Why Dempster's Fusion Rule is not a Generalization of Bayes Fusion Rule

Jean Dezert
Albena Tchamova
Deqiang Han
Jean-Marc Tacnet

Abstract—In this paper, we analyze Bayes fusion rule in details from a fusion standpoint, as well as the emblematic Dempster's rule of combination introduced by Shafer in his Mathematical Theory of evidence based on belief functions. We propose a new interesting formulation of Bayes rule and point out some of its properties. A deep analysis of the compatibility of Dempster's fusion rule with Bayes fusion rule is done. We show that Dempster's rule is compatible with Bayes fusion rule only in the very particular case where the basic belief assignments (bba's) to combine are Bayesian, and when the prior information is modeled either by a uniform probability measure, or by a vacuous bba. We show clearly that Dempster's rule becomes incompatible with Bayes rule in the more general case where the prior is truly informative (not uniform, nor vacuous). Consequently, this paper proves that Dempster's rule is not a generalization of Bayes fusion rule.

Keywords—*Information fusion, Probability theory, Bayes fusion rule, Dempster's fusion rule.*

I. INTRODUCTION

In 1979, Lotfi Zadeh questioned in [1] the validity of the Dempster's rule of combination [2], [3] proposed by Shafer in Dempster-Shafer Theory (DST) of evidence [4]. Since more than 30 years many strong debates [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15] on the validity of foundations of DST and Dempster's rule have bloomed. The purpose of this paper is not to discuss the validity of Dempster's rule, nor the foundations of DST which have been already addressed in previous papers [16], [17], [18]. In this paper, we just focus on the deep analysis of the real incompatibility of Dempster's rule with Bayes fusion rule. Our analysis supports Mahler's one briefly presented in [19].

This paper is organized as follows. In section II, we recall basics of conditional probabilities and Bayes fusion rule with its main properties. In section III, we recall the basics of belief functions and Dempster's rule. In section IV, we analyze in details the incompatibility of Dempster's rule with Bayes rule in general and its partial compatibility for the very particular case when prior information is modeled by a Bayesian uniform basic belief assignment (bba). Section V concludes this paper.

II. CONDITIONAL PROBABILITIES AND BAYES FUSION

In this section, we recall the definition of conditional probability [20], [21] and present the principle and the properties of Bayes fusion rule. We present the structure of this rule derived from the classical definition of the conditional probability in a new uncommon interesting form that will help us to analyze its partial similarity with Dempster's rule proposed by Shafer in his mathematical theory of evidence [4]. We will show clearly why Dempster's rule fails to be compatible with Bayes rule in general.

A. Conditional probabilities

Let us consider two random events X and Z. The conditional probability mass functions (pmfs) $P(X|Z)$ and $P(Z|X)$ are defined[1] (assuming $P(X) > 0$ and $P(Z) > 0$) by [20]:

$$P(X|Z) \triangleq \frac{P(X \cap Z)}{P(Z)} \quad \text{and} \quad P(Z|X) \triangleq \frac{P(X \cap Z)}{P(X)} \quad (1)$$

From Eq. (1), one gets $P(X \cap Z) = P(X|Z)P(Z) = P(Z|X)P(X)$, which yields to Bayes Theorem:

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)} \text{ and } P(Z|X) = \frac{P(X|Z)P(Z)}{P(X)} \quad (2)$$

where $P(X)$ is called the *a priori probability* of X, and $P(Z|X)$ is called the *likelihood of X*. The denominator $P(Z)$ plays the role of a normalization constant warranting that $\sum_{i=1}^{N} P(X = x_i|Z) = 1$. In fact $P(Z)$ can be rewritten as

$$P(Z) = \sum_{i=1}^{N} P(Z|X = x_i)P(X = x_i) \quad (3)$$

The set of the N possible exclusive and exhaustive outcomes of X is denoted $\Theta(X) \triangleq \{x_i, i = 1, 2, \ldots, N\}$.

B. Bayes parallel fusion rule

In fusion applications, we are often interested in computing the probability of an event X given two events Z_1 and Z_2 that have occurred. More precisely, one wants to compute $P(X|Z_1 \cap Z_2)$ knowing $P(X|Z_1)$ and $P(X|Z_2)$, where X can take N distinct exhaustive and exclusive states x_i, $i = 1, 2, \ldots, N$. Such type of problem is traditionally called a *fusion problem*. The computation of $P(X|Z_1 \cap Z_2)$ from

---

[1]For convenience and simplicity, we use the notation $P(X|Z)$ instead of $P(X = x|Z = z)$, and $P(Z|X)$ instead of $P(Z = z|X = x)$ where x and z would represent precisely particular outcomes of the random variables X and Z.

$P(X|Z_1)$ and $P(X|Z_2)$ cannot be done in general without the knowledge of the probabilities $P(X)$ and $P(X|Z_1 \cup Z_2)$ which are rarely given. However, $P(X|Z_1 \cap Z_2)$ becomes easily computable by assuming the following conditional statistical independence condition expressed mathematically by:

$$(A1): \qquad P(Z_1 \cap Z_2|X) = P(Z_1|X)P(Z_2|X) \qquad (4)$$

With such conditional independence condition (A1), then from Eq. (1) and Bayes Theorem one gets:

$$P(X|Z_1 \cap Z_2) = \frac{P(Z_1 \cap Z_2 \cap X)}{P(Z_1 \cap Z_2)} = \frac{P(Z_1 \cap Z_2|X)P(X)}{P(Z_1 \cap Z_2)}$$
$$= \frac{P(Z_1|X)P(Z_2|X)P(X)}{\sum_{i=1}^{N} P(Z_1|X=x_i)P(Z_2|X=x_i)P(X=x_i)}$$

Using again Eq. (2), we have:

$$P(Z_1|X) = \frac{P(X|Z_1)P(Z_1)}{P(X)} \text{ and } P(Z_2|X) = \frac{P(X|Z_2)P(Z_2)}{P(X)}$$

and the previous formula of conditional probability $P(X|Z_1 \cap Z_2)$ can be rewritten as:

$$P(X|Z_1 \cap Z_2) = \frac{\frac{P(X|Z_1)P(X|Z_2)}{P(X)}}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}} \qquad (5)$$

The rule of combination given by Eq. (5) is known as *Bayes parallel (or product) rule* and dates back to Bernoulli [22]. In the classification framework, this formula is also called the *Naive Bayesian Classifier* because it uses the assumption (A1) which is often considered as very unrealistic and too simplistic, and that is why it is called a *naive* assumption. The Eq. (5) can be rewritten as:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K(X, Z_1, Z_2)} \cdot P(X|Z_1) \cdot P(X|Z_2) \quad (6)$$

where the coefficient $K(X, Z_1, Z_2)$ is defined by:

$$K(X, Z_1, Z_2) \triangleq P(X) \cdot \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}$$
$$(7)$$

C. Symmetrization of Bayes fusion rule

The expression of Bayes fusion rule given by Eq. (5) can also be symmetrized in the following form that, quite surprisingly, rarely appears in the literature:

$$P(X|Z_1 \cap Z_2) = \frac{\frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}}}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}} \qquad (8)$$

or in an equivalent manner:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K'(Z_1, Z_2)} \cdot \frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}} \quad (9)$$

where the normalization constant $K'(Z_1, Z_2)$ is given by:

$$K'(Z_1, Z_2) \triangleq \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} \qquad (10)$$

We call the quantity $A_2(X = x_i) \triangleq \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}$ entering in Eq. (10) the *Agreement Factor* on $X = x_i$ of order 2, because only two posterior pmfs are used in the derivation. $A_2(X = x_i)$ corresponds to the posterior conjunctive consensus on the event $X = x_i$ taking into account the prior pmf of X. The denominator of Eq. (8) measures the level of the Global Agreement (GA) of the conjunctive consensus taking into account the prior pmf of X. It is denoted[2] GA_2.

$$GA_2 \triangleq \sum_{i_1, i_2=1|i_1=i_2}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X=x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}}$$
$$= \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} = K'(Z_1, Z_2)$$
$$(11)$$

In fact, with assumption (A1), the probability $P(X|Z_1 \cap Z_2)$ given in Eq. (9) is nothing but the simple ratio of the agreement factor $A_2(X)$ (conjunctive consensus) on X over the global agreement $GA_2 = \sum_{i=1}^{N} A_2(X = x_i)$, that is:

$$P(X|Z_1 \cap Z_2) = \frac{A_2(X)}{GA_2} \qquad (12)$$

The quantity GC_2 given in Eq. (13) measures the global conflict (i.e. the total conjunctive disagreement) taking into account the prior pmf of X.

$$GC_2 \triangleq \sum_{i_1, i_2=1|i_1 \neq i_2}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X=x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}} \quad (13)$$

• Generalization to $P(X|Z_1 \cap Z_2 \cap \ldots \cap Z_s)$

It can be proved that, when assuming conditional independence conditions, Bayes parallel combination rule can be generalized for combining $s > 2$ posterior pmfs as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{1}{K(X, Z_1, \ldots, Z_s)} \cdot \prod_{k=1}^{s} P(X|Z_k) \qquad (14)$$

where the coefficient $K(X, Z_1, \ldots, Z_s)$ is defined by:

$$K(X, Z_1, \ldots, Z_s) \triangleq P(X) \sum_{i=1}^{N} \frac{(\prod_{k=1}^{s} P(X=x_i|Z_k))}{P(X=x_i)} \qquad (15)$$

The symmetrized form of Eq. (14) is:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{1}{K'(Z_1, \ldots, Z_s)} \cdot \prod_{k=1}^{s} \frac{P(X|Z_k)}{\sqrt[s]{P(X)}} \qquad (16)$$

with the normalization constant $K'(Z_1, \ldots, Z_s)$ given by:

$$K'(Z_1, \ldots, Z_s) \triangleq \sum_{i=1}^{N} \prod_{k=1}^{s} \frac{P(X=x_i|Z_k)}{\sqrt[s]{P(X=x_i)}} \qquad (17)$$

---

[2]The index 2 is introduced explicitly in the notations because we consider only the fusion of two posterior pmfs.

The generalization of $A_2(X)$, GA_2, and GC_2 provides the agreement $A_s(X)$ of order s, the global agreement GA_s and the global conflict GC_s for s sources as follows:

$$A_s(X = x_i) \triangleq \prod_{k=1}^{s} \frac{P(X = x_i|Z_k)}{\sqrt[s]{P(X = x_i)}}$$

$$GA_s \triangleq \sum_{i_1,\ldots,i_s=1|i_1=\ldots=i_s}^{N} \frac{P(X = x_{i_1}|Z_1)}{\sqrt[s]{P(X = x_{i_1})}} \cdots \frac{P(X = x_{i_s}|Z_s)}{\sqrt[s]{P(X = x_{i_s})}}$$

$$GC_s \triangleq \sum_{i_1,\ldots,i_s=1}^{N} \frac{P(X = x_{i_1}|Z_1)}{\sqrt[s]{P(X = x_{i_1})}} \cdots \frac{P(X = x_{i_s}|Z_s)}{\sqrt[s]{P(X = x_{i_s})}} - GA_s$$

• **Symbolic representation of Bayes fusion rule**

The (symmetrized form of) Bayes fusion rule of two posterior probability measures $P(X|Z_1)$ and $P(X|Z_2)$, given in Eq. (9), requires an extra knowledge of the prior probability of X. For convenience, we denote symbolically this fusion rule as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2); P(X)) \quad (18)$$

Similarly, the (symmetrized) Bayes fusion rule of $s \geq 2$ probability measures $P(X|Z_k)$, $k = 1, 2, \ldots, s$ given by Eq. (16), which requires also the knowledge of $P(X)$, will be denoted as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X))$$

• **Particular case: Uniform a priori pmf**

If the random variable X is assumed as a priori uniformly distributed over the space of its N possible outcomes, then the probability of X is equal to $P(X = x_i) = 1/N$ for $i = 1, 2, \ldots, N$. In such particular case, all the prior probabilities values $\sqrt{P(X = x_i)} = \sqrt{1/N}$ and $\sqrt[s]{P(X = x_i)} = \sqrt[s]{1/N}$ can be simplified in Bayes fusion formulas Eq. (9) and Eq. (10). Therefore, Bayes fusion formula (9) reduces to:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{\sum_{i=1}^{N} P(X = x_i|Z_1)P(X = x_i|Z_2)} \quad (19)$$

By convention, Eq. (19) is denoted symbolically as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2)) \quad (20)$$

