A new distance between bodies of evidence based on Dempsterian specialisation matrices

Mehena Loudahi, John KLlein, Olivier Colot, Jean-Marc Vannobel

LAGIS - UMR CNRS 8219, Lille1 University, France

Abstract

Keywords:

1. Introduction

The evidence theory is a formal frame of reasoning with uncertainty. It can efficiently represent uncertain or/and imprecise pieces of information using some mathematical objects called belief functions. The evidence theory encompassess both the probabilisty and fuzzy set theories. It was initiated in 1967 by A. P. Dempster [] and further developed by G. Shaffer in 1976 in his seminal work [1]. The evidence theory is thus also known as the Dempster-Shafer theory or the belief function theory.
Information fusion as part of this framework has been widely used to provide more relevant solutions to decision making problems. Each state of belief of a source is represented by a belief function and source aggregation is obtained by applying a combination rule to their corresponding belief functions. The orthogonal sum of Dempster combines belief functions so as to retain common evidence while discarding conflicting evidence. It is a generalization of the intersection in Kantor's set theory. In this rule, conflicting evidence is actually supressed using a normalization factor relying on the so-called Dempster conflict degree. This degree is the amount of support given to conflicting hypotheses. It is also the first measure introduced to compute, in some way, the difference between some belief functions. This degree is unfortunately not fully appropriate to assess the dissimilarity between belief functions, since it can provide a non-null value when comparing two identical information sources.
In the past decades, the belief function communinity has shown a growing interest in determining meaningful dissimilarity measures between bodies of evidence. The need for such measures is explained by the fact that it would help for instance to compute approximations of belief functions [], to cluster belief functions [] or to estimate some parameters feeding refined combination rules []. All (dis)similarity measures attempt to describe the degree of (non-)alikeness between belief functions in a meaningful way for the widest range of applications. Indeed, the choice of a particular measure is most of the time application-dependent. Up to now, there is no formally well established measure assessing belief function difference to the user's satisfaction in all application fields. This is mainly explained by the fact that each application requires some properties and so far no similarity measure possess all of them. In fact, seeking a measure possessing all possible properties is likely to be futile as a first application field might well require property that is incompatible with the property required in a second application field.
As a conseqeuence, a plethora of similarity measures is found in the literature. Among the most widely known measures, the betting commitment distance was proposed by Tessem et al. [2] to evaluate the approximation of basic belief assignments. It is applied to the pignistic transform of the processed belief functions. The pignistic transform turns a belief function into the least committed probability distribution. The Tessem dissimilarity is thus calculated in a probability space. To solve the same problem, Bauer [3] introduced, in 1997, two other measures based on pignistic probabilities. In 2006, Tessem's distance is jointly used with Dempster's conflict in W. Liu's work [4] to define the a

Email address: mehena.loudahi@ed.univ-lille1.fr, john.klein@univ-lille1.fr, olivier.colot@univ-lille1.fr, jean-marc.vannobel@univ-lille1.fr (Mehena Loudahi, John KLlein, Olivier Colot, Jean-Marc Vannobel)

new conflict measure between basic belief assignments and investigate the consistency of applying Dempster's combination rule in particular cases. Zouhal and Denœux [5] also proposed a mean square distance based on the pignistic transform to compare between a belief function and an indicator vector. It is used to improve the performance of a *k-NN* classifier algorithm and Dempster's rule of combination. More recently, Petit-Renaud [] proposes a measure in the power set of belief functions based on the Hausdorff distance. Fixen and Mahler [6] proposed a "classification miss-distance" between belief functions based on the Bayesian *a priori* distribution matrix. Jousseme's distance [7] is based on the euclidean metric provided from the geometric interpretation of evidence theory. Jousselme used the Jaccard coefficient as a similarity measure between focal elements. Diaz et al. [8] proposed also a distance based on a similarity measure between focal elements. The properties of this measure are different from other approaches as it mainly investigates the proximity of the processed belief functions with the vacuous belief function, which stands for a state of ignorance. In 2011, D. Han et al. [9] introduced four dissimilarity measures in the evidence theory based on some lossy transformation of belief functions into fuzzy measures. Khatibi and Montazer [10] defined a distance based on belief intervals. Their approach relies on an approximate reasoning using the total degree of belief while considering problem uncertainty. A thorough survey about dissimilarity measures in the belief functions theory and their properties was presented in 2012 by Jousselme and Maupin [11]. The authors also provided generalizations of some distances thereby introducing families of new measures.

