mainstream literature is a revision theory of handling singular uncertain evidence [SHAFER, 1976], not so much an extension of Bayesian statistical prediction, although Dempster's pioneering works on upper and lower probabilities are motivated by statistical reasoning. The well-known Dempster's rule of conditioning, which is a special case of Dempster's rule of combination, can be viewed as a revision process. In general, prediction cannot be achieved using Dempster conditioning [DUBOIS AND DENOEUX, 2012]. Fagin and Halpern [FAGIN AND HALPERN, 1991] and Jaffray [JAFFRAY, 1992] provided two prediction-style conditioning rules which generalize Bayesian prediction by interpreting belief functions as inner measures and lower probabilities, respectively.

In this paper, we provide a *new* prediction-style conditioning rule which is *consistent* with the revision style of conditioning performed according to Dempster's rule of conditioning. For a given extended Dempster model $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$, the prediction conditioning is carried out for the induced mass function m_M while the revision rule is for the induced mass function m_D of the part $D = \langle (U, Pr), \Gamma, \Omega\rangle$ without the prior probability pr. These two conditioning are consistent in the sense of Proposition 3.6 when the certain knowledge C is represented by a categorical mass function with C as its only focal set. Our following rule for prediction-style conditioning provides a formula of how to compute conditional belief $m_M(\cdot|C)$ on the knowledge C. Generally, for each $E \subseteq \Omega$, we transfer a proportion $r_E \cdot m_M(E)$, where $0 \leq r_E \leq 1$, to $E \cap C$ and $(1 - r_E) \cdot m_M(E)$ to $E \cap \overline{C}$ [DUBOIS AND DENOEUX, 2012]. In particular, when $E\cap C = \emptyset$, we set $r_E = 0$, which contributes nothing to $E\cap C$; when $E \subseteq C$, we set $r_E = 1$ and leave the whole $m_M(E)$ to E. According to this idea, we obtain a *general* formula for conditioning:

$$m(A|C) := \frac{\sum_{E\cap C=A} r_E\cdot m_M(E)}{\sum_{E\cap C\neq\emptyset} r_E\cdot m_M(E)}.$$

It is easy to check that

- $m(A|C)$ is exactly the Dempster rule of conditioning in the case when $r_E = 1$ iff $E \cap C \neq \emptyset$;
- $m(A|C)$ is exactly the geometric rule of conditioning in the case when $r_E = 0$ iff $E \not\subseteq C$.

In this paper, we define a *new* rule for prediction-style conditioning between the above two by setting $r_E = \frac{pr(E\cap C)}{pr(E)}$:

$$m_M(A|C) = \frac{\sum_{E\cap C=A} \frac{pr(E\cap C)}{pr(E)} \cdot m_M(E)}{K}$$

where $K = \sum_{E\cap C\neq\emptyset} \frac{pr(E\cap C)}{pr(E)} \cdot m_M(E)$ is a normalization factor. It is easy to see that the conditioning $m_M(\cdot|C)$ is a special case of the new combination rule for absolute beliefs (Eq.(6)) when the knowledge C is represented by a categorical mass function with C as its only focal set.

Example 3.7 (Continue with Example 3.2) Assume that the jailer's preference over possible choices according to a's knowledge is represented by a uniform distribution pr on Ω. We obtain the induced mass function m_M : $m_M(\{(a,b),(a,c)\}) = 1/2$, $m_M(b,c) = 1/4$, $m_M(c,b) = 1/4$ and hence the corresponding beliefs $Bel_M(E_a|J_b) = 1/2 = Pl_M(E_a|J_b)$.