Similarly, $Bayes(P(X|Z_1), \ldots, P(X|Z_s))$ rule defined with an uniform a priori pmf of X will be given by:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s} P(X = x_i|Z_k)} \quad (21)$$

When $P(X)$ is uniform and from Eq. (19), one can redefine the global agreement and the global conflict as:

$$GA_2^{unif} \triangleq \sum_{i,j=1|i=j}^{N} P(X = x_i|Z_1)P(X = x_j|Z_2) \quad (22)$$

$$GC_2^{unif} \triangleq \sum_{i,j=1|i\neq j}^{N} P(X = x_i|Z_1)P(X = x_j|Z_2) \quad (23)$$

Because $\sum_{i=1}^{N} P(X = x_i|Z_1) = 1$ and $\sum_{j=1}^{N} P(X = x_j|Z_2) = 1$, then

$$1 = \left(\sum_{i=1}^{N} P(X = x_i|Z_1)\right)\left(\sum_{j=1}^{N} P(X = x_j|Z_2)\right)$$

$$= \sum_{i,j=1}^{N} P(X = x_i|Z_1)P(X = x_j|Z_2)$$

$$= \sum_{i,j=1|i=j}^{N} P(X = x_i|Z_1)P(X = x_j|Z_2)$$

$$+ \sum_{i,j=1|i\neq j}^{N} P(X = x_i|Z_1)P(X = x_j|Z_2)$$

Therefore, one has always $GA_2^{unif} + GC_2^{unif} = 1$ when $P(X)$ is uniform, and Eq. (19) can be expressed as:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{GA_2^{unif}} = \frac{P(X|Z_1)P(X|Z_2)}{1 - GC_2^{unif}} \quad (24)$$

By a direct extension, one will have:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{GA_s^{unif}} = \frac{\prod_{k=1}^{s} P(X|Z_k)}{1 - GC_s^{unif}} \quad (25)$$

$$GA_s^{unif} = \sum_{i_1,\ldots,i_s=1|i_1=\ldots=i_s}^{N} P(X = x_{i_1}|Z_1)\ldots P(X = x_{i_s}|Z_s)$$

$$GC_s^{unif} = 1 - GA_s^{unif}$$

Remark 1: The normalization coefficient corresponding to the global conjunctive agreement GA_s^{unif} can also be expressed using belief function notations [4] as:

$$GA_s^{unif} = \sum_{\substack{x_{i_1},\ldots,x_{i_s} \in \Theta(X) \\ x_{i_1} \cap \ldots \cap x_{i_s} \neq \emptyset}} P(X = x_{i_1}|Z_1)\ldots P(X = x_{i_s}|Z_s)$$

and the global disagreement, or total conflict level, is given by:

$$GC_s^{unif} = \sum_{\substack{x_{i_1},\ldots,x_{i_s} \in \Theta(X) \\ x_{i_1} \cap \ldots \cap x_{i_s} = \emptyset}} P(X = x_{i_1}|Z_1)\ldots P(X = x_{i_s}|Z_s)$$

D. Properties of Bayes fusion rule

In this subsection, we analyze Bayes fusion rule (assuming condition (A1) holds) from a pure algebraic standpoint. In fusion jargon, the quantities to combine come from sources of information which provide inputs that feed the fusion rule. In the probabilistic framework, a source s to combine corresponds to the posterior pmf $P(X|Z_s)$. In this subsection, we establish five interesting properties of Bayes rule. Contrary to Dempster's rule, we prove that Bayes rule is not associative in general.

• **(P1) : The pmf $P(X)$ is a neutral element of Bayes fusion rule when combining only two sources.**

Proof: A source is called a neutral element of a fusion rule if and only if it has no influence on the fusion result. $P(X)$ is a neutral element of Bayes rule if and only if

$Bayes(P(X|Z_1), P(X); P(X)) = P(X|Z_1)$. It can be easily verified that this equality holds by replacing $P(X|Z_2)$ by $P(X)$ and $P(X = x_i|Z_2)$ by $P(X = x_i)$ (as if the conditioning term Z_2 vanishes) in Eq. (5). One can also verify that $Bayes(P(X), P(X|Z_2); P(X)) = P(X|Z_2)$, which completes the proof.

Remark 2: When considering Bayes fusion of more than two sources, $P(X)$ doesn't play the role of a neutral element in general, except if $P(X)$ is uniform. For example, let us consider 3 pmfs $P(X|Z_1)$, $P(X|Z_2)$ and $P(X|Z_3)$ to combine with formula (14) with $P(X)$ not uniform. When Z_3 vanishes so that $P(X|Z_3) = P(X)$, we can easily check that:

$$Bayes(P(X|Z_1), P(X|Z_2), P(X); P(X))$$
$$\neq Bayes(P(X|Z_1), P(X|Z_2); P(X)) \quad (26)$$

• **(P2) : Bayes fusion rule is in general not idempotent.**

Proof: A fusion rule is idempotent if the combination of all same inputs is equal to the inputs. To prove that Bayes rule is not idempotent it suffices to prove that in general:

$$Bayes(P(X|Z_1), P(X|Z_1); P(X)) \neq P(X|Z_1)$$

From Bayes rule (5), when $P(X|Z_2) = P(X|Z_1)$ we clearly get in general

$$\frac{1}{P(X)} \frac{P(X|Z_1)P(X|Z_1)}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_1)}{P(X=x_i)}} \neq P(X|Z_1) \quad (27)$$

but when Z_1 and Z_2 vanish, because in such case Eq. (27) reduces to $P(X)$ on its left and right sides.

Remark 3: In the particular (two sources) degenerate case where Z_1 and Z_2 vanish, one has always: $Bayes(P(X), P(X); P(X)) = P(X)$. However, in the more general degenerate case (when considering more than 2 sources), one will have in general: $Bayes(P(X), P(X), \ldots, P(X); P(X)) \neq P(X)$, but when $P(X)$ is uniform, or when $P(X)$ is a "deterministic" probability measure such that $P(X = x_i) = 1$ for a given $x_i \in \Theta(X)$ and $P(X = x_j) = 0$ for all $x_j \neq x_i$.

• **(P3) : Bayes fusion rule is in general not associative.**

Proof: A fusion rule f is called associative if and only if it satisfies the associative law: $f(f(x,y),z) = f(x, f(y,z)) = f(y, f(x,z)) = f(x,y,z)$ for all possible inputs x, y and z. Let us prove that Bayes rule is not associative from a very simple example.

Example 1: Let us consider the simplest set of outcomes $\{x_1, x_2\}$ for X, with prior pmf:

$$P(X = x_1) = 0.2 \text{ and } P(X = x_2) = 0.8$$

and let us consider the three given sets of posterior pmfs:

$$\begin{cases} P(X = x_1|Z_1) = 0.1 \text{ and } P(X = x_2|Z_1) = 0.9 \\ P(X = x_1|Z_2) = 0.5 \text{ and } P(X = x_2|Z_2) = 0.5 \\ P(X = x_1|Z_3) = 0.6 \text{ and } P(X = x_2|Z_3) = 0.4 \end{cases}$$

Bayes fusion $Bayes(P(X|Z_1),)P(X|Z_2), P(X|Z_3); P(X))$ of the three sources altogether according to Eq. (16) provides:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2 \cap Z_3) = \frac{1}{K_{123}} \frac{0.1}{\sqrt[3]{0.2}} \frac{0.5}{\sqrt[3]{0.2}} \frac{0.6}{\sqrt[3]{0.2}} = 0.40 \\ P(X = x_2|Z_1 \cap Z_2 \cap Z_3) = \frac{1}{K_{123}} \frac{0.9}{\sqrt[3]{0.8}} \frac{0.5}{\sqrt[3]{0.8}} \frac{0.4}{\sqrt[3]{0.8}} = 0.60 \end{cases}$$

where the normalization constant K_{123} is given by:

$$K_{123} = \frac{0.1}{\sqrt[3]{0.2}} \frac{0.5}{\sqrt[3]{0.2}} \frac{0.6}{\sqrt[3]{0.2}} + \frac{0.9}{\sqrt[3]{0.8}} \frac{0.5}{\sqrt[3]{0.8}} \frac{0.4}{\sqrt[3]{0.8}} = 0.3750$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_2)$ using $Bayes(P(X|Z_1), P(X|Z_2); P(X))$. One has:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2) = \frac{1}{K_{12}} \frac{0.1}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} \approx 0.3077 \\ P(X = x_2|Z_1 \cap Z_2) = \frac{1}{K_{12}} \frac{0.9}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} \approx 0.6923 \end{cases}$$

where the normalization constant K_{12} is given by:

$$K_{12} = \frac{0.1}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} = 0.8125$$

Let us compute the fusion of $P(X|Z_2)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_2), P(X|Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|Z_2 \cap Z_3) = \frac{1}{K_{23}} \frac{0.5}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} \approx 0.8571 \\ P(X = x_2|Z_2 \cap Z_3) = \frac{1}{K_{23}} \frac{0.5}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 0.1429 \end{cases}$$

where the normalization constant K_{23} is given by:

$$K_{23} = \frac{0.5}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.5}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 1.75$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_1), P(X|Z_3); P(X))$. One has:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_3) = \frac{1}{K_{13}} \frac{0.1}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} = 0.4 \\ P(X = x_2|Z_1 \cap Z_3) = \frac{1}{K_{13}} \frac{0.9}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 0.6 \end{cases}$$

where the normalization constant K_{13} is given by:

$$K_{13} = \frac{0.1}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 0.75$$

Let us compute the fusion of $P(X|Z_1 \cap Z_2)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_1 \cap Z_2), P(X|Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|(Z_1 \cap Z_2) \cap Z_3) = \frac{1}{K_{(12)3}} \frac{0.3077}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|(Z_1 \cap Z_2) \cap Z_3) = \frac{1}{K_{(12)3}} \frac{0.6923}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{(12)3}$ is given by

$$K_{(12)3} = \frac{0.3077}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.6923}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 1.26925$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_2 \cap Z_3)$ using $Bayes(P(X|Z_1), P(X|Z_2 \cap Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|Z_1 \cap (Z_2 \cap Z_3)) = \frac{1}{K_{1(23)}} \frac{0.1}{\sqrt{0.2}} \frac{0.8571}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|Z_1 \cap (Z_2 \cap Z_3)) = \frac{1}{K_{1(23)}} \frac{0.9}{\sqrt{0.8}} \frac{0.1429}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{1(23)}$ is given by

$$K_{1(23)} = \frac{0.1}{\sqrt{0.2}} \frac{0.8571}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.1429}{\sqrt{0.8}} \approx 0.58931$$

Let us compute the fusion of $P(X|Z_1 \cap Z_3)$ with $P(X|Z_2)$ using $Bayes(P(X|Z_1 \cap Z_3), P(X|Z_2); P(X))$. One has

$$\begin{cases} P(X = x_1|(Z_1 \cap Z_3) \cap Z_2) = \frac{1}{K_{(13)2}} \frac{0.4}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|(Z_1 \cap Z_3) \cap Z_2) = \frac{1}{K_{(13)2}} \frac{0.6}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{(13)2}$ is given by

$$K_{(13)2} = \frac{0.4}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} + \frac{0.6}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} = 1.375$$

Therefore, one sees that even if in our example one has $f(x, f(y,z)) = f(f(x,y),z) = f(y, f(x,z))$ because $P(X|(Z_1 \cap Z_2) \cap Z_3) = P(X|Z_1 \cap (Z_2 \cap Z_3)) = P(X|Z_2 \cap (Z_1 \cap Z_3))$, Bayes fusion rule is not associative since:

$$\begin{cases} P(X|(Z_1 \cap Z_2) \cap Z_3) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_1 \cap (Z_2 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_2 \cap (Z_1 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \end{cases}$$

• **(P4) : Bayes fusion rule is associative if and only if $P(X)$ is uniform.**

Proof: If $P(X)$ is uniform, Bayes fusion rule is given by Eq. (21) which can be rewritten as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{P(X|Z_s) \prod_{k=1}^{s-1} P(X|Z_k)}{\sum_{i=1}^{N} P(X = x_i|Z_s) \prod_{k=1}^{s-1} P(X = x_i|Z_k)}$$

By introducing the term $1/\sum_{i=1}^{N} \prod_{k=1}^{s-1} P(X = x_i|Z_k)$ in numerator and denominator of the previous formula, it comes:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\frac{\prod_{k=1}^{s-1} P(X|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s-1} P(X=x_i|Z_k)} P(X|Z_s)}{\sum_{i=1}^{N} \frac{\prod_{k=1}^{s-1} P(X=x_i|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s-1} P(X=x_i|Z_k)} P(X = x_i|Z_s)}$$

which can be simply rewritten as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{P(X|Z_1 \cap \ldots \cap Z_{s-1}) P(X|Z_s)}{\sum_{i=1}^{N} P(X = x_i|Z_1 \cap \ldots \cap Z_{s-1}) P(X = x_i|Z_s)}$$

Therefore when $P(X)$ is uniform, one has:

$$Bayes(P(X|Z_1), \ldots, P(X|Z_s)) =$$
$$Bayes(Bayes(P(X|Z_1), \ldots, P(X|Z_{s-1})), P(X|Z_s))$$

The previous relation was based on the decomposition of $\prod_{k=1}^{s} P(X|Z_k)$ as $P(X|Z_s) \prod_{k=1}^{s-1} P(X|Z_k)$. This choice of decomposition was arbitrary and chosen only for convenience. In fact $\prod_{k=1}^{s} P(X|Z_k)$ can be decomposed in s different manners, as $P(X|Z_j) \prod_{k=1|k \neq j}^{s} P(X|Z_k)$, $j = 1, 2, \ldots s$ and the similar analysis can be done. In particular, when $s = 3$, we will have:

$$Bayes(P(X|Z_1), P(X|Z_2), P(X|Z_3)) =$$
$$Bayes(Bayes(P(X|Z_1), P(X|Z_2)), P(X|Z_3))$$
$$= Bayes(P(X|Z_1), Bayes(P(X|Z_2), P(X|Z_3)))$$

which completes the proof.