In this article, we also intend to explore another way of designing a distance between bodies of evidence. All previously mentionned distances are computed from objects called basic belief assignments (bba). Instead, a distance computed from the Dempsterian specialization matrices is proposed. Each bba defines a unique specialization matrix and conversely. This choice is justified by the fact a specialization matrix not only decribes the present state of belief but also the potential future states that could be reached if additionnal pieces of evidence were combined. A classical matrix distance is the Froebinius distance, which is simply the euclidian distance of matrices reshaped as column vectors. This metric appears to have properties that can be easily understood in terms of belief functions. In addition, its behavior has no equivalent in previously introduced distances.

The sequel of this article is organized as follows: section 2 recalls some fundamentals of evidence theory with a particular stress on matrix calculus as part of this framework. Section 3 gives a short review and definitions of previously introduced belief function distances. Our new matrix specialisation distance is also introduced along with its properties in this section. In section 4, a comparison of the newly introduced distance with existing ones is provided through worked out examples and property interpretations.

2. Belief functions fundamentals and matrix calculus

2.1. Basics of the evidence theory

In the Dempster-Shafer theory [1], an evidence model is defined in a frame of discernment Ω which contains a finite number of exclusive and exhaustive hypothesis. The knowledge is expressed by using the basic belief assignment attributed to each hypothesis of the power set 2^Ω containing all the possible disjunctions of the frame of discernment Ω. Then, a belief state is presented by a mapping m defined in $2^\Omega \longrightarrow [0, 1]$ satisfying the normality condition:

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{1}$$

The same information can already be expressed by other functions such as the belief function (*Bel*), representing the minimal degree of belief in an hypothesis, and the plausibility function which is used to express the incapacity degree of belief in the opposite of the hypothesis $A \subseteq \Omega$. These two functions are respectively represented by:

$$Bel(A) = \sum_{(B \subseteq A, B \neq \emptyset)} m(B) \tag{2}$$

$$Pl(A) = \sum_{(B \cap A \neq \emptyset,\ B \in 2^\Omega)} m(B) \tag{3}$$

The communality (q), also the implicability (b), are other forms to express the information treated in the framework of belief functions. They are very useful for calculation issues. They can respectively be calculated from the basic belief assignment by using the following one-to-one correspondences:

$$q(A) = \sum_{\substack{B \supseteq A \\ B \in 2^{\Omega}}} m(B) \tag{4}$$

$$b(A) = \sum_{B \subseteq A} m(B) = Bel(A) + m(\emptyset) \tag{5}$$

A subset A of Ω is called focal element of the function m if the basic belief assignment (bba) attributed to this subset is strictly positive. The set of all focal elements is defined as the *Kernel* of the mass function m.

The decision making in the belief function theory is principally based on the pignistic transformation of the bba's to probability distributions. This transformation is defined by P. Smets in the frame of the Transferable Belief Model. Each bba is so transformed into pignistic probability by using the formula:

$$BetP(\omega) = \sum_{A \subseteq \Omega, \omega \in A} \frac{1}{|A|} . \frac{m(A)}{1 - m(\emptyset)} \tag{6}$$

where $|A|$ is the cardinality of the subset A.
BetP can also be extended to all the elements of the power set 2^{Ω} by using:

$$BetP(A) = \sum_{\omega_i \in A} BetP(\omega_i) \tag{7}$$

The mainly used operator when fusing the belief functions is the Dempster rule of combination. This choice is justified since it respects the less commitment principle and accepts the total ignorance as neutral element. It is also commutative ans associative. This rule is seen as the normalized version of the conjunctive rule of combination used by P. Smets [12] in the transferable belief model. The simplest expression of the conjunctive rule using communality measures is given by:

$$q_{1 \textcircled{\cap} 2}(A) = q_1(A).q_2(A), \quad \forall A \subseteq \Omega \tag{8}$$

