Proceedings of the 10th
World Congress on Intelligent Control and Automation
July 6-8, 2012, Beijing, China

Hierarchical Proportional Redistribution Principle for Uncertainty Reduction and BBA Approximation*

Jean Dezert
ONERA
The French Aerospace Lab
F-91761 Palaiseau, France
jean.dezert@onera.fr

Deqiang Han
Inst. of Integrated Automation
Xi'an Jiaotong University
Xi'an, 710049, China
deqhan@gmail.com

Zhun-ga Liu
School of Automation,
NW Polytech. Univ.
Xi'an, 710072, China.
liuzhunga@gmail.com

Jean-Marc Tacnet
Cemagref-ETGR,
B.P. 76, 2 rue de la Papéterie
F-38402 St. Martin d'Hères, France
jean-marc.tacnet@cemagref.fr

Abstract—Dempster-Shafer evidence theory is very important in the fields of information fusion and decision making. However, it always brings high computational cost when the frames of discernments to deal with become large. To reduce the heavy computational load involved in many rules of combinations, the approximation of a general belief function is needed. In this paper we present a new general principle for uncertainty reduction based on hierarchical proportional redistribution (HPR) method which allows to approximate any general basic belief assignment (bba) at a given level of non-specificity, up to the ultimate level 1 corresponding to a Bayesian bba. The level of non-specificity can be adjusted by the users. Some experiments are provided to illustrate our proposed HPR method.

Index Terms—Belief functions, hierarchical proportional redistribution (HPR), evidence combination, belief approximation.

I. INTRODUCTION

Dempster-Shafer evidence theory, also called belief function theory [1], is an interesting and flexible tool to deal with imprecision and uncertainty for approximate reasoning. It has been widely used in many applications, e.g., information fusion, pattern recognition and decision making [2].

Although evidence theory is successful in uncertainty modeling and reasoning, high computational cost is a drawback which is often raised against evidence theory [2]. In fact, the computational cost of evidence combination increases exponentially with respect to the size of the frame of discernment (FOD) [3]–[5]. To resolve such a problem, three major types of approaches have been proposed by he researchers.

The first type is to propose efficient procedures for performing exact computations. For example, Kennes [6] proposed an optimal algorithm for Dempster's rule of combination. Barnett's work [7] and other works [8] are also the representatives.

The second type is composed of Monte-Carlo techniques. See details in the paper of Moral and Salmeron [9].

The third type is to approximate (or simplify) a belief function to a simpler one. The papers of Voorbraak [4], Dubois and Prade [10] are seminal works in this type of approaches. Tessem proposed the famous $k-l-x$ approximation approach

[3]. Grabisch proposed some approaches [11], which can build a bridge between belief functions and other types of uncertainty measures or functions, e.g., probabilities, possibilities and k-additive belief function (those belief functions whose cardinality of the focal elements are at most of k). Based on pignistic transformation in transferable belief model (TBM), Burger and Cuzzolin proposed two types of k-additive belief functions [12]. Denœux uses hierarchical clustering strategy to implement the inner and outer approximation of belief functions [13].

In this paper, we focus on the approximation approach of belief functions. This first reason obviously is that it can reduce the computational cost of evidence combination. Furthermore, human find that it is not intuitive to attach meaning to focal elements with large cardinality [14]. Belief approximation can either reduce the number or the cardinalities of focal elements, or both of them can be reduced. Thus by using belief function approximation, we can obtain a representation which is more intuitive and easier to process. We propose a new method called hierarchical proportional redistribution (HPR), which is a general principle for uncertainty reduction, to approximate any general basic belief assignment (bba) at a given level of non-specificity, up to the ultimate level 1 corresponding to a Bayesian bba. That is, our proposed approach can generate an intermediate object between probabilities and original belief function. The level of non-specificity can be controlled by the users through the adjusting of maximum cardinality of remaining focal element. Our proposed approach can be considered as a generalized k-additive belief approximation. Some experiments are provided to illustrate our proposed HPR approach and to compare it with other approximation approaches.

II. BASICS IN EVIDENCE THEORY

In Dempster-Shafer evidence theory [1], the frame of discernment (FOD) denoted by Θ is a basic concept. The elements in Θ are mutually exclusive. Suppose that 2^Θ denotes the powerset of FOD and define the function $m : 2^\Theta \to [0, 1]$ as the basic belief assignment (bba) satisfying:

$$\sum_{A \subseteq \Theta} m(A) = 1, \ m(\emptyset) = 0 \qquad (1)$$

A bba is also called a mass function. Belief function (Bel) and plausibility function (Pl) are defined below, respectively:

$$Bel(A) = \sum_{B \subseteq A} m(B) \qquad (2)$$

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \qquad (3)$$

Suppose that $m_1, m_2, ..., m_n$ are n mass functions, Dempster's rule of combination is defined in (4):

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \dfrac{\sum_{\cap A_i = A} \prod_{1 \leq i \leq n} m_i(A_i)}{\sum_{\cap A_i \neq \emptyset} \prod_{1 \leq i \leq n} m_i(A_i)}, & A \neq \emptyset \end{cases} \qquad (4)$$

Dempster's rule of combination is used in DST to implement the combination of bodies of evidence (BOEs).