• **(P5) : The levels of global agreement and global conflict between the sources do not matter in Bayes fusion rule.**

Proof: This property seems surprising at first glance, but, since the results of Bayes fusion is nothing but the ratio of the agreement on x_i ($i = 1, 2, \ldots, N$) over the global agreement factor, many distinct sources with different global agreements (and thus with different global conflicts) can yield same Bayes fusion result. Indeed, the ratio is kept unchanged when multiplying its numerator and denominator by same non null scalar value. Consequently, the absolute levels of global agreement between the sources (and therefore of global conflict

also) do not matter in Bayes fusion result. What really matters is only the proportions of relative agreement factors.

Example 2: To illustrate this property, let us consider Bayes fusion rule applied to two distinct sets[3] of sources represented by $Bayes(P(X|Z_1), P(X|Z_2); P(X))$ and by $Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$ with the following prior and posterior pmfs:

$$P(X = x_1) = 0.2 \text{ and } P(X = x_2) = 0.8$$

$$\begin{cases} P(X = x_1|Z_1) \approx 0.0607 \text{ and } P(X = x_2|Z_1) \approx 0.9393 \\ P(X = x_1|Z_2) \approx 0.6593 \text{ and } P(X = x_2|Z_2) \approx 0.3407 \end{cases}$$

$$\begin{cases} P'(X = x_1|Z_1) \approx 0.8360 \text{ and } P'(X = x_2|Z_1) \approx 0.1640 \\ P'(X = x_1|Z_2) \approx 0.0240 \text{ and } P'(X = x_2|Z_2) \approx 0.9760 \end{cases}$$

Applying Bayes fusion rule given by Eq. (5), one gets for $Bayes(P(X|Z_1), P(X|Z_2); P(X))$:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2) = \frac{0.2}{0.2+0.4} = 1/3 \\ P(X = x_2|Z_1 \cap Z_2) = \frac{0.4}{0.2+0.4} = 2/3 \end{cases} \qquad (28)$$

Similarly, one gets for $Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$

$$\begin{cases} P'(X = x_1|Z_1 \cap Z_2) = \frac{0.1}{0.1+0.2} = 1/3 \\ P'(X = x_2|Z_1 \cap Z_2) = \frac{0.2}{0.1+0.2} = 2/3 \end{cases} \qquad (29)$$

Therefore, one sees that $Bayes(P(X|Z_1), P(X|Z_2); P(X)) = Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$ even if the levels of global agreements (and global conflicts) are different. In this particular example, one has:

$$\begin{cases} (GA_2 = 0.60) \neq (GA'_2 = 0.30) \\ (GC_2 = 1.60) \neq (GC'_2 = 2.05) \end{cases} \qquad (30)$$

In summary, different sets of sources to combine (with different levels of global agreement and global conflict) can provide exactly the same result once combined with Bayes fusion rule. Hence the different levels of global agreement and global conflict do not really matter in Bayes fusion rule. What really matters in Bayes fusion rule is only the distribution of all the relative agreement factors defined as $A_s(X = x_i)/GA_s$.

III. Belief functions and Dempster's rule

The Belief Functions (BF) have been introduced in 1976 by Glenn Shafer in his mathematical theory of evidence [4], also known as Dempster-Shafer Theory (DST) in order to reason under uncertainty and to model epistemic uncertainties. We will not present in details the foundations of DST, but only the basic mathematical definitions that are necessary for the scope of this paper. The emblematic fusion rule proposed by Shafer to combine sources of evidences characterized by their basic belief assignments (bba) is Dempster's rule that will be analyzed in details in the sequel. In the literature over the years, DST has been widely defended by its proponents in arguing that: 1) Probability measures are particular cases of Belief

---

[3]The values chosen for $P(X|Z_1)$, $P(X|Z_2)$, $P'(X|Z_1)$, $P'(X|Z_2)$ here have been approximated at the fourth digit. They can be precisely determined such that the expressions for $P(X|Z_1 \cap Z_2)$ and $P'(X|Z_1 \cap Z_2)$ as given in Eqs. (28) and (29) hold. For example, the exact value of $P(x_1|Z_2)$ is obtained by solving a polynomial equation of degree 2 having as a possible solution $P(x_1|Z_2) = \frac{1}{2}(0.72 + \sqrt{0.72^2 - 4 \times 0.04}) = 0.659332590941915 \approx 0.6593$, etc.

functions; and 2) Dempster's fusion rule is a generalization of Bayes fusion rule. Although the statement 1) is correct because Probability measures are indeed particular (additive) Belief functions (called as Bayesian belief functions), we will explain why the second statement about Dempster's rule is incorrect in general.

A. Belief functions

Let Θ be a frame of discernment of a problem under consideration. More precisely, the set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ consists of a list of N exhaustive and exclusive elements θ_i, $i = 1, 2, \ldots, N$. Each θ_i represents a possible state related to the problem we want to solve. The exhaustivity and exclusivity of elements of Θ is referred as Shafer's model of the frame Θ. A basic belief assignment (bba), also called a belief mass function, $m(.) : 2^\Theta \to [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that verifies the following conditions [4]:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in 2^\Theta} m(X) = 1 \tag{31}$$

The quantity $m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The set $\mathcal{F}(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$ of all focal elements of a bba $m(.)$ is called the core of the bba. A bba $m(.)$ is said Bayesian if its focal elements are singletons of 2^Θ. The vacuous bba characterizing the total ignorance denoted[4] $I_t = \theta_1 \cup \theta_2 \cup \ldots \cup \theta_N$ is defined by $m_v(.) : 2^\Theta \to [0; 1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(I_t) = 1$.

From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined for $\forall X \in 2^\Theta$ as:

$$\begin{cases} Bel(X) = \sum_{Y \in 2^\Theta | Y \subseteq X} m(Y) \\ Pl(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m(Y) \end{cases} \tag{32}$$

$Bel(X)$ represents the whole mass of belief that comes from all subsets of Θ included in X. It is interpreted as the lower bound of the probability of X, i.e. $P_{\min}(X)$. $Bel(.)$ is a subadditive measure since $\sum_{\theta_i \in \Theta} Bel(\theta_i) \leq 1$. $Pl(X)$ represents the whole mass of belief that comes from all subsets of Θ compatible with X (i.e., those intersecting X). $Pl(X)$ is interpreted as the upper bound of the probability of X, i.e. $P_{\max}(X)$. $Pl(.)$ is a superadditive measure since $\sum_{\theta_i \in \Theta} Pl(\theta_i) \geq 1$. $Bel(X)$ and $Pl(X)$ are classically seen [4] as lower and upper bounds of an unknown probability $P(.)$, and one has the following inequality satisfied $\forall X \in 2^\Theta$: $Bel(X) \leq P(X) \leq Pl(X)$. The belief function $Bel(.)$ (and the plausibility function $Pl(.)$) built from any Bayesian bba $m(.)$ can be interpreted as a (subjective) conditional probability measure provided by a given source of evidence, because if the bba $m(.)$ is *Bayesian* the following equality always holds [4]: $Bel(X) = Pl(X) = P(X)$.

B. Dempster's rule of combination

Dempster's rule of combination, denoted DS rule[5] is a mathematical operation, represented symbolically by $\oplus$, which corresponds to the normalized conjunctive fusion rule. Based on Shafer's model of Θ, the combination of $s > 1$ independent and distinct sources of evidences characterized by their bba $m_1(.)$, ..., $m_s(.)$ related to the same frame of discernment Θ is denoted $m_{DS}(.) = [m_1 \oplus \ldots \oplus m_s](.)$. The quantity $m_{DS}(.)$ is defined mathematically as follows: $m_{DS}(\emptyset) \triangleq 0$ and $\forall X \neq \emptyset \in 2^\Theta$

$$m_{DS}(X) \triangleq \frac{m_{12\ldots s}(X)}{1 - K_{12\ldots s}} \tag{33}$$

where the conjunctive agreement on X is given by:

$$m_{12\ldots s}(X) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = X}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{34}$$

and where the global conflict is given by:

$$K_{12\ldots s} \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = \emptyset}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{35}$$

When $K_{12\ldots s} = 1$, the s sources are in total conflict and their combination cannot be computed with DS rule because Eq. (33) is mathematically not defined due to 0/0 indeterminacy [4]. DS rule is commutative and associative which makes it very attractive from engineering implementation standpoint.

It has been proved in [4] that the vacuous bba $m_v(.)$ is a neutral element for DS rule because $[m \oplus m_v](.) = [m_v \oplus m](.) = m(.)$ for any bba $m(.)$ defined on 2^Θ. This property looks reasonable since a total ignorant source should not impact the fusion result because it brings no information that can be helpful for the discrimination between the elements of the power set 2^Θ.

IV. ANALYSIS OF COMPATIBILITY OF DEMPSTER'S RULE WITH BAYES RULE

To analyze the compatibility of Dempster's rule with Bayes rule, we need to work in the probabilistic framework because Bayes fusion rule has been developed only in this theoretical framework. So in the sequel, we will manipulate only probability mass functions (pmfs), related with Bayesian bba's in the Belief Function framework. This perfectly justifies the restriction of singleton bba as a prior bba since we want to manipulate prior probabilities to make a fair comparison of results provided by both rules. If Dempster's rule is a true (consistent) generalization of Bayes fusion rule, it must provide same results as Bayes rule when combining Bayesian bba's, otherwise Dempster's rule cannot be fairly claimed to be a generalization of Bayes fusion rule. In this section, we analyze the real (partial or total) compatibility of Dempster's rule with Bayes fusion rule. Two important cases must be analyzed depending on the nature of the prior information $P(X)$ one has in hands for performing the fusion of the sources. These

---

[4]The set $\{\theta_1, \theta_2, \ldots, \theta_N\}$ and the complete ignorance $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_N$ are both denoted Θ in DST.

[5]We denote it DS rule because it has been proposed historically by Dempster [2], [3], and widely promoted by Shafer in the development of DST [4].

sources to combine will be characterized by the following Bayesian bba's:

$$\begin{cases} m_1(.) \triangleq \{m_1(\theta_i) = P(X = x_i|Z_1), i = 1, 2, \ldots, N\} \\ \vdots \quad \vdots \qquad\qquad\qquad \vdots \\ m_s(.) \triangleq \{m_s(\theta_i) = P(X = x_i|Z_s), i = 1, 2, \ldots, N\} \end{cases}$$
(36)

The prior information is characterized by a given bba denoted as $m_0(.)$ that can be defined either on 2^Θ, or only on Θ if we want to deal for the needs of our analysis with a Bayesian prior. In the latter case, if $m_0(.) \triangleq \{m_0(\theta_i) = P(X = x_i), i = 1, 2, \ldots, N\}$ then $m_0(.)$ plays the same role as the prior pmf $P(X)$ in the probabilistic framework.