Given two bba's m_1 et m_2, the conjunctive combination rule can also be defined as follow:

$$m_{1 \textcircled{\cap} 2}(A) = \sum_{B \cap C = A} m_1(B).m_2(C) \tag{9}$$

The concept of Dempsterain specialisation is a particular case of the specialisation concept defined by Yager in 1986 in the belief functions theory. According to him, a basic belief assignment m is a specialization of another basic belief assignment m_1, if it is, at least, as committed as m_1. This situation is then denoted by $m \sqsubseteq m_1$. The matrix representation of the Dempstrian specialization is introduced into the belief functions theory in 1992 by F. Klawonn and P. Smets [13]. It was then used in 2001 in the works of P.A. Monney [14].

To simplify the calculation issues laid to the complexity of the different algorithms developed in the belief functions theory, P. Smets [12] introduced a matrix framework which is very useful because it makes more easier these algorithmic calculus. This framework is reviewed in the sequel.

2.2. *Basic belief assignments and matrix calculus*

In 2001, A.L. Jousselme used a geometrical representation in her works about dissimilarity in the belief functions theory. This representation considers the basic belief assignments as vectors defined in vector space spanned by the different elements of the power set 2^{Ω}. The basis vectors of this space are organized following an order called *binary* or *natural*. This representation was after developed in the works of F. Cuzzolin [15, 16].

The matrix notation is very useful in the belief function theory algorithms. It allows jointly to better establish the linearity relationships between the basic operators of belief functions and to widely simplify the the algorithmic calculus of the belief functions operations. This well established notation is introduced in 2002 by P. Smets [12]. He proposed then a framework to apply the matrix calculus into belief functions. These works were after used by F. Pichon [17].

2.2.1. The natural order of basic belief assignments

For algorithmic aims, a basic belief assignment m (also the four other related forms of information representation bel, pl, b and q) defined in a frame of discernment Ω is conventionally seen as a stochastic column vector about $2^{|\Omega|}$ elements. The elements of this vector can, in fact, be organized in any arbitrary way. But a certain particular order renders the the algorithmic calculations more accessible and easily reachable. This order is called **natural order**.

In an ordered frame of discernment of cardinality $N = |\Omega|$, one can attribute a binary representation defined over N digits to each subset $A \subseteq \Omega$. The hypothesis $\omega_i \in A$ are so encoded by the binary value 1 in the so related binary representation of A. In fact, the i^{th} element of the vector $\mathbf{m}$ corresponds to the subset A of Ω. This last is encoded by the binary representation of the value $i-1$ defined over N digits. Suppose, for instance, a frame of discernment $\Omega = \{a, b, c\}$. The table 1 summarizes the order and the binary representation of the elements of the mass vector $\mathbf{m}$. For example, the eighth element corresponds to the subset $\Omega = \{a, b, c\}$ because of the binary representation of the number $8 - 1 = 7$ is 111. Also, the binary representation 100 concerns the subset $\{c\} \subset \Omega$ which is the fifth element in the binary order sins $4 + 1 = 5$.

Table 1: Binary order of the mass vector $\mathbf{m}$ while $\Omega = \{a, b, c\}$

Position	$c\ b\ a$	Ω	m
1	000	$\emptyset$	$m(\emptyset)$
2	001	$\{a\}$	$m(\{a\})$
3	010	$\{b\}$	$m(\{b\})$
4	011	$\{a, b\}$	$m(\{a, b\})$
5	100	$\{c\}$	$m(\{c\})$
6	101	$\{a, c\}$	$m(\{a, c\})$
7	110	$\{b, c\}$	$m(\{b, c\})$
8	111	$\{a, b, c\}$	$m(\{a, b, c\})$

For the sequel, the following notations are used :