By using our definition of conditioning rule $m_M(\cdot|C)$ on the *certain* knowledge C, we define its corresponding Jeffrey's rule when the prior knowledge is uncertain and is represented by a probability function pr_e on a coarsening of Ω: $(\Omega, \mathcal{B})$ where $\mathcal{B}$ is a subalgebra of the powerset 2^Ω with its atoms forming a partition of Ω. Let $At(\mathcal{B})$ denote the set of atoms of $\mathcal{B}$. A mass function m'_M on $(\Omega, 2^\Omega)$ is said to be obtained from m_M by *belief kinematics* on $(\Omega, \mathcal{B})$ if, for any $B \in At(\mathcal{B})$,

$$m_M(A|B) = m'_M(A|B) \text{ for all } A \subseteq \Omega. \qquad (8)$$

m'_M is called the mass function proposed by *Jeffrey's rule* if it is obtained as follows: for any $A \subseteq \Omega$,

$$m'_M(A) = \sum_{B\in At(\mathcal{B})} m_M(A|B)pr_e(B), \qquad (9)$$

Intuitively, the above principle of belief kinematics on $(\Omega, \mathcal{B})$ says that, even though m_M and m'_M may disagree on propositions on $(\Omega, \mathcal{B})$, they agree on their relevance to every proposition $A \subseteq \Omega$.

4 EXTENDED TRANSFERABLE BELIEF MODELS WITH PROBABILISTIC PRIORS

Definition 4.1 Let m_M be the induced mass function of an extended Dempster model $\langle (U, Pr), \Gamma, (\Omega, pr)\rangle$. Its associated *pignistic probability function* $Betp_{m_M}$ on Ω is defined as follows: for any $A \subseteq \Omega$,

$$Betp_{m_M}(A) = \sum_{E\subseteq\Omega} m_M(E)\frac{pr(E\cap A)}{pr(E)} \qquad (10)$$

The transformation between m_M and $Betp_{m_M}$ is called *the generalized pignistic transformation*. When the context is clear, we simply call it pignistic transformation. $\lhd$

Since $m_M(A) = \frac{m_D(A)pr(A)}{\sum_{E\subseteq\Omega} m_D(E)pr(E)}$, the pignistic probability function can be expressed in terms of the mass function m_D for belief updates: $Betp_{m_M}(A) = \frac{\sum_{E\subseteq\Omega} m_D(E)pr(E\cap A)}{\sum_{E\subseteq\Omega} m_D(E)pr(E)}$. Note that Smets' pignistic transformation is not a special case of the above defined generalized pignistic transformation when the prior probability pr is the uniform distribution on Ω.

Example 4.2 (Continue with Example 3.7) We may *complete* the above partial model $\langle\Omega, m\rangle$ and obtain a probabilistic model according to the uniform distribution pr. When a is to be executed, the "chances" of the jailer's saying b or c are equal. So a will distribute the mass $m(E_a)$ equally between (a,b) and (a,c). Then we have $m(b,c) = m(c,b) = 1/3$ and $m(a,b) = m(a,c) = 1/6$, which is exactly the probability function according to Smets' pignistic transformation. Also we obtain the corresponding beliefs $Bel(E_a|J_b) = 1/3 = Pl(E_a|J_b)$, which is the same as expected according to Bayesian reasoning. However, this distribution is not the same as the one obtained according to the above generalized pignistic transformation in Eq.(10). Instead, $Betp_{m_M}(a,b) = Betp_{m_M}(a,c) = Betp_{m_M}(c,b) = Betp_{m_M}(b,c) = 1/4$.

Assume that $m_1, \cdots, m_l$ are induced mass functions on (Ω, pr) and $p_1, \cdots, p_l$ are non-negative numbers such that $\sum_{i=1}^{l} p_i = 1$. It is interesting to note that pignistic transformation $Betp$ satisfies the following *linearity property*:

$$Betp(\sum_{i=1}^{l} p_i m_i) = \sum_{i=1}^{l} p_i Betp(m_i). \qquad (11)$$

This property is both the major requirement that led Smets to the solution for the pignistic transformation [SMETS, 2005] and the crucial step to show the commutativity of the diagrams in the following Theorem 4.5. In addition to the linearity property, Smets proposed other requirements: credal-pignistic link, projectivity, continuity, efficiency, anonymity and impossible event [SMETS, 2005]. These requirements lead to the *unique* solution of Smets' pignistic transformation. One can check that our *generalized* pignistic transformation meets all these requirements except the anonymity one. The anonymity requirement rephrases a general form of insufficient reason principle and hence is equivalent to the constraint that the prior probability in the extended Dempster model is uniform.