When considering a non vacuous prior $m_0(.) \neq m_v(.)$, we denote Dempster's combination of s sources symbolically as:

$$m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_0(.))$$

When the prior bba is vacuous $m_0(.) = m_v(.)$ then $m_0(.)$ has no impact on Dempster's fusion result, and so we denote symbolically Dempster's rule as:

$$m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_v(.))$$
$$= DS(m_1(.), \ldots, m_s(.))$$

A. Case 1: Uniform Bayesian prior

It is important to note that Dempster's fusion formula proposed by Shafer in [4] and recalled in Eq. (33) makes no real distinction between the nature of sources to combine (if they are posterior or prior information). In fact, the formula (33) reduces exactly to Bayes rule given in Eq. (25) if the bba's to combine are Bayesian and if the prior information is either uniform or vacuous. Stated otherwise the following functional equality holds

$$DS(m_1(.), \ldots, m_s(.); m_0(.)) \equiv$$
$$Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \quad (37)$$

as soon as all bba's $m_i(.)$, $i = 1, 2, \ldots, s$ are Bayesian and coincide with $P(X|Z_i)$, $P(X)$ is uniform, and either the prior bba $m_0(.)$ is vacuous $(m_0(.) = m_v(.))$, or $m_0(.)$ is the uniform Bayesian bba.

Example 3: Let us consider $\Theta(X) = \{x_1, x_2, x_3\}$ with two distinct sources providing the following Bayesian bba's

$$\begin{cases} m_1(x_1) = P(X = x_1|Z_1) = 0.2 \\ m_1(x_2) = P(X = x_2|Z_1) = 0.3 \\ m_1(x_3) = P(X = x_3|Z_1) = 0.5 \end{cases} \text{and} \begin{cases} m_2(x_1) = 0.5 \\ m_2(x_2) = 0.1 \\ m_2(x_3) = 0.4 \end{cases}$$

• If we choose as prior $m_0(.)$ the vacuous bba, that is $m_0(x_1 \cup x_2 \cup x_3) = 1$, then one will get

$$\begin{cases} m_{DS}(x_1) & = \frac{1}{1-K_{12}^{vacuous}} m_1(x_1) m_2(x_1) m_0(x_1 \cup x_2 \cup x_3) \\ & = \frac{1}{1-0.67} 0.2 \cdot 0.5 \cdot 1 = \frac{0.10}{0.33} \approx 0.3030 \\ m_{DS}(x_2) & = \frac{1}{1-K_{12}^{vacuous}} m_1(x_2) m_2(x_2) m_0(x_1 \cup x_2 \cup x_3) \\ & = \frac{1}{1-0.67} 0.3 \cdot 0.1 \cdot 1 = \frac{0.03}{0.33} \approx 0.0909 \\ m_{DS}(x_3) & = \frac{1}{1-K_{12}^{vacuous}} m_1(x_3) m_2(x_3) m_0(x_1 \cup x_2 \cup x_3) \\ & = \frac{1}{1-0.67} 0.5 \cdot 0.4 \cdot 1 = \frac{0.20}{0.33} \approx 0.6061 \end{cases}$$

with

$$K_{12}^{vacuous} = 1 - m_1(x_1) m_2(x_1) m_0(x_1 \cup x_2 \cup x_3)$$
$$- m_1(x_2) m_2(x_2) m_0(x_1 \cup x_2 \cup x_3)$$
$$- m_1(x_3) m_2(x_3) m_0(x_1 \cup x_2 \cup x_3) = 0.67$$

• If we choose as prior $m_0(.)$ the uniform Bayesian bba given by $m_0(x_1) = m_0(x_2) = m_0(x_3) = 1/3$, then we get

$$\begin{cases} m_{DS}(x_1) & = \frac{1}{1-K_{12}^{uniform}} m_1(x_1) m_2(x_1) m_0(x_1) \\ & = \frac{1}{1-0.89} 0.2 \cdot 0.5 \cdot 1/3 = \frac{0.10/3}{0.11} \approx 0.3030 \\ m_{DS}(x_2) & = \frac{1}{1-K_{12}^{uniform}} m_1(x_2) m_2(x_2) m_0(x_2) \\ & = \frac{1}{1-0.89} 0.3 \cdot 0.1 \cdot 1/3 = \frac{0.03/3}{0.11} \approx 0.0909 \\ m_{DS}(x_3) & = \frac{1}{1-K_{12}^{uniform}} m_1(x_3) m_2(x_3) m_0(x_3) \\ & = \frac{1}{1-0.89} 0.5 \cdot 0.4 \cdot 1/3 = \frac{0.20/3}{0.11} \approx 0.6061 \end{cases}$$

where the degree of conflict when $m_0(.)$ is Bayesian and uniform is now given by $K_{12}^{uniform} = 0.89$.

Clearly $K_{12}^{uniform} \neq K_{12}^{vacuous}$, but the fusion results obtained with two distinct priors $m_0(.)$ (vacuous or uniform) are the same because of the algebraic simplification by 1/3 in Dempster's fusion formula when using uniform Bayesian bba. When combining Bayesian bba's $m_1(.)$ and $m_2(.)$, the vacuous prior and uniform prior $m_0(.)$ have therefore no impact on the result. Indeed, they contain no information that may help to prefer one particular state x_i with respect to the other ones, even if the level of conflict is different in both cases. So, the level of conflict doesn't matter at all in such Bayesian case. As already stated, what really matters is only the distribution of relative agreement factors. It can be easily verified that we obtain same results when applying Bayes Eq. (14), or (16).

Only in such very particular cases (i.e. Bayesian bba's, and vacuous or Bayesian uniform priors), Dempster's rule is fully consistent with Bayes fusion rule. So the claim that Dempster's is a generalization of Bayes rule is true in this very particular case only, and that is why such claim has been widely used to defend Dempster's rule and DST thanks to its compatibility with Bayes fusion rule in that very particular case. Unfortunately, such compatibility is only partial and not general because it is not longer valid when considering the more general cases involving non uniform Bayesian prior bba's as shown in the next subsection.

B. Case 2: Non uniform Bayesian prior

Let us consider Dempster's fusion of Bayesian bba's with a Bayesian non uniform prior $m_0(.)$. In such case it is easy to check from the general structures of Bayes fusion rule (16) and Dempster's fusion rule (33) that these two rules are incompatible. Indeed, in Bayes rule one divides each posterior source $m_i(x_j)$ by $\sqrt[s]{m_0(x_j)}$, $i = 1, 2, \ldots s$, whereas the prior source $m_0(.)$ is combined in a pure conjunctive manner by Dempster's rule with the bba's $m_i(.)$, $i = 1, 2, \ldots s$, as if $m_0(.)$ was a simple additional source. This difference of processing prior information between the two approaches explains clearly the incompatibility of Dempster's rule with Bayes rule when Bayesian prior bba is not uniform. This incompatibility is illustrated in the next simple example. Mahler and Fixsen have already proposed in [23], [24], [25] a modification of

Dempster's rule to force it to be compatible with Bayes rule when combining Bayesian bba's. The analysis of such modified Dempster's rule is out of the scope of this paper.

Example 4: Let us consider the same frame $\Theta(X)$, and same bba's $m_1(.)$ and $m_2(.)$ as in the Example 3. Suppose that the prior information is Bayesian and non uniform as follows: $m_0(x_1) = P(X = x_1) = 0.6$, $m_0(x_2) = P(X = x_2) = 0.3$ and $m_0(x_3) = P(X = x_3) = 0.1$. Applying Bayes rule (12) yields:

$$\begin{cases} P(x_1|Z_1 \cap Z_2) &= \frac{A_2(x_1)}{GA_2} = \frac{0.2 \cdot 0.5/0.6}{2.2667} = \frac{0.1667}{2.2667} \approx 0.0735 \\ P(x_2|Z_1 \cap Z_2) &= \frac{A_2(x_2)}{GA_2} = \frac{0.3 \cdot 0.1/0.3}{2.2667} = \frac{0.1000}{2.2667} \approx 0.0441 \\ P(x_3|Z_1 \cap Z_2) &= \frac{A_2(x_3)}{GA_2} = \frac{0.5 \cdot 0.4/0.1}{2.2667} = \frac{2.0000}{2.2667} \approx 0.8824 \end{cases}$$

Applying Dempster's rule yields $m_{DS}(x_i) \neq P(x_i|Z_1 \cap Z_2)$ because:

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-0.9110} \cdot 0.2 \cdot 0.5 \cdot 0.6 = \frac{0.060}{0.089} \approx 0.6742 \\ m_{DS}(x_2) &= \frac{1}{1-0.9110} \cdot 0.3 \cdot 0.1 \cdot 0.3 = \frac{0.009}{0.089} \approx 0.1011 \\ m_{DS}(x_3) &= \frac{1}{1-0.9110} \cdot 0.5 \cdot 0.4 \cdot 0.1 = \frac{0.020}{0.089} \approx 0.2247 \end{cases}$$

Therefore, one has in general[6]:

$$DS(m_1(.), \ldots, m_s(.); m_0(.)) \neq \\ Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \quad (38)$$

V. Conclusions

In this paper, we have analyzed in details the expression and the properties of Bayes rule of combination based on statistical conditional independence assumption, as well as the emblematic Dempster's rule of combination of belief functions introduced by Shafer in his Mathematical Theory of evidence. We have clearly explained from a theoretical standpoint, and also on simple examples, why Dempster's rule is not a generalization of Bayes rule in general. The incompatibility of Dempster's rule with Bayes rule is due to its impossibility to deal with non uniform Bayesian priors in the same manner as Bayes rule does. Dempster's rule turns to be compatible with Bayes rule only in two very particular cases: 1) if all the Bayesian bba's to combine (including the prior) focus on same state (i.e. there is a perfect conjunctive consensus between the sources), or 2) if all the bba's to combine (excluding the prior) are Bayesian, and if the prior bba cannot help to discriminate a particular state of the frame of discernment (i.e. the prior bba is either vacuous, or Bayesian and uniform). Except in these two very particular cases, Dempster's rule is totally incompatible with Bayes rule. Therefore, Dempster's rule cannot be claimed to be a generalization of Bayes fusion rule, even when the bba's to combine are Bayesian.

Acknowledgment

---

[6]but in the very degenerate case when manipulating *deterministic* Bayesian bba's, which is of little practical interest from the fusion standpoint.

References

[1] L.A. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, CA, U.S.A., 1979.

[2] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist., Vol. 38, pp. 325–339, 1967.

[3] A. Dempster, A generalization of bayesian inference, J. R. Stat. Soc. B 30, pp. 205–247, 1968.

[4] G. Shafer, A Mathematical theory of evidence, Princeton University Press, Princeton, NJ, U.S.A., 1976.

[5] L.A. Zadeh, Book review: A mathematical theory of evidence, The AI Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[6] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The AI Magazine, Vol. 7, No. 2, pp. 85–90, 1986.

[7] J. Lemmer, Confidence factors, empiricism and the Dempster-Shafer theory of evidence, Proc. of UAI '85, pp. 160–176, Los Angeles, CA, U.S.A., July 10–12, 1985.

[8] J. Pearl, Do we need higher-order probabilities, and, if so, what do they mean?, Proc. of UAI '87, pp. 47–60, Seattle, WA, U.S.A., July 10–12, 1987.

[9] F. Voorbraak, On the justification of Demspster's rule of combination, Dept. of Phil., Utrecht Univ., Logic Group, Preprint Ser., No. 42, Dec. 1988.

[10] G. Shafer, Perspectives on the theory and practice of belief functions, IJAR, Vol. 4, No. 5–6, pp. 323–362, 1990.

[11] J. Pearl, Reasoning with belief functions: an analysis of compatibility, IJAR, Vol. 4, No. 5–6, pp. 363–390, 1990.

[12] J. Pearl, Rejoinder of comments on "Reasoning with belief functions: An analysis of compatibility", IJAR, Vol. 6, No. 3, pp. 425–443, May 1992.

[13] G.M. Provan, The validity of Dempster-Shafer belief functions, IJAR, Vol. 6., No. 3, pp. 389–399, May 1992.

[14] P. Wang, A defect in Dempster-Shafer theory, Proc. of UAI '94, pp. 560–566, Seattle, WA, U.S.A., July 29–31, 1994.

[15] A. Gelman, The boxer, the wrestler, and the coin flip: a paradox of robust bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp. 146–150, 2006.