- The vectors are column vectors and the matrices are square. Their length is equal to $2^{|\Omega|}$.
- The vectors and matrices are written in bold notation. A matrix can be represented by $\mathbf{M} = [M_{ij}]$, also by the notation $\mathbf{M} = [M(A, B)],\ \ \forall A, B \in \Omega$. The row and column indexes i and j are those corresponding to the subsets B_i and B_j of Ω ordered following the binary order. For instance, if $B_i = \{a\}$ and $B_j = \{a, c\}$ in $\Omega = \{a, b, c\}$, then the indexes are 2 and 6 respectively.
- $\mathbf{I}$ is the identity matrix, its components are null except those of the principal diagonal which are equal to 1.
- $\mathbf{J}$ is the square matrix which all components are null except those of the secondary diagonal that are equal to 1. The principal property of the matrix $\mathbf{J}$ is its ability to invert the rows of some matrix $\mathbf{M}$ when the product $\mathbf{J}.\mathbf{M}$ is computed. It inverts also the order of the columns when the product is $\mathbf{M}.\mathbf{J}$ is realized.

Definition 1. *The concept of negation of a basic belief assignment m in the frame of discernment Ω is defined by Dubois and Prad [12] as :*

$$\bar{m}(A) = m(\bar{A}) \tag{10}$$

where $\bar{m}$ is the negation of m and the subset $\bar{A}$ is the complementary of $A \subseteq \Omega$ in Ω.

Then, one can easily obtain, in this case, the communality of the basic belief assignment $\bar{m}$ as follow:

$$\begin{array}{rcl} \bar{q}(A) & = & \displaystyle\sum_{B \supseteq A} \bar{m}(B) \\ \text{but we know that :} & & if\ \ B \supseteq A \Leftrightarrow \bar{B} \subseteq \bar{A} \\ \Rightarrow \bar{q}(A) & = & \displaystyle\sum_{\bar{B} \subseteq \bar{A}} m(\bar{B}) \\ \bar{q}(A) & = & b(\bar{A}) \end{array}$$

In the same way, one can establish:

$$\bar{b}(A) = q(\bar{A})$$

The previous results are rewritten in the matrix representation as follow:

$$\bar{\mathbf{m}} = \mathbf{J}.\mathbf{m}, \quad \bar{\mathbf{q}} = \mathbf{J}.\mathbf{b}, \quad \bar{\mathbf{b}} = \mathbf{J}.\mathbf{q}$$

2.2.2. The Möbius transform

All the possible transformations between the different representations of evidence in the Dempster-Shafer theory can be written using the matrix notations. For example, given a basic belief assignment m, its transformation into communality function is given by the equation:

$$q(A) = \sum_{\Omega \supseteq B \supseteq A} m(B)$$

This latter can be expressed by:

$$q(A) = \sum_{\Omega \supseteq B} M(A, B).m(B)$$

while $M(A, B) = \begin{cases} 1 & if \;\; A \subseteq B \\ 0 & otherwise. \end{cases}$

The elements $M(A, B)$ are the components of the matrix $\mathbf{M}$ called incidence matrix.

Definition 2. *Given an ordered frame of discernment $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, the incidence matrix $\mathbf{M}$ is square and its elements are given by:*

$$M_{ij} = \begin{cases} 1 & if \;\; B_i \subseteq B_j \\ 0 & otherwise. \end{cases} \tag{11}$$

- *The matrix $\mathbf{M}$ is upper triangular. All its diagonal elements are one.*
- *The matrix relationships are subsequently provided:* $\mathbf{q} = \mathbf{M}.\mathbf{m} \Rightarrow \mathbf{m} = \mathbf{M}^{-1}.\mathbf{q}$
- *The communality matrix is diagonal and defined by $\mathbf{Q}$ which elements are:* $Q_{ij} = \begin{cases} q(B_i) & if \;\; i = j \\ 0 & otherwise. \end{cases}$

For the example concerning the ordered frame of discernment $\Omega = \{a, b, c\}$, the incidence matrix $\mathbf{M}$ is given in the table 2.