Definition 4.3 An *extended transferable belief model* (ETBM) $\mathbf{M} = \langle M, Betp\rangle$ is a two level mental model: the credal level where beliefs are represented by an extended Dempster model $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$, and the pignistic level where the pignistic probability function is obtained from the induced mass function m_M of M by the generalized pignistic transformation $Betp$. $\lhd$

Smets' transferable belief model is a special case of the above defined extended transferable belief model when the prior probability is uniform.

Theorem 4.4 *Let $Cond^p$ and $Cond$ denote the above defined prediction style conditioning operator for mass functions and the standard one for Bayesian probability functions, respectively. We have that the following diagram commutes:*

$$\begin{array}{ccc} (m_M, C) & \xrightarrow{Cond^p} & m_M(\cdot|C) \\ \Big\downarrow {\scriptstyle Betp} & & \Big\downarrow {\scriptstyle Betp} \\ (Pr, C) & \xrightarrow{Cond} & Pr(\cdot|C) \end{array}$$

Theorem 4.5 *Let m_M, pr_e and m'_M be as in Eq.(9). Probability measures Pr and Pr' denote the pignistic probability functions of m_M and m'_M, respectively. Then the following diagram commutes:*

$$\begin{array}{ccc} (m_M, pr_e) & \xrightarrow{J} & m'_M \\ \Big\downarrow {\scriptstyle Betp} & & \Big\downarrow {\scriptstyle Betp} \\ (Pr, pr_e) & \xrightarrow{J} & Pr' \end{array}$$

where the first J is the Jeffrey conditioning for mass functions as defined in Eq.(9) and the second J denotes the standard Jeffrey conditioning in Bayesian probability theory. In other words, our Jeffrey's rule is nothing but the linearity property in Eq.(11).

The above two theorems tell us that in extended transferable belief models the two new conditioning rules are consistent with pignistic transformation; in other words, the following two strategies are equivalent: we can revise the pignistic probabilities which are transformed from the prior beliefs with Bayes rule applied to the (certain or uncertain) knowledge, or revise the prior beliefs at the credal level by the above two conditioning rules and recompute the pignistic transformation.

However, from Example 1.1, we know that marginalization or coarsening is *inconsistent* with pignistic transformation. That is to say, pignistic transformation is sensitive to the choice of frame of discernment, which causes the partial dissociation between the credal and pignistic levels. In the remainder of this section, we show that, in an extended TBM, these two levels can be coordinated by transforming its prior probability function.

Let $M = \langle (U, Pr), \Gamma, (\Omega, pr)\rangle$ be a given extended Dempster model. Let m_M and m_D denote the induced mass functions for absolute beliefs and belief updates, respectively. Let $(\Omega, \mathcal{B})$ be a coarsening of Ω where $\mathcal{B}$ is a subalgebra of the powerset of Ω with its atoms $\mathcal{C} := \{B_1, \cdots, B_n\}$ forming a partition of Ω. So each element of $\mathcal{B}$ is a disjoint union of some atoms from the basis $\mathcal{C}$. Correspondingly, the coarsening Γ^C of the multivalued mapping Γ must be defined in the following way: for any $u \in U$, $\Gamma^C(u) = \mathbf{B}(\Gamma(u))$ where $\mathbf{B}$ denotes the operation of taking upper approximation in the subalgebra $\mathcal{B}$. The natural associated prior probability function pr_0^C in the coarsening frame is given by $pr_0^C(B) := pr(B)$ for all $B \in \mathcal{B}$. Consider the coarsened extended Dempster model $M_0^C = \langle (U, Pr), \Gamma^C, ((\Omega, \mathcal{B}), pr_0^C)\rangle$. It is easy to check that the associated belief function for belief update remains unchanged: for any $B \in \mathcal{B}$, $(Bel_D)_{\mathcal{B}}(B) = Bel_D(B)$ where Bel_D and $(Bel_D)_{\mathcal{B}}$ are the belief functions corresponding to the mass functions m_D and $(m_D)_{\mathcal{B}}$, respectively. But the pignsitic probabilities may change: $Betp_{m_{M_0^C}}(B) \neq Betp_{m_M}(B)$ for some $B \in \mathcal{B}$.