[16] F. Smarandache, J. Dezert (Editors), Applications and advances of DSmT for information fusion, Vol. 3, ARP, U.S.A., 2009. http://fs.gallup.unm.edu/DSmT.htm

[17] J. Dezert, P. Wang, A. Tchamova, On the validity of Dempster-Shafer theory, Proc. of Fusion 2012 Int. Conf., Singapore, July 9–12, 2012.

[18] A. Tchamova, J. Dezert, On the Behavior of Dempster's rule of combination and the foundations of Dempster-Shafer theory, (best paper award), Proc. of 6th IEEE Int. Conf. on Intelligent Systems IS '12, Sofia, Bulgaria, Sept. 6–8, 2012.

[19] R.P. Mahler, Statistical multisource-multitarget information fusion, Chapter 4, Artech House, 2007.

[20] X.R. Li, Probability, random signals and statistics, CRC Press, 1999.

[21] P.E. Pfeiffer, Applied probability, Connexions Web site, Aug. 31, 2009. http://cnx.org/content/col10708/1.6/

[22] G. Shafer, Non-additive probabilities in the work of Bernoulli and Lambert, in Archive for History of Exact Sciences, C. Truesdell (Ed.), Springer-Verlag, Berlin, Vol. 19, No. 4, pp. 309–370, 1978.

[23] R.P. Mahler, Using a priori evidence to customize Dempster-Shafer theory, Proc. of 6th Nat. Symp. on Sensor Fusion, Vol. 1, pp. 331–345, Orlando, FL, U.S.A., April 13–15, 1993.

[24] R.P. Mahler, Classification when a priori evidence is ambiguous, Proc. SPIE Conf. on Opt. Eng. in Aerospace Sensing (Automatic Object Recognition IV), pp. 296–304, Orlando, FL, U.S.A., April 4–8, 1994.

[25] D. Fixsen, R.P. Mahler, The modified Dempster-Shafer approach to classification, IEEE Trans. on SMC, Part A, Vol. 27, No. 1, pp. 96–104, 1997.

On the Consistency of PCR6 with the Averaging Rule and its Application to Probability Estimation

Florentin Smarandache
Jean Dezert

Abstract—Since the development of belief function theory introduced by Shafer in seventies, many combination rules have been proposed in the literature to combine belief functions specially (but not only) in high conflicting situations because the emblematic Dempster's rule generates counter-intuitive and unacceptable results in practical applications. Many attempts have been done during last thirty years to propose better rules of combination based on different frameworks and justifications. Recently in the DSmT (Dezert-Smarandache Theory) framework, two interesting and sophisticate rules (PCR5 and PCR6 rules) have been proposed based on the Proportional Conflict Redistribution (PCR) principle. These two rules coincide for the combination of two basic belief assignments, but they differ in general as soon as three or more sources have to be combined altogether because the PCR used in PCR5 and in PCR6 are different. In this paper we show why PCR6 is better than PCR5 to combine three or more sources of evidence and we prove the coherence of PCR6 with the simple Averaging Rule used classically to estimate the probability based on the frequentist interpretation of the probability measure. We show that such probability estimate cannot be obtained using Dempster-Shafer (DS) rule, nor PCR5 rule.

Keywords: Information fusion, belief functions, PCR6, PCR5, DSmT, frequentist probability.

I. INTRODUCTION

In this paper, we work with belief functions [1] defined from the finite and discrete frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$. In Dempster-Shafer Theory (DST) framework, basic belief assignments (bba's) provided by the distinct sources of evidence are defined on the fusion space $2^\Theta = (\Theta, \cup)$ consisting in the power-set of Θ, that is the set of elements of Θ and those generated from Θ with the union set operator. Such fusion space assumes that the elements of Θ are non-empty, exhaustive and exclusive, which is called Shafer's model of Θ. More generally, in Dezert-Smarandache Theory (DSmT) [2], the fusion space denoted G^Θ can also be either the hyper-power set $D^\Theta = (\Theta, \cup, \cap)$ (Dedekind's lattice), or super-power set[1] $S^\Theta = (\Theta, \cup, \cap, c(.))$ depending on the underlying model of the frame of discernment we choose to fit with the nature of the problem. Details on DSm models are given in [2], Vol. 1.

We assume that $s \geq 2$ basic belief assignments (bba's) $m_i(.)$, $i = 1, 2, \ldots, s$ provided by s distinct sources of evidences defined on the fusion space G^Θ are available and we need to combine them for a final decision-making purpose.

For doing this, many rules of combination have been proposed in the literature, the most emblematic ones being the simple Averaging Rule, Dempster-Shafer (DS) rule, and more recently the PCR5 and PCR6 fusion rules.

The contribution of this paper is to analyze in deep the behavior of PCR5 and PCR6 fusion rules and to explain why we consider more preferable to use PCR6 rule rather than PCR5 rule for combining several distinct sources of evidence altogether. We will show in details the strong relationship between PCR6 and the averaging fusion rule which is commonly used to estimate the probabilities in the classical frequentist interpretation of probabilities.

This paper is organized as follows. In section II, we briefly recall the background on belief functions and the main fusion rules used in this paper. Section III demonstrates the consistency of PCR6 fusion rule with the Averaging Rule for binary masses in total conflict as well as the ability of PCR6 to discriminate asymmetric fusion cases for the fusion of Bayesian bba's. Section IV shows that PCR6 can also be used to estimate empirical probability in a simple (coin tossing) random experiment. Section V will conclude and open challenging problem about the recursivity of fusion rules formulas that are sought for efficient implementations.

II. BACKGROUND ON BELIEF FUNCTIONS

A. Basic belief assignment

Lets' consider a finite discrete frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, $n > 1$ of the fusion problem under consideration and its fusion space G^Θ which can be chosen either as 2^Θ, D^Θ or S^Θ depending on the model that fits with the problem. A basic belief assignment (bba) associated with a given source of evidence is defined as the mapping $m(.) : G^\Theta \rightarrow [0, 1]$ satisfying $m(\emptyset) = 0$ and $\sum_{A \in G^\Theta} m(A) = 1$. The quantity $m(A)$ is called mass of belief of A committed by the source of evidence. If $m(A) > 0$ then A is called a focal element of the bba $m(.)$. When all focal elements are singletons and $G^\Theta = 2^\Theta$ then $m(.)$ is called a Bayesian bba [1] and it is homogeneous to a (possibly subjective) probability measure. The vacuous bba representing a totally ignorant source is defined as $m_v(\Theta) = 1$. Belief and plausibility functions are defined by

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in G^\Theta}} m(B) \quad \text{and} \quad \text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in G^\Theta}} m(B) \quad (1)$$

---

[1]$\cap$ and $c(.)$ are respectively the set intersection and complement operators.

B. Fusion rules

The main information fusion problem in the belief function frameworks (DST or DSmT) is how to combine efficiently several distinct sources of evidence represented by $m_1(.)$, $m_2(.)$, ..., $m_s(.)$ ($s \geq 2$) bba's defined on G^Θ. Many rules have been proposed for such task – see [2], Vol. 2, for a detailed list of fusion rules – and we focus here on the following ones: 1) the Averaging Rule because it is the simplest one and it is used to empirically estimate probabilities in random experiment, 2) DS rule because it was historically proposed in DST, and 3) PCR5 and PCR6 rules because they were proposed in DSmT and have shown to provide better results than the DS rule in all applications where they have been tested so far. So we just briefly recall how these rules are mathematically defined.

• Averaging fusion rule $m_{1,2,...,s}^{Average}(.)$

For any X in G^Θ, $m_{1,2,...,s}^{Average}(X)$ is defined by

$$m_{1,2,...,s}^{Average}(X) = Average(m_1, m_2, \ldots, m_s) \triangleq \frac{1}{s} \sum_{i=1}^{s} m_i(X) \tag{2}$$

Note that the vacuous bba $m_v(\Theta) = 1$ is not a neutral element for this rule. This Averaging Rule is commutative but it is not associative because in general

$$m_{1,2,3}^{Average}(X) = \frac{1}{3}[m_1(X) + m_2(X) + m_3(X)]$$

is different from

$$m_{(1,2),3}^{Average}(X) = \frac{1}{2}[\frac{m_1(X) + m_2(X)}{2} + m_3(X)]$$

which is also different from

$$m_{1,(2,3)}^{Average}(X) = \frac{1}{2}[m_1(A) + \frac{m_2(X) + m_3(X)}{2}]$$

and also from

$$m_{(1,3),2}^{Average}(X) = \frac{1}{2}[\frac{m_1(X) + m_3(X)}{2} + m_2(X)]$$

In fact, it is easy to prove that the following recursive formula holds

$$m_{1,2,...,s}^{Average}(X) = \frac{s-1}{s} m_{1,2,...,s-1}^{Average}(X) + \frac{1}{s} m_s(X) \tag{3}$$

This simple averaging fusion rule has been used since more than two centuries for estimating empirically the probability measure in random experiments [3], [4].

• Dempster-Shafer fusion rule $m_{1,2,...,s}^{DS}(.)$

In DST framework, the fusion space G^Θ equals the power-set 2^Θ because Shafer's model of the frame Θ is assumed. The combination of $s \geq 2$ distinct sources of evidences characterized by the bba's $m_i(.)$, $i = 1, 2, \ldots, s$, is done with DS rule as follows [1]: $m_{1,2,...,s}^{DS}(\emptyset) = 0$ and for all $X \neq \emptyset$ in 2^Θ

$$m_{1,2,...,s}^{DS}(X) \triangleq \frac{1}{K_{1,2,...,s}} \sum_{\substack{X_1,X_2,...,X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s = X}} \prod_{i=1}^{s} m_i(X_i) \tag{4}$$

where the numerator of (4) is the mass of belief on the conjunctive consensus on X, and where $K_{1,2,...,s}$ is a normalization constant defined by

$$K_{1,2,...,s} = \sum_{\substack{X_1,X_2,...,X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s \neq \emptyset}} \prod_{i=1}^{s} m_i(X_i) = 1 - m_{1,2,...,s}(\emptyset)$$

The total degree of conflict between the s sources of evidences is defined by

$$m_{1,2,...,s}(\emptyset) = \sum_{\substack{X_1,X_2,...,X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s = \emptyset}} \prod_{i=1}^{s} m_i(X_i)$$

The sources are said in total conflict when $m_{1,2,...,s}(\emptyset) = 1$.

The vacuous bba $m_v(\Theta) = 1$ is a neutral element for DS rule and DS rule is commutative and associative. It remains the milestone fusion rule of DST. The doubts on the validity of such fusion rule has been discussed by Zadeh in 1979 [5]–[7] based on a very simple example with two highly conflicting sources of evidence. Since 1980's, many criticisms have been done about the behavior and justification of such DS rule. More recently, Dezert et al. in [8], [9] have put in light other counter-intuitive behaviors of DS rule even in low conflicting cases and showed serious flaws in logical foundations of DST.

• PCR5 and PCR6 fusion rules

To work in general fusion spaces G^Θ and to provide better fusion results in all (low or high conflicting) situations, several fusion rules have been developed in DSmT framework [2]. Among them, two fusion rules called PCR5 and PCR6 based on the proportional conflict redistribution (PCR) principle have been proved to work efficiently in all different applications where they have been used so far. The PCR principle transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved.

The general principle of PCR consists:

1) to apply the conjunctive rule;
2) calculate the total or partial conflicting masses;
3) then redistribute the (total or partial) conflicting mass proportionally on non-empty sets according to the integrity constraints one has for the frame Θ.

Because the proportional transfer can be done in two different ways, this has yielded to two different fusion rules. The PCR5 fusion rule has been proposed by Smarandache and Dezert in [2], Vol. 2, Chap. 1, and PCR6 fusion rule has been proposed by Martin and Osswald in [2], Vol. 2, Chap. 2.

We will not present in deep these two fusion rules since they have already been discussed in details with many examples in the aforementioned references but we only give their expressions for convenience here.