Table 2: Incidence matrix $\mathbf{M}$ when $\Omega = \{a, b, c\}$

	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	$\{a,b,c\}$
$\emptyset$	1	1	1	1	1	1	1	1
$\{a\}$	0	1	0	1	0	1	0	1
$\{b\}$	0	0	1	1	0	0	1	1
$\{a,b\}$	0	0	0	1	0	0	0	1
$\{c\}$	0	0	0	0	1	1	1	1
$\{a,c\}$	0	0	0	0	0	1	0	1
$\{b,c\}$	0	0	0	0	0	0	1	1
$\{a,b,c\}$	0	0	0	0	0	0	0	1

We can easily consider that the matrix $\mathbf{M}$ is upper triangular and it is built using the elementary constructing bloc:

$$\mathbf{M^1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The matrix $\mathbf{M^1}$ represents the incidence matrix when $|\Omega| = 1$. It is also clear that $length(\mathbf{M}) = 2^{|\Omega|}$. For an ordered frame of discernment with i elements, the computation of the incidence matrix can, in fact, be recursively computed from the incidence matrix of a frame containing $i-1$ elements. This construction is possible by using the Kronecker product of the constructing bloc $\mathbf{M^1}$ by the incidence matrix of the frame containing $i-1$ elements[12]:

$$\mathbf{M^i} = \mathbf{Kron}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \mathbf{M^{i-1}}\right) \tag{12}$$

In the same way, the matrix transformation of the basic belief assignment m into implicability function b is possible by using the matrix $\mathbf{B}$. we also can write $\mathbf{b} = \mathbf{B}.\mathbf{m}$ [12]. This matrix can be calculated from the matrix $\mathbf{M}$ as follow $\mathbf{B} = \mathbf{J}.\mathbf{M}.\mathbf{J}$ [12]. It can also be computed, like for the incidence matrix, by using the Kronecker product. The constructing bloc used in this case is:

$$\mathbf{B^1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

2.3. Dempsterian specialisation matrices

The concept of specialisation is basically founded on the redistribution of the masses of evidence after bringing new information knowledge. The impact of the new knowledge changes the initial commitment related to the problem and produces a new basic belief assignment at least as committed as the initial distribution [13, 14].

Definition 3. *A specialisation matrix* $\mathbf{S}$ *is a stochastic matrix [12] following the columns which coefficients* $S(A, B) = 0 \ \forall \ A \nsubseteq \ B$.

So, given a basic belief assignment m_0 which attributes the subset B the value $m_0(B)$. If $S(A, B) \in [0, 1]$ is the amount of the mass $m_0(B)$ which is transferred to the subset $A \subseteq B$ after specialisation and contributes on constructing $m_1(A)$, then :

$$m_1(A) = \sum_{B \subseteq \Omega} S(A, B).m_0(B) \tag{13}$$

In the aim of conserving the total commitment $m_0(B)$ after the transfer, the two following conditions are necessary:

$$\begin{cases} \sum_{A \subseteq B} S(A, B) = 1, \quad B \subseteq \Omega \\ S(A, B) > 0 \ \Rightarrow \ A \subseteq B \end{cases}$$

The basic belief assignment m_1 is called specialisation of m_0. It is achieved by the following matrix notation:

$$\mathbf{m_1} = \mathbf{S}.\mathbf{m_0}$$

where $\mathbf{S}$ is a square matrix which elements $S_{ij} = S(B_i, B_j)$ and $B_i, B_j \subseteq \Omega$.

2.3.1. The conjunctive combination rule

The conjunctive combination rule, certainly the most used rule in the information merging problems, provides a bba which is at least as committed as the merged bba's. It is so a form of specialisation process. It can be expressed using the concept of matrix representation. In fact, the conjunctive revision of a basic belief assignment m_1 by another basic belief assignment m_2 is realized by a specialisation matrix called Dempsterian [13] and denoted $\mathbf{S_{m_2}}$. This specialisation matrix is seen as a function of the basic belief assignment m_2. The conjunctive combination rule is written as:

$$\begin{aligned} m_1 \circledcap m_2(A) &= \sum_{B \cap C = A} m_1(B).m_2(C) && (14) \\ &= \sum_{B \subseteq \Omega} \Big(\sum_{B \cap C = A} m_2(C) \Big).m_1(B) && (15) \\ &= \sum_{B \subseteq \Omega} S_{m_2}(A, B).m_1(B) && (16) \end{aligned}$$

One can easily see that the elements $S_{m_2}(A, B)$ are only depending on the basic belief assignment m_2. It is also clear that these elements are the components of the so called Dempsterian specialisation matrix of m_2 since it is emerged from the non-normalized Dempster's combination rule.