In order to coordinate pignistic probabilities with coarsening, we need to *transform* the prior probability function pr_0^C to a new prior probability pr^C on the coarsening $(\Omega, \mathcal{B})$ such that the pignistic probabilities on the new coarsening frame $M^C := \langle (U, Pr), \Gamma^C, ((\Omega, \mathcal{B}), pr^C)\rangle$ are the same as those on the original extended Dempster model $M(= \langle (U, Pr), \Gamma, (\Omega, pr)\rangle)$: for all $B_i \in \mathcal{C}$, $Betp_{m_M}(B_i) = Betp_{m_{M^C}}(B_i)$. This equality is equivalent to the following one:

$$\frac{\sum_{B\in\mathcal{B}} (m_D)_{\mathcal{B}}(B) pr^C(B \cap B_i)}{\sum_{B\in\mathcal{B}} (m_D)_{\mathcal{B}}(B) pr^C(B)} = Betp_{m_M}(B_i). \quad (12)$$

Let $Pl(B_i)$ denote the sum $\sum_{B_i \subseteq B} (m_D)_{\mathcal{B}}$, where $1 \leq i \leq n$. It is easy to see that $\sum_{B\in\mathcal{B}} (m_D)_{\mathcal{B}}(B) pr^C(B) = \sum_{1\leq i\leq n} pr^C(B_i) Pl(B_i)$ and $\sum_{B\in\mathcal{B}} (m_D)_{\mathcal{B}}(B) pr^C(B \cap B_i) = pr^C(B_i)Pl(B_i)$. So the equality (12) is reduced to the following form: for any $1 \leq i \leq n$,

$$\frac{pr^C(B_i)}{\sum_{1\leq i\leq n} pr^C(B_i)Pl(B_i)} = \frac{Betp_{m_M}(B_i)}{Pl(B_i)}. \quad (13)$$

In this equation, $pr^C(B_i)$ is the only unknown quantity. Since there are n equations with n unknowns in the group $\mathtt{G}$ of Eq.(13), this group has at least one solution. But we don't know whether this solution is nonnegative or not. Now we provide a *constructive* solution to $\mathtt{G}$. Let K denote $\sum_B pr^C(B_i)Pl(B_i)$ and $a_i = \frac{Betp_{m_M}(B_i)}{Pl(B_i)}$. The above group of equations can be simplified as follows: $pr^C(B_i) = a_i K, 1 \leq i \leq n$. Since $\sum_{1\leq i\leq n} pr^C(B_i) = 1$, we get the following equation by adding the equations in $\mathtt{G}$ together: $1 = (a_1 + a_2 + \cdots + a_n)K$. So we get: $K = \frac{1}{\sum_{1\leq i\leq n} \frac{Betp_{m_M}(B_i)}{Pl_i}}$. Finally we solve $\mathtt{G}$ and obtain the following solutions: for any $1 \leq i \leq n$,

$$pr^C(B_i) = \frac{\frac{Betp_{m_M}(B_i)}{Pl(B_i)}}{\sum_{1\leq i\leq n} \frac{Betp_{m_M}(B_i)}{Pl_i(B_i)}}. \quad (14)$$

Theorem 4.6 *The above defined coarsening frame $M^C = \langle (U, Pr), \Gamma^C, ((\Omega, \mathcal{B}), pr^C)\rangle$ with the prior probability pr^C given in Eq.(14) is consistent with pignistic transformation. Let $m_{m_{M^C}}$ be the induced mass function of M^C and m_{D^C} be the induced mass function of the Dempster part $D^C := \langle (U, Pr), \Gamma^C, (\Omega, \mathcal{B})\rangle$. Then we have:*

- $(m_D)_{\mathcal{B}}(B) = m_{D^C}(B)$;
- $Betp_{m_{M^C}}(B) = Betp_{m_M}(B)$ *for all* $B \in \mathcal{B}$.

So pr^C serves as a coordinator between believing represented by $m_{D^C}(= (m_D)_{\mathcal{B}})$ and betting by $Betp_{m_{M^C}}$ on M^C by recording the sensitivity of the pignistic probabilities derived from m_{D^C}. Pignistic transformation provides a credal-pignistic link (Assumption 3.1 in [SMETS, 2005]); pr^C here offers another credal-pignistic link between pignistic probabilities ($Betp_{m_M}(B_i)$) and plausibility $Pl(B_i)$ (defined in terms of m_D) for belief update.