Evidence theory has been widely used in many application fields due to its capability of approximate reasoning and processing of uncertain information. However, as referred in introduction section, there also exists the drawback of high computational cost in evidence combination. Several approaches have been proposed accordingly, which includes efficient algorithms [6]–[8] for evidence combination, the Monte-Carlo techniques and the approach of belief function approximation [9]. We prefer to use the belief approximation approach [10]–[13] to reduce the computational cost needed in the combination operation because the approximation approach reduces the computational cost and also allow to deal with smaller-size focal elements, which is more intuitive for human to catch the meaning [14]. In the next section, we recall some well-known basic approximation approaches.

III. BBA APPROXIMATION APPROACHES

1) $k - l - x$ approach: The approach of $k - l - x$ was proposed by Tessem [3]. The simplified or compact bba obtained by using $k - l - x$ satisfies:

1) keep no less than k focal elements;
2) keep no more than l focal elements;
3) the mass assignment to be deleted is no greater than x.

In algorithm of $k - l - x$, the focal elements of a original bba are sorted according to their mass assignments. Such algorithm chooses the first p focal elements such that $k \leq p \leq l$ and such that the sum of the mass assignments of these first p focal elements is no less than $1 - x$. The deleted mass assignments are redistributed to the other focal elements through a normalization.

2) k-additive belief function approximation: Given a bba $m : 2^\Theta \rightarrow [0, 1]$, the k-additive belief function [11], [12] induced by the mass assignment is defined in Eq.(5). Suppose that $B \subseteq \Theta$,

$$\begin{cases} m_k(B) = m(B) + \sum_{A \supset B, A \subseteq \Theta, |A| > k} \frac{m(A) \cdot |B|}{\mathcal{N}(|A|, k)}, & \forall |B| \leq k \\ m_k(B) = 0, & \forall |B| > k \end{cases} \qquad (5)$$

where

$$\mathcal{N}(|A|, k) = \sum_{j=1}^{k} \binom{|A|}{j} \cdot j = \sum_{j=1}^{k} \frac{|A|!}{(j-1)!(|A|-j)!} \qquad (6)$$

is average cardinality of the subsets of A of size at most k.

It can be seen that for k-additive belief approximation, the maximum cardinality of available focal elements is no greater than k.

In this section, $k - l - x$ approach and k-additive belief function approximation approach are introduced, which will be compared with our proposed approach introduced in next section. There also other types of bba approximation approximation approaches, see details in related references.

IV. HIERARCHICAL PROPORTIONAL REDISTRIBUTION APPROXIMATION

In this paper we propose a hierarchical bba approximation approach called hierarchical proportional redistribution (HPR), which provides a new way to reduce step by step the mass committed to uncertainties. Ultimately an approximate measure of subjective probability can be obtained if needed, i.e. a so-called Bayesian bba in [1]. It must be noticed that this procedure can be stopped at any step in the process and thus it allows to reduce the number of focal elements in a given bba in a consistent manner to diminish the size of the core of a bba and thus reduce the complexity (if needed) when applying also some complex rules of combinations. We present here the general principle of hierarchical and proportional reduction of uncertainties in order to obtain approximate bba's at different non-specificity level we expect. The principle of redistribution of uncertainty to more specific elements of the core at any given step of the process follows the proportional redistribution already proposed in the (non hierarchical) DSmP transformation proposed recently in [5]. Thus the proposed HPR can be considered as a bba approximation approach inspired by the idea of DSmP.

Let's first introduce two new notations for convenience and for concision:

1) Any element of cardinality $1 \leq k \leq n$ of the power set 2^Θ will be denoted, by convention, by the generic notation $X(k)$. For example, if $\Theta = \{A, B, C\}$, then $X(2)$ can denote the following partial uncertainties $A \cup B$, $A \cup C$ or $B \cup C$, and $X(3)$ denotes the total uncertainty $A \cup B \cup C$.

2) The proportional redistribution factor (ratio) of width n involving elements X and Y of the powerset is defined as (for $X \neq \emptyset$ and $Y \neq \emptyset$)

$$R_s(Y, X) \triangleq \frac{m(Y) + \epsilon \cdot |X|}{\sum_{\substack{Y \subset X \\ |X| - |Y| = s}} m(Y) + \epsilon \cdot |X|} \qquad (7)$$

where ϵ is a small positive number introduced here to deal with particular cases where $\sum_{\substack{Y \subset X \\ |X| - |Y| = s}} m(Y) = 0$.