The matrix notation of the conjunctive combination rule is :

$$\mathbf{m_1} \textcircled{\cap} \mathbf{m_2} = \mathbf{S_{m_2}}.\mathbf{m_1} \tag{17}$$

Definition 4. *Given a basic belief assignment m defined on a frame of discernment Ω, the related Dempsterian specialisation matrix is square which elements are obtained by :*

$$(S_m)_{ij} = \sum_{C \cap B_j = B_i} m(C) \tag{18}$$

- *The matrix* $\mathbf{S_m}$ *is upper triangular.*
- $\sum_{1 \leq i \leq 2^{|\Omega|}} (S_m)_{ij} = 1$ *for* $j \in \{1, 2, ..., 2^{|\Omega|}\}$.

The conjunctive combination rule must satisfy the the commutativity and the associativity properties. Its matrix formulation (equation (17)) highlights its linear aspect. Obviously, it is very easy to prove that the family of Dempsterian specialisation matrices is associative and commutative.

Theorem 1. *Given two Dempsterian specialisation matrices* $\mathbf{S_{m_1}}$ *and* $\mathbf{S_{m_2}}$ *obtained from two basic belief assignments* $\mathbf{m_1}$ *and* $\mathbf{m_2}$ *respectively. If the matrix* $\mathbf{S_{m_{12}}}$ *is so the Dempsterian specialisation matrix of the basic belief assignment* $\mathbf{m_{12}} = \mathbf{m_1} \textcircled{\cap} \mathbf{m_2}$ *resulting from the conjunctive combination of* $\mathbf{m_1}$ *and* $\mathbf{m_2}$*. Then, the following equality holds:*

$$\mathbf{S_{m_{12}}} = \mathbf{S_{m_1}}.\mathbf{S_{m_2}} = \mathbf{S_{m_2}}.\mathbf{S_{m_1}}$$

Proof. Given three basic belief assignments $\mathbf{m_1}$, $\mathbf{m_2}$ and $\mathbf{m_3}$. The result of their conjunctive combination is written:

$$\begin{aligned} \mathbf{m_{1 \textcircled{\cap} 2 \textcircled{\cap} 3}} &= \mathbf{m_1} \textcircled{\cap} \mathbf{m_2} \textcircled{\cap} \mathbf{m_3} \\ &= \mathbf{S_{m_1}}.\mathbf{m_2} \textcircled{\cap} \mathbf{m_3} \\ &= \mathbf{S_{m_1}}.\mathbf{S_{m_2}}.\mathbf{m_3} \end{aligned}$$

moreover, we know the conjunctive combination low is commutative and associative, then:

$$\begin{aligned} \mathbf{m_{1 \textcircled{\cap} 2 \textcircled{\cap} 3}} &= \mathbf{m_2} \textcircled{\cap} \mathbf{m_1} \textcircled{\cap} \mathbf{m_3} \\ &= \mathbf{S_{m_2}}.\mathbf{S_{m_1}}.\mathbf{m_3} \end{aligned}$$

We can therefore statute that the Dempsterian specialisation matrix of the basic belief assignment m_{12} which results on the combination of m_1 and m_2 is:

$$\mathbf{S_{m_{12}}} = \mathbf{S_{m_1}}.\mathbf{S_{m_2}} = \mathbf{S_{m_2}}.\mathbf{S_{m_1}}$$

□

Theorem 2. *given a Dempsterian specialisation matrix* $\mathbf{S_m}$ *and a communality matrix* $\mathbf{Q}$ *of some basic belief assignment* $\mathbf{m}$ *defined on Ω. If* $\mathbf{M}$ *is the incidence matrix related to this frame of discernment,* $\mathbf{S_m}$ *can be written as:*

$$\mathbf{S_m} = \mathbf{M^{-1}} * \mathbf{Q} * \mathbf{M} \tag{19}$$

Proof. Note that this theorem was proven in [14] using the set properties of the inclusion and intersection of focal elements. In this work, we use the matrix calculus formalism to provide a very simplified demonstration.