As for Example 1.1, according to the above formulation, we have that $Betp_{m_M^D}(D_1) = \frac{1}{3}$ and $Betp_{m_M^D}(\{D_2, D_3\}) = \frac{2}{3}$. So, since pignistic probabilities are insensitive to the choice of frame, $Betp_{m_M^C}(C_1) = \frac{1}{3}$ and $Betp_{m_M^C}(C_2) = \frac{2}{3}$. Moreover, we get that $Pl(C_1) = 1$ and $Pl(C_2) = 1$. Finally we obtain the prior probability on the frame $\Omega^C : pr^C(C_1) = \frac{1}{3}$ and $pr^C(C_2) = \frac{2}{3}$.

5 RELATED WORKS AND CONCLUSIONS

Yen ([YEN, 1986]) extended the multivalued mapping in the DS theory to a probabilistic one that uses conditional probabilities to express the uncertain associations. He also proposed a combination similar to our rule in Eq.(6) and discussed its relationship to Dempster's rule of combination. Moreover, he distinguished between mass functions for belief update and those for absolute beliefs. Such a distinction motivated our definition of generalized pignistic transformation in extended TBM. But his framework differs from ours in that Yen considered probabilistic multivalued mapping while our probabilistic extension is about prior knowledge base. Our method of combining evidence with prior knowledge is similar to [MAHLER, 1996, FIXSEN AND MAHLER, 1997]. Mahler proposed a similar combination rule and investigated its relationship with Bayesian parallel combination. More importantly, he pointed out the connection between his combination rule and pignistic transformation. He extended DS theory mainly from the perspective of random sets while we stick to the Dempster-model approach. Our work essentially differs from those papers in that we focus on both the partial dissociation of betting from believing and the (in)sensitivity of pignistic probabilities to the choice of frame of discernment. Wilson [WILSON, 1993] did study the sensitivity problem of pignistic probabilities in TBM. But he stayed within the DS theory without considering any probabilistic extension.

In order to translate DS models into probability models which are consistent with belief-function semantics (especially Dempster's rule of combination), Cobb and Shenoy [COBB AND SHENOY, 2006] proposed another probability transformation method called *plausibility transformation* as an alternative to pignistic transformation. Plausibility transformation enjoys many interesting properties. The most important one is the so-called regularity property, i.e., plausibility transformation turns Dempster combination of belief functions into "pointwise" combination of probability functions. But, as Cobb and Shenoy [COBB AND SHENOY, 2006] pointed out, another important operation in DS belief networks, coarsening (or marginalization), is not invariant under this transformation. In fact there is no probability transformation for DS models with Dempster's rule of combination that enjoys the regularity property and makes coarsening invariant [COBB AND SHENOY, 2006]. For a more compressive survey of probability transformation, one may refer to [CUZZOLIN, 2015]. There are many proposals for Jeffrey's rule in DS theory [MA ET AL., 2010, MA ET AL., 2011, SMETS, 1993, ZHOU ET AL., 2014]. But none of these Jeffrey's rules was proposed from the perspective of pignistic transformation as in this paper. Our proposed conditioning rules are consistent with pignistic transformation.

In order to focus on pignistic transformation, we simplify the presentation in this paper by taking a closed world assumption, which is different from Smets' open world assumption for TBM. Moreover, here we choose to represent beliefs with Dempster models, which is opposed to Smets' TBM without probabilistic interpretation. So we would like to investigate the extension of TBM with probabilistic priors under the open-world assumption and its probabilistic interpretation.

Acknowledgements

The first author is partly supported by Key project for basic research from the Ministry of Science and Technology of China (Grant No. 2012CB316205), NSF of China (Grant No. 61370053) and the RUC foundation (Grant No. 2012030005). The second author is supported by ARC Discovery Project (ARC DP130102764), NSF of China (Grant Nos. 61428208 and 61472412), AMSS-UTS Joint Research Laboratory for Quantum Computation, Chinese Academy of Sciences, and the CAS/SAFEA International Partnership Program for Creative Research Team.