By convention, we will denote $R(Y, X) \triangleq R_1(Y, X)$ when we use the proportional redistribution factors of width $s = 1$.

The HPR is obtained by a step by step (recursive) proportional redistribution of the mass $m(X(k))$ of a given uncertainty $X(k)$ (partial or total) of cardinality $2 \leq k \leq n$ to all the least

specific elements of cardinality $k-1$, i.e. to all possible $X(k-1)$, until $k = 2$ is reached. The proportional redistribution is done from the masses of belief committed to $X(k-1)$ as done classically in DSmP transformation. The "hierarchical" masses $m_h(.)$ are recursively (backward) computed as follows. Here $m_{h(k)}$ represents the approximate bba obtained at the step $n-k$ of HPR, i.e., it has the maximum focal element cardinality of k.

$$m_{h(n-1)}(X(n-1)) = m(X(n-1)) +$$
$$\sum_{\substack{X(n) \supset X(n-1) \\ X(n),X(n-1) \in 2^\Theta}} [m(X(n)) \cdot R(X(n-1), X(n))];$$

$$m_{h(n-1)}(A) = m(A), \forall |A| < n - 1$$
(8)

$m_{h(n-1)}(\cdot)$ is the bba obtained at the first step of HPR $(n - (n-1) = 1)$, the maximum focal element cardinality of $m_{h(n-1)}$ is $n-1$.

$$m_{h(n-2)}(X(n-2)) = m(X(n-2)) +$$
$$\sum_{\substack{X(n-1) \supset X(n-2) \\ X(n-2),X(n-1) \in 2^\Theta}} [m_{h(n-1)}(X(n-1)) \cdot R(X(n-2), X(n-1))]$$

$$m_{h(n-2)}(A) = m_{h(n-1)}(A), \forall |A| < n - 2$$
(9)

$m_{h(n-2)}(\cdot)$ is the bba obtained at the second step of HPR $(n - (n-2) = 2)$, the maximum focal element cardinality of $m_{h(n-2)}$ is $n-2$.

This hierarchical proportional redistribution process can be applied similarly (if one wants) to compute $m_{h(n-3)}(.)$, $m_{h(n-4)}(.)$, ..., $m_{h(2)}(\cdot)$, $m_{h(1)}(\cdot)$ with

$$m_{h(2)}(X(2)) = m(X(2)) +$$
$$\sum_{\substack{X(3) \supset X(2) \\ X(3),X(2) \in 2^\Theta}} [m_{h(3)}(X(3)) \cdot R(X(2), X(3))]$$

$$m_{h(2)}(A) = m_{h(3)}(A), \forall |A| < n - 2$$
(10)

$m_{h(2)}(\cdot)$ is the bba obtained at the first step of HPR $(n - 2)$, the maximum focal element cardinality of $m_{h(2)}$ is 2.

Mathematically, for any $X(1) \in \Theta$, i.e. any $\theta_i \in \Theta$ a Bayesian belief function can be obtained by HPR approach in deriving all possible steps of proportional redistributions of partial ignorances in order to get

$$m_{h(1)}(X(1)) = m(X(1)) +$$
$$\sum_{\substack{X(2) \supset X(1) \\ X(1),X(2) \in 2^\Theta}} [m_{h(2)}(X(2)) \cdot R(X(1), X(2))] \quad (11)$$

In fact, $m_{h(1)}(\cdot)$ is a probability transformation, called here the Hierarchical DSmP (HDSmP). Since $X(n)$ is unique and corresponds only to the full ignorance $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n$, the expression of $m_h(X(n-1))$ in Eq.(10) just simplifies as

$$m_{h(n-1)}(X(n-1)) = m_h(X(n-1)) +$$
$$m(X(n)) \cdot R(X(n-1), X(n)) \quad (12)$$

Because of the full proportional redistribution of the masses of uncertainties to the elements least specific involved in these uncertainties, no mass of belief is lost during the step by step hierarchical process and thus at any step of HPR, the sum of masses of belief is kept to one, and of course the Hierarchial DSmP also satisfies $\sum_{X(1) \in 2^\Theta} m_h(X(1)) = 1$.

Remark: For some reasons depending of applications, it is also possible to easily modify this HPR approach with little effort into a constrained HPR version (CHPR for short) by forcing the masses of some partial ignorances of cardinality $k+1$ to be (proportionally) redistributed back only to a subset of the partial ignorances of cardinality k included in them. This possibility has not be detailed here due to space limitation constraint and its little technical interest.

V. EXAMPLES

In this section we show in details how HPR can be applied on very simple different examples. So let's examine the three following examples based on a simple 3D frame of discernment $\Theta = \{\theta_1, \theta_2, \theta_3\}$ satisfying Shafer's model.