The general formula of PCR5 for the combination of $s \geq 2$ sources is given by $m_{1,2,\ldots,s}^{PCR5}(\emptyset) = 0$ and for $X \neq \emptyset$ in G^{Θ}

$$m_{1,2,\ldots,s}^{PCR5}(X) = m_{1,2,\ldots,s}(X) +$$

$$\sum_{\substack{2 \leq t \leq s \\ 1 \leq r_1,\ldots,r_t \leq s \\ 1 \leq r_1 < r_2 < \ldots < r_{t-1} < (r_t = s)}} \sum_{\substack{X_{j_2},\ldots,X_{j_t} \in G^{\Theta} \setminus \{X\} \\ \{j_2,\ldots,j_t\} \in \mathcal{P}^{t-1}(\{1,\ldots,n\}) \\ X \cap X_{j_2} \cap \ldots \cap X_{j_s} = \emptyset \\ \{i_1,\ldots,i_s\} \in \mathcal{P}^s(\{1,\ldots,s\})}}$$

$$\frac{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(X)^2) \cdot [\prod_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]}{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(X)) + [\sum_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]} \quad (5)$$

where i, j, k, r, s and t in (5) are integers. $m_{1,2,\ldots,s}(X)$ corresponds to the conjunctive consensus on X between s sources and where all denominators are different from zero. If a denominator is zero, that fraction is discarded; $\mathcal{P}^k(\{1,2,\ldots,n\})$ is the set of all subsets of k elements from $\{1,2,\ldots,n\}$ (permutations of n elements taken by k), the order of elements doesn't count.

The general formula of PCR6 proposed by Martin and Osswald for the combination of $s \geq 2$ sources is given by $m_{1,2,\ldots,s}^{PCR6}(\emptyset) = 0$ and for $X \neq \emptyset$ in G^{Θ}

$$m_{1,2,\ldots,s}^{PCR6}(X) = m_{1,2,\ldots,s}(X) +$$

$$\sum_{i=1}^{s} m_i(X)^2 \sum_{\substack{\overset{s-1}{\underset{k=1}{\cap}} Y_{\sigma_i(k)} \cap X \equiv \emptyset \\ (Y_{\sigma_i(1)},\ldots,Y_{\sigma_i(s-1)}) \in (G^{\Theta})^{s-1}}}$$

$$\left(\frac{\prod_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})}{m_i(X) + \sum_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})} \right) \quad (6)$$

where σ_i counts from 1 to s avoiding i:

$$\begin{cases} \sigma_i(j) = j & \text{if } j < i, \\ \sigma_i(j) = j+1 & \text{if } j \geq i, \end{cases} \quad (7)$$

Since Y_i is a focal element of expert/source i, $m_i(X) + \sum_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)}) \neq 0$.

The general PCR5 and PCR6 formulas (5)–(6) are rather complicate and not very easy to understand. From the implementation point of view, PCR6 is much simple to implement than PCR5. For convenience, very basic (not optimized) Matlab codes of PCR5 and PCR6 fusion rules can be found in [2], [10] and from the toolboxes repository on the web [11]. The PCR5 and PCR6 fusion rules are commutative but not associative, like the averaging fusion rule, but the vacuous belief assignment is a neutral element for these PCR fusion rules.

The PCR5 and PCR6 fusion rules simplify greatly and coincide for the combination of two sources ($s = 2$). In such simplest case, one always gets the resulting bba $m_{PCR5/6}(.) = m_{1,2}^{PCR6}(.) = m_{1,2}^{PCR5}(.)$ expressed as $m_{PCR5/6}(\emptyset) = 0$ and for all $X \neq \emptyset$ in G^{Θ}

$$m_{PCR5/6}(X) = \sum_{\substack{X_1,X_2 \in G^{\Theta} \\ X_1 \cap X_2 = X}} m_1(X_1)m_2(X_2) +$$

$$\sum_{\substack{Y \in G^{\Theta} \setminus \{X\} \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)} \right] \quad (8)$$

where all denominators in (8) are different from zero. If a denominator is zero, that fraction is discarded. All propositions/sets are in a canonical form.

Example 1: See [2], Vol.2, Chap. 1 for more examples.

Let's consider the frame of discernment $\Theta = \{A, B\}$ of exclusive elements. Here Shafer's model holds so that $G^{\Theta} = 2^{\Theta} = \{\emptyset, A, B, A \cup B\}$. We consider two sources of evidences providing the following bba's

$$m_1(A) = 0.6 \quad m_1(B) = 0.3 \quad m_1(A \cup B) = 0.1$$

$$m_2(A) = 0.2 \quad m_2(B) = 0.3 \quad m_2(A \cup B) = 0.5$$

Then the conjunctive consensus yields :

$$m_{12}(A) = 0.44 \quad m_{12}(B) = 0.27 \quad m_{12}(A \cup B) = 0.05$$

with the conflicting mass

$$m_{12}(A \cap B = \emptyset) = m_1(A)m_2(B) + m_1(B)m_2(A)$$
$$= 0.18 + 0.06 = 0.24$$

One sees that only A and B are involved in the derivation of the conflicting mass, but not $A \cup B$. With PCR5/6, one redistributes the partial conflicting mass 0.18 to A and B proportionally with the masses $m_1(A)$ and $m_2(B)$ assigned to A and B respectively, and also the partial conflicting mass 0.06 to A and B proportionally with the masses $m_2(A)$ and $m_1(B)$ assigned to A and B respectively, thus one gets two weighting factors of the redistribution for each corresponding set A and B respectively. Let x_1 be the conflicting mass to be redistributed to A, and y_1 the conflicting mass redistributed to B from the first partial conflicting mass 0.18. This first partial proportional redistribution is then done according

$$\frac{x_1}{0.6} = \frac{y_1}{0.3} = \frac{x_1 + y_1}{0.6 + 0.3} = \frac{0.18}{0.9} = 0.2$$

whence $x_1 = 0.6 \cdot 0.2 = 0.12$, $y_1 = 0.3 \cdot 0.2 = 0.06$. Now let x_2 be the conflicting mass to be redistributed to A, and y_2 the conflicting mass redistributed to B from the second the partial conflicting mass 0.06. This second partial proportional redistribution is then done according

$$\frac{x_2}{0.2} = \frac{y_2}{0.3} = \frac{x_2 + y_2}{0.2 + 0.3} = \frac{0.06}{0.5} = 0.12$$

whence $x_2 = 0.2 \cdot 0.12 = 0.024$, $y_2 = 0.3 \cdot 0.12 = 0.036$. Thus one finally gets:

$$m_{PCR5/6}(A) = 0.44 + 0.12 + 0.024 = 0.584$$
$$m_{PCR5/6}(B) = 0.27 + 0.06 + 0.036 = 0.366$$
$$m_{PCR5/6}(A \cup B) = 0.05 + 0 = 0.05$$

• The difference between PCR5 and PCR6 fusion rules

For the two sources case, PCR5 and PCR6 fusion rules coincide. As soon as three (or more) sources are involved in the fusion process, PCR5 and PCR6 differ in the way the proportional conflict redistribution is done. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, $m_3(B) = 0.1$.

– With PCR5, the mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ corresponding to a conflict is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization requires

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

that is

$$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

Thus

$$\begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

– With the PCR6 fusion rule, the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_B^{PCR6}}{m_2(B) + m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + (m_2(B) + m_3(B))}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_B^{PCR6}}{0.3 + 0.1} = \frac{0.018}{0.6 + (0.3 + 0.1)} = 0.018$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$\begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_B^{PCR6} = (0.3 + 0.1) \cdot 0.018 = 0.0072 \end{cases}$$

In [2], Vol. 2, Chap. 2, Martin and Osswald have proposed PCR6 based on intuitive considerations and the authors have shown through simulations that PCR6 is more stable than PCR5 in term of decision for combining $s > 2$ sources of evidence. Based on these results and the relative "simplicity" of implementation of PCR6 over PCR5, PCR6 has been considered more interesting/efficient than PCR5 for combining 3 (or more) sources of evidences.

III. CONSISTENCY OF PCR6 WITH THE AVERAGING RULE

In this section we show why we also consider PCR6 as better than PCR5 for combining bba's. But here, our argumentation is not based on particular simulation results and decision-making as done by Martin and Osswald, but on a theoretical analysis of the structure of PCR6 fusion rule itself. In particular, we show the full consistency of PCR6 rule with the averaging fusion rule used to empirically estimate probabilities in random experiments. For doing this, it is necessary to simplify the original PCR6 fusion formula (6). Such simplification has already been proposed in [12] and the PCR6 fusion rule can be in fact rewritten as

$$m_{1,2,\ldots,s}^{PCR6}(X) = m_{1,2,\ldots,s}(X) +$$
$$\sum_{k=1}^{s-1} \sum_{\substack{X_{i_1}, X_{i_2}, \ldots, X_{i_k} \in G^\Theta \setminus X \\ (\cap_{j=1}^k X_{i_j}) \cap X = \emptyset}} \sum_{(i_1, i_2, \ldots, i_k) \in \mathcal{P}^s(\{1, \ldots, s\})}$$
$$[m_{i_1}(X) + m_{i_2}(X) + \ldots + m_{i_k}(X)] \cdot$$
$$\cdot \frac{m_{i_1}(X) \ldots m_{i_k}(X) m_{i_{k+1}}(X_{i_{k+1}}) \ldots m_{i_s}(X_{i_s})}{m_{i_1}(X) + \ldots + m_{i_k}(X) + m_{i_{k+1}}(X_{i_{k+1}}) + \ldots + m_{i_s}(X_{i_s})}$$

(9)

where $\mathcal{P}^s(\{1, \ldots, s\})$ is the set of all permutations of the elements $\{1, 2, \ldots, s\}$. It should be observed that X_{i_1}, $X_{i_2}, \ldots, X_{i_s}$ may be different from each other, or some of them equal and others different, etc.

We wrote this PCR6 general formula (9) in the style of PCR5, different from Arnaud Martin & Christophe Oswald's notations, but actually doing the same thing. In order not to complicate the formula of PCR6, we did not use more summations or products after the third Sigma.

We now are able to establish the consistency of general PCR6 formula with the Averaging fusion rule for the case of binary bba's through the following theorem 1.

Theorem 1: When $s \geq 2$ sources of evidences provide binary bba's on G^Θ whose total conflicting mass is 1, then the PCR6 fusion rule coincides with the averaging fusion rule. Otherwise, PCR6 and the averaging fusion rule provide in general different results.

Proof 1: All $s \geq 2$ bba's are assumed binary, i.e. $m(X) = 0$ or 1 (two numerical values 0 and 1 only are allowed) for any bba $m(.)$ and for any set X in the focal elements. A focal element in this case is an element X such that at least one of the s binary sources assigns a mass equals to 1 to X. Let's suppose the focal elements are F_1, $F_2, \ldots, F_n$. Then the set of bba's to combine can be expressed as in the Table I. where

Table I.   LIST OF BBA'S TO COMBINE.

bba's \ Focal elem.	F_1	F_2	$\ldots$	F_n
$m_1(.)$	$\star$	$\star$	$\ldots$	$\star$
$m_2(.)$	$\star$	$\star$	$\ldots$	$\star$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m_s(.)$	$\star$	$\star$	$\ldots$	$\star$

• all $\star$ are 0's or 1's;

• on each row there is only a 1 (since the sum of all masses of a bba is equal to 1) and all the other elements are 0's;

- also each column has at least an 1 (since all elements are focals; and if there was a column corresponding for example to the set F_p having only 0's, then it would result that the set F_p is not focal, i.e. that all $m(F_p) = 0$).

Using PCR6, we first need to apply the conjunctive rule to all s sources, and the result is a sum of products of the form $m_1(X_1)m_2(X_2)\ldots m_s(X_s)$ where $X_1, X_2,\ldots,X_s$, are the focal elements $F_1, F_2,\ldots,F_n$ in various permutations, with $s \geq n$. If $s > n$ some focal elements X_i are repeated in the product $m_1(X_1)m_2(X_2)\ldots m_s(X_s)$. But there is only one product of the form $m_1(X_1)m_2(X_2)\ldots m_s(X_s) = 1$ which is not equal to zero, i.e. that product which has each factor equals to "1" (i.e. the product that collects from each row the existing single 1). Since the total conflicting mass is equal to 1, it means that this product represents the total conflict. In this case the PCR6 formula (9) becomes

$$m_{1,2,\ldots,s}^{PCR6}(X) = 0+$$
$$\sum_{k=1}^{s-1} \sum_{\substack{X_{i_1},X_{i_2},\ldots,X_{i_k}\in G^\Theta\setminus X \\ (\bigcap_{j=1}^k X_{i_j})\cap X=\emptyset}} \sum_{(i_1,i_2,\ldots,i_k)\in\mathcal{P}^s(\{1,\ldots,s\})}$$
$$[1+1+\ldots+1]\cdot\frac{1\cdot 1\cdot\ldots\cdot 1\cdot 1\cdot\ldots\cdot 1}{1+1+\ldots+1+1+\ldots+1} \quad (10)$$

The previous expression can be rewritten as

$$m_{1,2,\ldots,s}^{PCR6}(X) = \sum_{k=1}^{s-1} \sum_{\substack{X_{i_1},X_{i_2},\ldots,X_{i_k}\in G^\Theta\setminus X \\ (\bigcap_{j=1}^k X_{i_j})\cap X=\emptyset}} \sum_{\substack{(i_1,i_2,\ldots,i_k) \\ \in\mathcal{P}^s(\{1,\ldots,s\})}} k\cdot\frac{1}{s}$$

which is equal to k/s since there is only one possible non-null product of the form $m_1(X_1)m_2(X_2)\ldots m_s(X_s)$, and all other products are equal to zero. Therefore, we finally get:

$$m_{1,2,\ldots,s}^{PCR6}(X) = \frac{k}{s} \quad (11)$$

where "k" is the number of bba's $m(.)$ which give $m(X) = 1$. Therefore PCR6 in this case reduces to the average of masses, which completes the proof 1 of the theorem.