References

[Cobb and Shenoy, 2006] Cobb, B. and Shenoy, P. (2006). On the plausibility transformation method for translating belief function models to probability models. *Int. J. Approx. Reasoning*, 41(3):314–330.

[Cuzzolin, 2015] Cuzzolin, F. (2015). *Geometry of Uncertainty*. Springer. to appear.

[Dubois and Denoeux, 2012] Dubois, D. and Denoeux, T. (2012). Conditioning in dempster-shafer theory: Prediction vs. revision. In Denoeux, T. and Masson, M.-H., editors, *Belief Functions*, volume 164 of *Advances in Soft Computing*, pages 385–392. Springer.

[Fagin and Halpern, 1991] Fagin, R. and Halpern, J. (1991). A new approach to updating beliefs. In Bonissone, P., M., H., L., K., and J., L., editors, *UAI*, pages 347–374. Elsevier.

[Fixsen and Mahler, 1997] Fixsen, D. and Mahler, R. (1997). The modified dempster-shafer approach to classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 27(1):96–104.

[Grünwald and Halpern, 2003] Grünwald, P. and Halpern, J. (2003). Updating probabilities. *J. Artif. Intell. Res. (JAIR)*, 19:243–278.

[Halpern, 2005] Halpern, J. (2005). *Reasoning about Uncertainty*. MIT Press.

[Heckerman, 1985] Heckerman, D. (1985). Probabilistic interpretation for mycin's certainty factors. In Kanal, L. N. and Lemmer, J. F., editors, *UAI '85: Proceedings of the First Annual Conference on Uncertainty in Artificial Intelligence, Los Angeles, CA, USA, July 10-12, 1985*, pages 167–196. Elsevier.

[Jaffray, 1992] Jaffray, J. (1992). Bayesian updating and belief functions. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5):1144–1152.

[Ma et al., 2010] Ma, J., Liu, W., Dubois, D., and Prade, H. (2010). Revision rules in the theory of evidence. In *ICTAI (1)*, pages 295–302. IEEE Computer Society.

[Ma et al., 2011] Ma, J., Liu, W., Dubois, D., and Prade, H. (2011). Bridging jeffrey's rule, agm revision and dempster conditioning in the theory of evidence. *International J. on AI Tools*, 20 (4):691–720.

[Mahler, 1996] Mahler, R. (1996). Combining ambiguous evidence with respect to ambiguous a priori knowledge. i. boolean logic. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 26(1):27–41.

[Shafer, 1976] Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, N.J.

[Smets, 1993] Smets, P. (1993). Jeffrey's rule of conditioning generalized to belief functions. In Heckerman, D. and Mamdani, E. H., editors, *UAI*, pages 500–505. Morgan Kaufmann.

[Smets, 2005] Smets, P. (2005). Decision making in the tbm: the necessity of the pignistic transformation. *Int. J. Approx. Reasoning*, 38(2):133–147.

[Smets and Kennes, 1994] Smets, P. and Kennes, R. (1994). The transferable belief model. *Artif. Intell.*, 66(2):191–234.

[Snow, 1998] Snow, P. (1998). The vulnerability of the transferable belief model to dutch books. *Artif. Intell.*, 105(1-2):345–354.

[Wilson, 1993] Wilson, N. (1993). Decision-making with belief functions and pignistic probabilities. In Clarke, M., Kruse, R., and Moral, S., editors, *ECSQARU'93, Granada, Spain, November 8-10, 1993, Proceedings*, volume 747 of *Lecture Notes in Computer Science*, pages 364–371. Springer.

[Yen, 1986] Yen, J. (1986). A reasoning model based on an extended dempster-shafer theory. In Kehler, T., editor, *UAI, 1986. Volume 1: Science.*, pages 125–131. Morgan Kaufmann.

[Zhou et al., 2014] Zhou, C., Wang, M., and Qin, B. (2014). Belief-kinematics jerrefy's rules in the theory of evidence. In Zhang, N. and Tian, J., editors, *UAI 2014*, pages 917–926. AUAI Press.