A. Example 1

Let's consider the following bba:

$$m(\theta_1) = 0.10, \quad m(\theta_2) = 0.17, \quad m(\theta_3) = 0.03,$$
$$m(\theta_1 \cup \theta_2) = 0.15, \quad m(\theta_1 \cup \theta_3) = 0.20,$$
$$m(\theta_2 \cup \theta_3) = 0.05, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30.$$

We apply the hierarchical proportional redistribution (HPR) principle with $\epsilon = 0$ in this example because there is no mass of belief equal to zero. It can be verified that the result obtained with small positive ϵ parameter remains (as expected) numerically very close to that obtained with $\epsilon = 0$.

The first step of HPR consists in redistributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1 \cup \theta_2 \cup \theta_3$. Applying the Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses.

$$m_{h(2)}(\theta_1 \cup \theta_2) = m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3))$$
$$= 0.15 + (0.30 \cdot 0.375) = 0.2625$$

because $R(\theta_1 \cup \theta_2, X(3)) = \frac{0.15}{0.15+0.20+0.05} = 0.375$.
Similarly, one gets

$$m_{h(2)}(\theta_1 \cup \theta_3) = m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3))$$
$$= 0.20 + (0.30 \cdot 0.5) = 0.35$$

because $R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.15+0.20+0.05} = 0.5$, and also

$$m_{h(2)}(\theta_2 \cup \theta_3) = m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3))$$
$$= 0.05 + (0.30 \cdot 0.125) = 0.0875$$

because $R(\theta_2 \cup \theta_3, X(3)) = \frac{0.05}{0.15+0.20+0.05} = 0.125$.

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$m_{h(1)}(\theta_1) = m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3)$$
$$\approx 0.10 + (0.2625 \cdot 0.3703) + (0.35 \cdot 0.7692)$$
$$= 0.10 + 0.0972 + 0.2692 = 0.4664$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.10}{0.10 + 0.17} \approx 0.3703$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.10}{0.10 + 0.03} \approx 0.7692$$

Similarly, one gets

$$m_{h(1)}(\theta_2) = m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3)$$
$$\approx 0.10 + (0.2625 \cdot 0.6297) + (0.0875 \cdot 0.85)$$
$$= 0.17 + 0.1653 + 0.0744 = 0.4097$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.10 + 0.17} \approx 0.6297$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.03} = 0.85$$

and also

$$m_{h(1)}(\theta_3) = m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3)$$
$$\approx 0.03 + (0.35 \cdot 0.2307) + (0.0875 \cdot 0.15)$$
$$= 0.03 + 0.0808 + 0.0131 = 0.1239$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.03}{0.10 + 0.03} \approx 0.2307$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.03}{0.17 + 0.03} = 0.15$$

Hence, the result of final step of HPR is:

$$m_{h(1)}(\theta_1) = 0.4664, \quad m_{h(1)}(\theta_2) = 0.4097,$$
$$m_{h(1)}(\theta_3) = 0.1239.$$

We can easily verify that

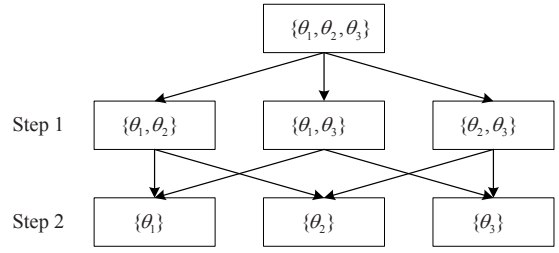$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

Figure 1. Illustration of Example 1

Table I
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 1.

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k = 3$	$k = 2$	$k = 1$
θ_1	0.1000	0.1000	0.4664
θ_2	0.1700	0.1700	0.4097
θ_3	0.0300	0.0300	0.1239
$\theta_1 \cup \theta_2$	0.1500	0.2625	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3500	0.0000
$\theta_2 \cup \theta_3$	0.0500	0.0875	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

The procedure can be illustrated in Fig. 1 below. The approximate bba at each step with different maximum focal elements' cardinality are listed in Table I.

To compare our proposed HPR with the approach of $k - l - x$, we set the parameters of $k - l - x$ to obtain bba's with equal focal element number with HPR at each step. In Example 1, for HPR at first step, it can obtain a bba with 6 focal elements. Thus we set $k = l = 6, x = 0.4$ for $k - l - x$ to obtain a bba with 6 focal elements. Similarly, for HPR at second step, it can obtain a bba with 3 focal elements. Thus we set $k = l = 3, x = 0.4$ for $k - l - x$. Based on the approach of $k - l - x$, the results are in Table II.