Proof 2: A second method of proving this theorem can also be done as follows. Let $m_1(.), m_2(.), \ldots, m_s(.)$, for $s \geq 3$, be bba's of the sources of information to combine and denote $\mathcal{F} = \{F_1, F_2,\ldots,F_n\}$, for $n \geq 2$, the set of all focal elements. All sources give only binary masses, i.e. $m_k(F_l) = 0$ or $m_k(F_l) = 1$ for any $k \in \{1, 2,\ldots,s\}$ and any $l \in \{1, 2,\ldots,n\}$. Since each $F_i, 1 \leq i \leq n$, is a focal element, there exists at least a bba $m_{i_o}(.)$ such that $m_{i_o}(F_i) = 1$, otherwise (i.e. if all sources gave the mass of F_i be equal to zero) F_i would not be focal. Without reducing the generality of the theorem, we can regroup the masses (since we combine all of them at once, so their order doesn't matter), as in Table II. Of course $i_1 + i_2 + \ldots + i_n = s$, since the s bba's are the same but reordered, and $i_1 \geq 1, i_2 \geq 1, \ldots,$ and $i_n \geq 1$. The total conflicting mass according to the theorem hypothesis $m_{1,2,\ldots,s}(\emptyset)$ is 1. With the PCR6 fusion rule we transfer the conflict mass back to focal elements $F_1, F_2,\ldots F_n$ respectively according to PCR

Table II.   LIST OF REORDERED BINARY BBA'S.

bba's \ Focal elem.	F_1	F_2	...	F_n	$\emptyset$
$m_{r_1}(.)$	1	0	...	0	0
$m_{r_2}(.)$	1	0	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{r_{i_1}}(.)$	1	0	...	0	0
$m_{s_1}(.)$	0	1	...	0	0
$m_{s_2}(.)$	0	1	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{s_{i_2}}(.)$	0	1	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{u_1}(.)$	0	0	...	1	0
$m_{u_2}(.)$	0	0	...	1	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{u_{i_n}}(.)$	0	0	...	1	0
$m_{1,2,\ldots,s}(.)$	0	0	...	0	1

principle such that:

$$\frac{x_{F_1}}{\underbrace{1+1+\ldots+1}_{i_1 \text{ times}}} = \frac{x_{F_2}}{\underbrace{1+1+\ldots+1}_{i_2 \text{ times}}} = \ldots$$
$$= \frac{x_{F_n}}{\underbrace{1+1+\ldots+1}_{i_n \text{ times}}} = \frac{m_{1,2,\ldots,s}(\emptyset)}{i_1+i_2+\ldots+i_n} = \frac{1}{s}$$

whence $x_{F_1} = i_1/s, x_{F_2} = i_2/s, \ldots, x_{F_n} = i_n/s$. Therefore $m_{1,2,\ldots,s}^{PCR6}(F_1) = i_1/s, m_{1,2,\ldots,s}^{PCR6}(F_2) = i_2/s, \ldots m_{1,2,\ldots,s}^{PCR6}(F_n) = i_n/s$. But averaging the masses $m_1(.), m_2(.), \ldots, m_s(.)$ is equivalent to averaging each column of $F_1, F_2, \ldots F_n$. Hence average of column F_1 is i_1/s, average of column F_2 is $i_2/s, \ldots,$ average of column F_n is i_n/s. Therefore, in case of binary bba's which are globally totally conflicting, PCR6 rule is equal to the Averaging Rule. This completes the proof 2 of the theorem.

Note that using PCR5 fusion rule, we also transfer the total conflicting mass that is equal to 1 to $F_1, F_2, \ldots, F_n$ respectively, but we replace the addition "+" with the multiplication "." in the above proportionalizations:

$$\frac{x_{F_1}}{\underbrace{1\cdot 1\cdot\ldots\cdot 1}_{i_1 \text{ times}}} = \frac{x_{F_2}}{\underbrace{1\cdot 1\cdot\ldots\cdot 1}_{i_2 \text{ times}}} = \ldots = \frac{x_{F_n}}{\underbrace{1\cdot 1\cdot\ldots\cdot 1}_{i_n \text{ times}}} = \frac{m_{1,2,\ldots,s}(\emptyset)}{\underbrace{1+1+\ldots+1}_{n \text{ times}}} = \frac{1}{n}$$

so that $x_{F_1} = 1/n, x_{F_2} = 1/n, \ldots, x_{F_n} = 1/n$ and therefore

$$m_{1,2,\ldots,s}^{PCR5}(F_1) = m_{1,2,\ldots,s}^{PCR5}(F_2) = \ldots = m_{1,2,\ldots,s}^{PCR5}(F_n) = 1/n$$

Corollary 1: When $s \geq 2$ sources of evidences provide binary bba's on G^Θ with at least two focal elements, and all focal elements are disjoint two by two, then PCR6 fusion rule coincides with the Averaging Rule.

This Corollary is true because if all focal elements are disjoint two by two then the total conflict is equal to 1.

Examples 2: where PCR6 rule equals the Averaging Rule.

Let's consider the frame $\Theta = \{A, B\}$ with Shafer's model and the bba's to combine as given in Table III.

Table III. LIST OF BBA'S TO COMBINE FOR EXAMPLE 2.

bba's \ Focal elem.	A	B	$A \cup B$	$A \cap B = \emptyset$
$m_1(.)$	1	0	0	
$m_2(.)$	0	1	0	
$m_3(.)$	0	0	1	
$m_{1,2,3}(.)$	0	0	0	1

Since we have binary masses, and their total conflict is 1, we expect getting the same result for PCR6 and the Averaging Rule according to our Theorem 1. The PCR principle gives us

$$\frac{x_A}{1} = \frac{y_B}{1} = \frac{z_{A \cup B}}{1} = \frac{m_{1,2,3}(\emptyset)}{1+1+1} = \frac{1}{3}$$

Hence $x_A = y_B = z_{A \cup B} = \frac{1}{3}$, so that

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}(A) + x_A = 0 + \frac{1}{3} = \frac{1}{3}$$
$$m_{1,2,3}^{PCR6}(B) = m_{1,2,3}(B) + y_B = 0 + \frac{1}{3} = \frac{1}{3}$$
$$m_{1,2,3}^{PCR6}(A \cup B) = m_{1,2,3}(A \cup B) + z_{A \cup B} = 0 + \frac{1}{3} = \frac{1}{3}$$

Interestingly, PCR5 gives the same result as PCR6 in this case since one makes the same proportionalizations as for PCR6. Using the Averaging Rule (2), we get

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (1 + 0 + 0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(B) = \frac{1}{3} \cdot (0 + 1 + 0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B) = \frac{1}{3} \cdot (0 + 0 + 1) = \frac{1}{3}$$

So we see that PCR6 rule equals the Averaging Rule as proved in the theorem because the bba's are binary and the intersection of all focal elements is empty since $A \cap B \cap (A \cup B) = \emptyset \cap (A \cup B) = \emptyset$ because $A \cap B = \emptyset$ since Shafer's model has been assumed for the frame Θ.

Examples 3: where PCR6 differs from the Averaging Rule.

Let's consider the frame $\Theta = \{A, B, C\}$ with Shafer's model and the bba's to combine as given in Table IV.

Table IV. LIST OF BBA'S TO COMBINE FOR EXAMPLE 3.

bba's \ Focal elem.	A	$A \cup B$	$A \cup B \cup C$	$\emptyset$
$m_1(.)$	1	0	0	
$m_2(.)$	0	1	0	
$m_3(.)$	0	0	1	
$m_{1,2,3}(.)$	1	0	0	

Clearly, in this case the focal elements are nested and the condition on emptiness of intersection of all focal elements is not satisfied because one has $A \cap (A \cup B) \cap (A \cup B \cup C) = A \neq \emptyset$, so that the theorem cannot be applied in such case. The total conflicting mass is not 1. One can verify in such example that PCR6 rule differs from the Averaging Rule because one gets

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}(A) = 1$$
$$m_{1,2,3}^{PCR6}(A \cup B) = m_{1,2,3}(A \cup B) = 0$$
$$m_{1,2,3}^{PCR6}(A \cup B \cup C) = m_{1,2,3}(A \cup B \cup C) = 0$$

since there is no conflicting mass to redistribute to apply PCR principle, whereas the averaging fusion rule gives

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (1 + 0 + 0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B) = \frac{1}{3} \cdot (0 + 1 + 0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B \cup C) = \frac{1}{3} \cdot (0 + 0 + 1) = \frac{1}{3}$$

Examples 4 (Bayesian non-binary bba's): where PCR6 differs from the Averaging Rule.

Let's consider the frame $\Theta = \{A, B\}$ with Shafer's model and the Bayesian bba's to combine as given in Table V.

Table V. LIST OF BBA'S TO COMBINE FOR EXAMPLE 4.

bba's \ Focal elem.	A	B	$A \cap B = \emptyset$
$m_1(.)$	0.2	0.8	0
$m_2(.)$	0.6	0.4	0
$m_3(.)$	0.7	0.3	0
$m_{1,2,3}(.)$	0.084	0.096	0.820

The total conflicting mass $m_{1,2,3}(A \cap B = \emptyset) = 0.82 = 1 - m_1(A)m_2(A)m_3(A) - m_1(B)m_2(B)m_3(B)$ equals the sum of partial conflicting masses that will be redistributed through PCR principle in PCR6

$$m_{1,2,3}(A \cap B = \emptyset) = \underbrace{m_1(A)m_2(B)m_3(B)}_{0.024}$$
$$+ \underbrace{m_2(A)m_1(B)m_3(B)}_{0.144} + \underbrace{m_3(A)m_1(B)m_2(B)}_{0.224}$$
$$+ \underbrace{m_1(B)m_2(A)m_3(A)}_{0.336} + \underbrace{m_2(B)m_1(A)m_3(A)}_{0.056}$$
$$+ \underbrace{m_3(B)m_1(A)m_2(A)}_{0.036} = 0.82$$

Applying PCR principle for each of these six partial conflicts, one gets:

- for $m_1(A)m_2(B)m_3(B) = 0.2 \cdot 0.4 \cdot 0.3 = 0.024$
$$\frac{x_1(A)}{0.2} = \frac{y_1(B)}{0.4 + 0.3} = \frac{0.024}{0.2 + 0.3 + 0.4}$$
whence $x_1(A) \approx 0.005333$ and $y_1(B) \approx 0.018667$.

- for $m_2(A)m_1(B)m_3(B) = 0.6 \cdot 0.8 \cdot 0.3 = 0.144$
$$\frac{x_2(A)}{0.6} = \frac{y_2(B)}{0.8 + 0.3} = \frac{0.144}{0.6 + 0.8 + 0.3}$$
whence $x_2(A) \approx 0.050824$ and $y_2(B) \approx 0.093176$.

- for $m_3(A)m_1(B)m_2(B) = 0.7 \cdot 0.8 \cdot 0.4 = 0.224$
$$\frac{x_3(A)}{0.7} = \frac{y_3(B)}{0.8 + 0.4} = \frac{0.224}{0.7 + 0.8 + 0.4}$$
whence $x_3(A) \approx 0.082526$ and $y_3(B) \approx 0.141474$.