Table II
EXPERIMENTAL RESULTS OF $k - l - x$ FOR EXAMPLE 1

Focal elements	$m(\cdot)$ obtained by $k - l - x$	
	$k = l = 6$	$k = l = 3$
θ_1	0.1031	0.0000
θ_2	0.1753	0.2573
θ_3	0.0000	0.0000
$\theta_1 \cup \theta_2$	0.1546	0.0000
$\theta_1 \cup \theta_3$	0.2062	0.2985
$\theta_2 \cup \theta_3$	0.0515	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3093	0.4478

B. Example 2

Let's consider the following bba:

$$m(\theta_1) = 0, \quad m(\theta_2) = 0.17, \quad m(\theta_3) = 0.13,$$
$$m(\theta_1 \cup \theta_2) = 0.20, \quad m(\theta_1 \cup \theta_3) = 0.20,$$
$$m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$$

The first step of HPR consists in redistributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, and $\theta_1 \cup \theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1 \cup \theta_2 \cup \theta_3$.

Applying Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$m_{h(2)}(\theta_1 \cup \theta_2) = m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3))$$
$$= 0.20 + (0.3 \cdot 0.5) = 0.35$$

because

$$R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20}{0.20 + 0.20 + 0.00} = 0.5$$

Similarly, one gets

$$m_{h(2)}(\theta_1 \cup \theta_3) = m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3))$$
$$= 0.20 + (0.3 \cdot 0.5) = 0.35$$

because

$$R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.20 + 0.20 + 0.00} = 0.5$$

and also

$$m_{h(2)}(\theta_2 \cup \theta_3) = m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3))$$
$$= 0.00 + (0.3 \cdot 0.0) = 0.0$$

because

$$R(\theta_2 \cup \theta_3, X(3)) = \frac{0.0}{0.20 + 0.20 + 0.00} = 0$$

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$m_{h(1)}(\theta_1) = m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3)$$
$$\approx 0.00 + (0.35 \cdot 0.00) + (0.35 \cdot 0.00)$$
$$= 0.00 + 0.00 + 0.00 = 0.00$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.00}{0.00 + 0.17} = 0.00$$
$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.00}{0.00 + 0.13} = 0.00$$

Similarly, one gets

$$m_{h(1)}(\theta_2) = m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3)$$
$$\approx 0.17 + (0.35 \cdot 1) + (0.00 \cdot 0.5667)$$
$$= 0.17 + 0.35 + 0.00 = 0.52$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.00 + 0.17} = 1$$
$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.13} \approx 0.5667$$

and also

$$m_{h(1)}(\theta_3) = m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3)$$
$$\approx 0.13 + (0.35 \cdot 1) + (0.00 \cdot 0.4333)$$
$$= 0.13 + 0.35 + 0.00 = 0.48$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.13}{0.13 + 0.00} = 1$$
$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.13}{0.17 + 0.13} \approx 0.4333$$

Hence, the result of final step of HPR is

$$m_{h(1)}(\theta_1) = 0.00, \quad m_{h(1)}(\theta_2) = 0.52, \quad m_{h(1)}(\theta_3) = 0.48$$

and we can easily verify that

$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

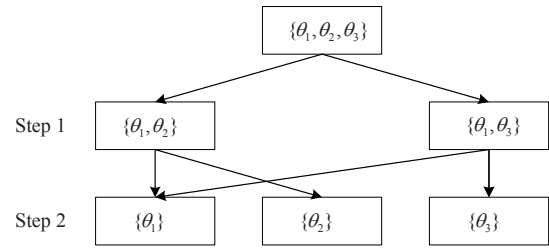The HPR procedure of Example 2 with $\epsilon = 0$ is Fig. 2.

Figure 2. Illustration of Example 2.

If one takes $\epsilon = 0$, there is no mass that will be reassigned to $\{\theta_2 \cup \theta_3\}$ as illustrated in Fig. 2. But if one takes $\epsilon > 0$, HPR procedure of Example 2 is the same as that illustrated in Fig. 1, i.e., there also exist masses redistributed to $\{\theta_2 \cup \theta_3\}$ as illustrated in Fig. 1. That's the difference between Fig. 1 and Fig. 2.

Suppose that $\epsilon = 0.001$, the HPR calculation procedure is as follows.

The first step of HPR consists in distributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$. Applying the Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$m_{h(2)}(\theta_1 \cup \theta_2) = m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3))$$
$$= 0.20 + (0.3 \cdot 0.4963) \approx 0.3489$$

because

$$R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)}$$
$$= 0.4963$$

$$m_{h(2)}(\theta_1 \cup \theta_3) = m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3))$$
$$= 0.20 + (0.3 \cdot 0.4963) = 0.3489$$

668

because

$$R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)}$$
$$= 0.4963$$

$$m_{h(2)}(\theta_2 \cup \theta_3) = m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3))$$
$$= 0.00 + (0.3 \cdot 0.0073) = 0.0022$$

because

$$R(\theta_2 \cup \theta_3, X(3)) = \frac{0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)}$$
$$= 0.0073$$

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$m_{h(1)}(\theta_1) = m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3)$$
$$\approx 0.00 + (0.3489 \cdot 0.0115) + (0.3489 \cdot 0.0149)$$
$$= 0.00 + 0.0040 + 0.0052 = 0.0092$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)}$$
$$= 0.0115$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)}$$
$$= 0.0149$$

Similarly, one gets

$$m_{h(1)}(\theta_2) = m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3)$$
$$\approx 0.17 + (0.3489 \cdot 0.9885) + (0.0022 \cdot 0.5658)$$
$$= 0.17 + 0.3449 + 0.0012 = 0.5161$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)}$$
$$= 0.9885$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)}$$
$$\approx 0.5658$$

and also

$$m_{h(1)}(\theta_3) = m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3)$$
$$+ m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3)$$
$$\approx 0.13 + (0.3489 \cdot 0.9851) + (0.0022 \cdot 0.4342)$$
$$= 0.13 + 0.3437 + 0.0009 = 0.4746$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.13 + 0.001 \cdot 2}{(0.13 + 0.001 \cdot 2) + (0.00 + 0.001 \cdot 2)}$$
$$= 0.9851$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.13 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)}$$
$$\approx 0.4342$$

Hence, the final result of HPR approximation is

$$m_{h(1)}(\theta_1) = 0.0092, \quad m_{h(1)}(\theta_2) = 0.5161,$$
$$m_{h(1)}(\theta_3) = 0.4746$$

and we can easily verify that

$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

The bba's obtained in each step are listed in Table III ($\epsilon = 0$) and Table IV ($\epsilon = 0.001$)

Table III
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 2 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k = 3$	$k = 2$	$k = 1$
θ_1	0.0000	0.0000	0.0000
θ_2	0.1700	0.1700	0.5200
θ_3	0.1300	0.1300	0.4800
$\theta_1 \cup \theta_2$	0.2000	0.3500	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3500	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

Table IV
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 2 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k = 3$	$k = 2$	$k = 1$
θ_1	0.0000	0.0000	0.0092
θ_2	0.1700	0.1700	0.5141
θ_3	0.1300	0.1300	0.4746
$\theta_1 \cup \theta_2$	0.2000	0.3489	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3489	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0022	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

When using $k - l - x$ approach, the results are in Table V.

Table V
EXPERIMENTAL RESULTS OF $k - l - x$ FOR EXAMPLE 2

Focal elements	$m(\cdot)$ obtained by $k - l - x$	
	$k = l = 6$	$k = l = 3$
θ_1	0.0000	0.0000
θ_2	0.1700	0.0000
θ_3	0.1300	0.0000
$\theta_1 \cup \theta_2$	0.2000	0.2857
$\theta_1 \cup \theta_3$	0.2000	0.2857
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.4286

C. Example 3

Let's consider the following bba:

$$m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0.70,$$
$$m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0,$$
$$m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$$

In this example, the mass assignments for all the focal elements with cardinality size 2 equal to zero. For HPR, when $\epsilon > 0$, $m(\theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$. Because the ratios are

$$R(\theta_1 \cup \theta_2, X(3)) = R(\theta_1 \cup \theta_3, X(3)) = R(\theta_2 \cup \theta_3, X(3))$$
$$= \frac{0.00 + 0.001 \cdot 3}{(0.00 + 0.001 \cdot 3) \cdot 3} = 0.3333$$

For HPR, when $\epsilon = 0$, it can not be executed directly. This can show the necessity for the using of ϵ.

The bba's obtained through $\text{HPR}_{\epsilon=0.001}$ at different steps are listed in Table VI

Table VI
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 3 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k = 3$	$k = 2$	$k = 1$
θ_1	0.0000	0.0000	0.0503
θ_2	0.0000	0.0000	0.0503
θ_3	0.7000	0.7000	0.8994
$\theta_1 \cup \theta_2$	0.0000	0.1000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.1000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.1000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

When using $k - l - x$ approach, the results are in Table VII.

Table VII
EXPERIMENTAL RESULTS OF $k - l - x$ FOR EXAMPLE 3

Focal elements	$m(\cdot)$ obtained by $k - l - x$	
	$k = l = 6$	$k = l = 3$
θ_1	0.0000	0.0000
θ_2	0.0000	0.0000
θ_3	0.7000	0.7000
$\theta_1 \cup \theta_2$	0.0000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.3000

D. Example 4 (vacuous bba)

Let's consider the following bba:

$$m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0,$$
$$m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0,$$
$$m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 1$$

In this example, the mass assignments for all the focal elements with cardinality size less than 3 equal to zero. For HPR, when $\epsilon > 0$, $m(\theta_1 \cup \theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$.