- for $m_1(B)m_2(A)m_3(A) = 0.8 \cdot 0.6 \cdot 0.7 = 0.336$
$$\frac{x_4(A)}{0.6 + 0.7} = \frac{y_4(B)}{0.8} = \frac{0.336}{0.8 + 0.6 + 0.7}$$
whence $x_4(A) \approx 0.208000$ and $y_4(B) \approx 0.128000$.

- for $m_2(B)m_1(A)m_3(A) = 0.4 \cdot 0.2 \cdot 0.7 = 0.056$

$$\frac{x_5(A)}{0.2+0.7} = \frac{y_5(B)}{0.4} = \frac{0.056}{0.4+0.2+0.7}$$

whence $x_5(A) \approx 0.038769$ and $y_5(B) \approx 0.017231$.

- for $m_3(B)m_1(A)m_2(A) = 0.3 \cdot 0.2 \cdot 0.6 = 0.036$

$$\frac{x_6(A)}{0.2+0.6} = \frac{y_6(B)}{0.3} = \frac{0.036}{0.3+0.2+0.6}$$

whence $x_6(A) \approx 0.026182$ and $y_6(B) \approx 0.009818$.

Therefore, with PCR6 one finally gets

$$m_{1,2,3}^{PCR6}(A) = 0.084 + \sum_{i=1}^{6} x_i(A) = 0.495634$$

$$m_{1,2,3}^{PCR6}(B) = 0.096 + \sum_{i=1}^{6} y_i(A) = 0.504366$$

whereas the Averaging Rule (2) will give us

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (0.2+0.6+0.7) = \frac{1.5}{3} = 0.5$$
$$m_{1,2,3}^{Average}(B) = \frac{1}{3} \cdot (0.8+0.4+0.3) = \frac{1.5}{3} = 0.5$$

In this example, the intersection of focal elements is empty but the bba's to combine are not binary. Therefore the total conflict between sources is not total and the theorem doesn't apply and so PCR6 results differ from the Averaging Rule.

It however can happen that in some very particular symmetric cases PCR6 coincides with the Averaging Rule. For example, if we consider the bba's as given in the Table VI. In such case the opinion of source #1 totally balances opinion of source #3, and the opinion of source #2 cannot support A more than B (and reciprocally), so that the fusion problem is totally symmetrical. In this example, it is expected that the final fusion result should commit an equal mass of belief to A and to B. And indeed, it can be easily verified that one gets in such case

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}^{Average}(A) = 0.5$$
$$m_{1,2,3}^{PCR6}(B) = m_{1,2,3}^{Average}(B) = 0.5$$

which makes perfectly sense. Note that the Averaging Rule provides same result on example 4 which is somehow questionable because example 4 doesn't present an inherent symmetrical structure. In our opinion PCR6 presents the advantage to respond more adequately to the change of inherent internal structure (asymmetry) of bba's to combine, which is not well captured by the simple averaging fusion rule.

Table VI.   A BAYESIAN NON-BINARY SYMMETRIC EXAMPLE.

bba's \ Focal elem.	A	B	$A \cap B = \emptyset$
$m_1(.)$	0.2	0.8	0
$m_2(.)$	0.5	0.5	0
$m_3(.)$	0.8	0.2	0
$m_{1,2,3}(.)$	0.08	0.08	0.84

IV.   APPLICATION TO PROBABILITY ESTIMATION

Let's review a simple coin tossing random experiment. When we flip a coin [13], there are two possible outcomes. The coin could land showing a head (H) or a tail (T). The list of all possible outcomes is called the sample space and correspond to the frame $\Theta = \{H, T\}$. There exist many interpretations of probability [14] that are out of the scope of this paper. We focus here on the estimation of the probability measure $P(H)$ of a given coin (biased or not) based on n outcomes of a coin tossing experiment. The long-run frequentist interpretation of probability [15] considers that the probability of an event A is its relative frequency of occurrence over time after repeating the experiment a large number of times under similar circumstances, that is

$$P(A) = \lim_{n \to \infty} \frac{n(A)}{n} \qquad (12)$$

where $n(A)$ denotes the number of occurrences of an event A in $n > 0$ trials. In practice however, we usually estimate the probability of an event A based only on a limited number of data (observations) that are available, and so we estimate the idealistic $P(A)$ defined in (12), by classical Laplace's probability definition

$$\hat{P}(A|n(A), n) = \frac{n(A)}{n} \qquad (13)$$

Naturally, $\hat{P}(A) \geq 0$ because $n(A) \geq 0$ and $n > 0$, and $\hat{P}(A) \leq 1$ because we cannot get $n(A) > n$ in a series of n trials. $P(A) + P(\bar{A}) = 1$ because $\frac{n(A)}{n} + \frac{n(\bar{A})}{n} = \frac{n(A)}{n} + \frac{n-n(A)}{n} = 1$ where $\bar{A}$ is the complement of A in the sample space.

It is interesting to note that the classical estimation of the probability measure given by (13) corresponds in fact to the simple averaging fusion rule of distinct pieces of evidence represented by binary masses. For example, let's take a coin and flip it $n = 8$ times and assume for instance that we observe the following series of outcomes $\{o_1 = H, o_2 = H, o_3 = T, o_4 = H, o_5 = T, o_6 = H, o_7 = H, o_8 = T\}$, so that $n(H) = 5$ and $n(T) = 3$. Then these observations can be associated with distinct sources of evidences providing to the following basic (binary) belief assignments:

Table VII.   OUTCOMES OF A COIN TOSSING EXPERIMENT.

bba's \ Focal elem.	H	T
$m_1(.)$	1	0
$m_2(.)$	1	0
$m_3(.)$	0	1
$m_4(.)$	1	0
$m_5(.)$	0	1
$m_6(.)$	1	0
$m_7(.)$	1	0
$m_8(.)$	0	1

It is clear that the probability estimate in (13) equals the averaging fusion rule (2) and in such example because

$$\hat{P}(H|\{o_1, o_2, \ldots, o_8\}) = \frac{n(H)}{n} = \frac{5}{8} \qquad \text{by eq. (13)}$$
$$= \frac{1}{8}(1+1+0+1+0+1+1+0)$$
$$= m_{1,2,\ldots,8}^{Average}(H) \qquad \text{by eq. (2)}$$

$$\hat{P}(T|\{o_1, o_2, \ldots, o_8\}) = \frac{n(T)}{n} = \frac{3}{8} \qquad \text{by eq. (13)}$$
$$= \frac{1}{8}(0 + 0 + 1 + 0 + 1 + 0 + 0 + 1)$$
$$= m_{1,2,\ldots,8}^{Average}(T) \qquad \text{by eq. (2)}$$

Because all the bba's to combine here are binary and are in total conflict, our theorem 1 of Section III applies, and PCR6 fusion rule in this case coincides with the averaging fusion rule. Therefore we can use PCR6 to estimate the probabilities that the coin will land on H or T at the next toss given the series of observations. More precisely,

$$\begin{cases} m_{1,2,\ldots,8}^{PCR6}(H) = m_{1,2,\ldots,8}^{Average}(H) = \hat{P}(H|\{o_1, o_2, \ldots, o_8\}) \\ m_{1,2,\ldots,8}^{PCR6}(T) = m_{1,2,\ldots,8}^{Average}(T) = \hat{P}(T|\{o_1, o_2, \ldots, o_8\}) \end{cases}$$

We must insist on the fact that Dempster-Shafer (DS) rule of combination (4) cannot be used at all in such very simple case to estimate correctly the probability measure because DS rule doesn't work (because of division by zero) in total conflicting situations. PCR5 rule can be applied to combine these 8 bba's but is unable to provide a consistent result with the classical probability estimates because one will get

$$\frac{x_H}{1 \cdot 1 \cdot 1 \cdot 1 \cdot 1} = \frac{y_T}{1 \cdot 1 \cdot 1} = \frac{m_{1,2,\ldots,8}(\emptyset)}{(1 \cdot 1 \cdot 1 \cdot 1 \cdot 1) + (1 \cdot 1 \cdot 1)} = \frac{1}{1+1} = 0.5$$

and therefore the PCR5 fusion result is

$$\begin{cases} m_{1,2,\ldots,8}^{PCR5}(H) = x_H = 0.5 \neq (m_{1,2,\ldots,8}^{PCR6}(H) = 5/8) \\ m_{1,2,\ldots,8}^{PCR5}(T) = y_T = 0.5 \neq (m_{1,2,\ldots,8}^{PCR6}(T) = 3/8) \end{cases}$$

Remark: The PCR6 fusion result is valid if and only if PCR6 rule is applied globally, and not sequentially. If PCR6 is sequentially applied, it becomes equivalent with PCR5 sequentially applied and it will generate incorrect results for combining $s > 2$ sources of evidence. Because of the ability of PCR6 to estimate frequentist probabilities in a random experiment, we strongly recommend PCR6 rather than PCR5 as soon as $s \geq 2$ bba's have to be combined altogether.

V. CONCLUSIONS AND CHALLENGE

In this paper, we have proved that PCR6 fusion rule coincides with the Averaging Rule when the bba's to combine are binary and in total conflict. Because of such nice property, PCR6 is able to provide a frequentist probability measure of any event occurring in a random experiment, contrariwise to other fusion rules like DS rule, PCR5 rule, etc. Except the Averaging Rule of course since it is the basis of the frequentist probability interpretation. In a more general context with non-binary bba's, PCR6 is quite complicate to apply to combine globally $s > 2$ sources of evidences, and a general recursive formula of PCR6 would be very convenient. It can be mathematically reformulated as follows: Let R be a fusion rule and assume we have s sources that provide $m_1, m_2, \ldots, m_{s-1}, m_s$ respectively on a fusion space G^Θ. Find a function (or an operator) T such that: $T(R(m_1, m_2, \ldots m_{s-1}), m_s) = R(m_1, m_2, \ldots, m_{s-1}, m_s)$, or by simplifying the notations $T(R_{s-1}, m_s) = R_s$, where R_i means the fusion rule R applied to i masses all together. For example, if R equals the Averaging Rule, the function T is defined according to the relation (3) by $T(R_{s-1}, m_s) = \frac{s-1}{s} R_{s-1} + \frac{1}{s} m_s = R_s$, and if R equals

DS rule one has $T(R_{s-1}, m_s) = DS(R_{s-1}, m_s)$ because of the associativity of DS rule. What is the T operator associated with PCR6? Such very important open challenging question is left for future research works.

REFERENCES

[1] G. Shafer, A mathematical theory of evidence, Princeton Univ. Press, Princeton, NJ, U.S.A., 1976.

[2] F. Smarandache, J. Dezert (Editors), Advances and applications of DSmT for information fusion, American Research Press, Rehoboth, NM, U.S.A., Vol. 1–3, 2004–2009. http://fs.gallup.unm.edu//DSmT.htm

[3] X.R. Li, Probability, random signals and statistics, CRC Press, 1999.

[4] P.E. Pfeiffer, Applied probability, Connexions Web site. http://cnx.org/content/col10708/1.6/

[5] L.A. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, CA, U.S.A., 1979.

[6] L.A. Zadeh, Book review: A mathematical theory of evidence, The AI Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[7] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The AI Magazine, Vol. 7, No. 2, 1986.

[8] J. Dezert, P. Wang, A. Tchamova, On the validity of Dempster-Shafer theory, Proc. of Fusion 2012 Int. Conf., Singapore, July 9–12, 2012.

[9] A. Tchamova, J. Dezert, On the behavior of Dempster's Rule of combination and the foundations of Dempster-Shafer theory, (Best paper awards), 6th IEEE Int. Conf. on Int. Syst. (IS '12), Sofia, Bulgaria, Sept. 6–8, 2012.

[10] F. Smarandache, J. Dezert, J.-M. Tacnet, Fusion of sources of evidence with different importances and reliabilities, Proc. of Fusion 2010 Int. Conf., Edinburgh, UK, July 26–29, 2010.

[11] http://bfas.iutlan.univ-rennes1.fr/wiki/index.php/Toolboxs

[12] F. Smarandache, J. Dezert, Importance of sources using the repeated fusion method and the proportional conflict redistribution rules #5 and #6, in Multispace & Multistructure. Neutrosophic transdisciplinarity (100 collected papers of sciences), Vol. IV, pp. 349–354, North-European Scientific Publishers Hanko, Finland, 2010.

[13] http://www.random.org/coins/ or http://shazam.econ.ubc.ca/flip/

[14] Free online Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/probability-interpret/

[15] R. Von Mises, Probability, statistics and the truth, Dover, New York, Second revised English Edition, 1957.