Similarly, the mass assignments for focal elements with cardinality of 2 obtained in intermediate step will be divided equally and redistributed to singletons. This is due to $\epsilon > 0$.

For HPR, when $\epsilon = 0$, it can not be executed directly. This can show the necessity for the using of ϵ. The bba's obtained through $\text{HPR}_{\epsilon=0.001}$ at different steps are listed in Table VIII.

Table VIII
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 4 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k = 3$	$k = 2$	$k = 1$
θ_1	0.0000	0.0000	0.3333
θ_2	0.0000	0.0000	0.3333
θ_3	0.0000	0.0000	0.3333
$\theta_1 \cup \theta_2$	0.0000	0.3333	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.3333	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.3333	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	1.0000	0.0000	0.0000

When using $k - l - x$ approach, the results are in Table IX.

Table IX
EXPERIMENTAL RESULTS OF $k - l - x$ FOR EXAMPLE 3

Focal elements	$m(\cdot)$ obtained by $k - l - x$	
	$k = l = 6$	$k = l = 3$
θ_1	0.0000	0.0000
θ_2	0.0000	0.0000
θ_3	0.0000	0.0000
$\theta_1 \cup \theta_2$	0.0000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	1.0000	1.0000

From the results of Example 1 – Example 4, we can see that based on $k - l - x$, the users can control the number of focal elements but can not control the maximum cardinality of focal elements. Although based on $k - l - x$, the number of focal elements can be reduced, the focal elements with big cardinality might also be remained. This is not good for further reducing computational cost and not good for human to catch the meaning.

E. Example 5

More generally, an approximation method 1 (giving $m_1(.)$) is considered better than a method 2 (giving $m_2(.)$) if both conditions are fulfilled: 1) if Jousselme's distance of $m_1(.)$ to original bba $m(.)$ is smaller than the distance of $m_2(.)$ to original bba $m(.)$, i.e. $d(m_1, m) < d(m_2, m)$; 2) if the approximate non-specificity value $U(m_1)$ is closer (and lower) to the true non-specificity value $U(m)$ than $U(m_2)$, where Jousselme's distance is defined in [16], and non-specificity [17] is given by $U(m) = \sum_{A \subseteq \Theta} m(A) \log_2 |A|$.

In this example, we make a comparison between HPR (method 1) and k-additive approach (method 2). We consider the FoD $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and we generate randomly $L = 30$ bba's by using the algorithm given below [15]:

Input: Θ : Frame of discernment;

N_{max}: Maximum number of focal elements

Output: Bel: Belief function (under the form of a bba, m)
Generate the power set of Θ: $\mathcal{P}(\Theta)$;

Generate a random permutation of $\mathcal{P}(\Theta) \to \mathcal{R}(\Theta)$;
Generate a integer between 1 and $N_{max} \to k$;
FOReach First k elements of $\mathcal{R}(\Theta)$ do
Generate a value within $[0,1] \to m'_k$;
END Normalize the vector $m'(.) = [m'_1, ..., m'_k] \to m(.)$
(that is $m(A_k) = m_k$);
Algorithm 1: Random generation of bba.

We compute and plot $d(m_1^j, m)$, $d(m_2^j, m)$, $U(m)$, $U(m_1^j)$ and $U(m_2^j)$ for several levels of approximation for $j = 1, 2, \ldots, L$ (where j is the index of the Monte-Carlo run). The results are shown in Fig. 3 and indicate clearly the superiority of HPR over the k-additive approach.
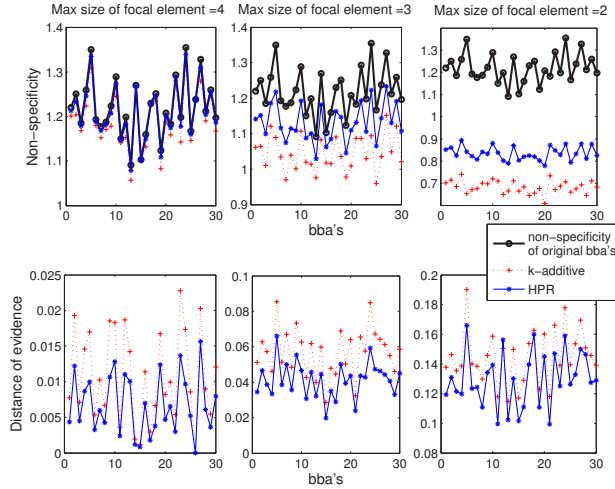

Figure 3.   Illustration of Example 5.

We further use the Normalized Mean Square Error (NMSE) statistics defined by

$$NMSE_i = \frac{1}{L} \sum_{j=1}^{L} \frac{(U(m_i^j) - U(m))^2}{\text{Var}(\vec{e}_i)} \qquad (13)$$

to evaluate the global quality of the approximation of the non-specificity by HPR (if $i = 1$) and by k-additive method (if $i = 2$). $\vec{e}_i = [e_i^1, \ldots, e_i^j, \ldots, e_i^L]$ is the approximation error vector of method #i where $e_i^j = U(m_i^j) - U(m)$, for $j = 1, ..., L$. $\text{Var}(\vec{e}_i)$ is the variance of $\vec{e}_i$. The NMSE results are given in Table X below.

Table X
NMSE RESULTS OF EXAMPLE 5

Max size of focal element	4	3	2
k-additive method	3.9003	21.8118	69.0191
HPR method	**3.9003**	**19.0264**	**61.9468**

Table X shows that HPR outperforms k-additive method since it provides a lower NMSE, which means that in terms of information loss, HPR is better (it generates less loss) than the k-additive approximation method.

VI. CONCLUSIONS

We have proposed a new interesting and useful hierarchical method, called HPR, to approximate any bba. The non-specificity degree can be easily controlled by the user. Some examples were provided to show how HPR works, and to show its rationality and advantage in comparison with some well-known bba approximation approaches. In future works, we will compare this HPR method with more bba approximation methods. In this paper, we have used only the distance of evidence and non-specificity as performance criteria. We plan to develop a more efficient evaluation criteria for capturing more aspects of the information expressed in a bba to measure the global performances of a method, and to design a better bba approximation approach (if possible).

REFERENCES

[1] G. Shafer, A Mathematical Theory of Evidence, Princeton, NJ: Princeton University, 1976.
[2] P. Smets, "Practical uses of belief functions", in K. B. Lskey and H. Prade, Editors, Uncertianty in Artificial Intelligence 15 (UAI 99), Stockholm, Sweden, pp. 612-621, 1999.
[3] B. Tessem, "Approximations for efficient computation in the theory of evidence", Artificial Intelligence, Vol. 61, no. 2, pp. 315-329, June 1993.
[4] F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory", Int. J. Man-Machine Studies, Vol. 30, pp. 525-536, 1989.
[5] F. Smarandache, J. Dezert (Editors), Applications and Advances of DSmT for Information Fusion (Vol 3), Rehoboth, NM: American Research Press, 2009. http://www.gallup.unm.edu/~smarandache/DSmT-book3.pdf.
[6] R. Kennes, "Computational aspects of the Möbius transform of graphs", IEEE Transactions on SMC, Vol. 22, pp. 201-223, 1992.
[7] J.A. Barnett, "Computational methods for a mathematical theory of evidence", in Proceedings of IJCAI-81, Vancouver, pp. 868-875, 1981.
[8] G. Shafer, R. Logan, "Implementing Dempster's rule for hierarchical evidence", Artificial Intelligence, Vol. 33, pp. 271-298, 1987.
[9] S. Moral, A. Salmeron, "A Monte Carlo algorithm for combining Dempster-Shafer belief based on approximate pre-computation", in A. Hunter and S. Pearsons, Editors, Symbolic and quantitative approaches to reasoning and uncertainty (ECSQARU'99), London, UK, pp. 305-315, 1999.
[10] D. Dubois, H. Prade, "An alternative approach to the handling of subnormal possiblity distributions", Fuzzy Sets and Systems, Vol. 24, pp. 123-126, 1987.
[11] M. Grabisch, "Upper approximation of non-additive measures by k-additive measures - the case of belief functions", in Proc. of 1st Int. Symp. on Imprecise Proba. and their applications, Ghent, Belgium, June 1999.
[12] T. Burger, F. Cuzzolin, "Two k-additive generalizations of the pignistic transform", submitted in 2011 to Fuzzy Sets and Systems, available on line: http://cms.brookes.ac.uk/staff/FabioCuzzolin/files/fss11kadditive.pdf
[13] T. Denœux, "Inner and outer approximation of belief structures using a hierarchical clustering approach", Int. J. of Uncertainty, Fuzziness, and Knowledge-based Systems, Vol. 9, no. 4, pp. 437-460, 2001.
[14] T. Burger, "Defining new approximations of belief functions by means of Dempster's combination", in Proc. of the 1st International Workshop on the Theories of Belief Functions (WTBF 2010), Brest, France, March 31st - April 2nd, 2010.
[15] A.-L. Jousselme, P. Maupin, "On some properties of distances in evidence theory", in Proc. of the 1st Workshop on Theory of Belief Functions(WTBF2010), Brest, France, March 31st - April 2nd, 2010.
[16] A.-L. Jousselme, D. Grenier, E. Bosse, "A new distance between two bodies of evidence", Information Fusion, Vol. 2, no. 2, pp. 91-101, 2001.
[17] D. Dubois, H. Prade, "A note on measures of specificity for fuzzy sets", International Journal of General Systems, Vol. 10, no. 4, pp. 279-283, 1985.