Given $\mathbf{m_{12}} = \mathbf{m_1} \circledcirc \mathbf{m_2}$. We can then calculate the communality vector $\mathbf{q_{12}}$ corresponding to $\mathbf{m_{12}}$ as:

$$
\begin{array}{rcl}
\mathbf{q_{12}} & = & \mathbf{M}.\mathbf{m_{12}} \\
 & = & \mathbf{M}.\mathbf{S_{m_1}}.\mathbf{m_2} \\
 & = & \mathbf{M}.\mathbf{S_{m_1}}.\mathbf{M^{-1}}.\mathbf{q_2} \\
\text{because :} & & \mathbf{q_2} = \mathbf{M}.\mathbf{m_2} \\
\text{but, the vector } \mathbf{q_{12}} & = & \mathbf{diag}(\mathbf{q_1}).\mathbf{q_2} \\
 & = & \mathbf{Q_1}.\mathbf{q_2} \\
\text{Then, we obtain :} & & \\
\mathbf{Q_1}.\mathbf{q_2} & = & \mathbf{M}.\mathbf{S_{m_1}}.\mathbf{M^{-1}}.\mathbf{q_2} \\
\text{Finally} & & \\
\mathbf{S_{m_1}} & = & \mathbf{M^{-1}} * \mathbf{Q_1} * \mathbf{M}
\end{array}
$$

□

2.3.2. Dempsterian specialisation matrix and discounting

Proposition 1. *Given a basic belief assignment* $\mathbf{m}$ *defined over a frame of discernment* Ω *and* $\alpha \in [0, 1]$ *a discounting factor. The Dempsterian specialisation matrix of the discounted bba using the factor* α *and denoted* ${}^{\alpha}m$ *is:*

$$ {}^{\alpha}\mathbf{S_m} = (1 - \alpha).\mathbf{S_m} + \alpha.\mathbf{S_{m_\Omega}} \tag{20} $$

where $\mathbf{S_m}$ *is the Dempsterain specialisation matrix of the bba m and* $\mathbf{I}$ *the identity matrix.*

Proof. The basic belief assignment resulting from the discounting operation of $\mathbf{m}$ using the factor α is:

$$ {}^{\alpha}\mathbf{m} = (1 - \alpha).\mathbf{m} + \alpha.\mathbf{m_\Omega} \tag{21} $$

where $\mathbf{m_\Omega}$ is the vacuous belief assignment defined on Ω.
The communality of the discounted basic belief assignment is defined as ${}^{\alpha}\mathbf{q}$ such that:

$$ {}^{\alpha}\mathbf{q} = \mathbf{M}.{}^{\alpha}\mathbf{m} = \mathbf{M}.\Big((1 - \alpha).\mathbf{m} + \alpha.\mathbf{m_\Omega}\Big) $$

then:

$$ {}^{\alpha}\mathbf{q} = (1 - \alpha).\mathbf{q} + \alpha.\mathbf{q_\Omega} $$

$$ {}^{\alpha}\mathbf{Q} = (1 - \alpha).\mathbf{Q} + \alpha.\mathbf{Q_\Omega} \tag{22} $$

It is obvious that the communality matrix of the vacuous basic belief assignment is equal to the identity matrix: $\mathbf{Q_\Omega} = \mathbf{I}$. The dempsterian specialisation matrix associated to the discounted basic belief assignment ${}^{\alpha}\mathbf{m}$ is obtained by:

$$
\begin{array}{rcl}
{}^{\alpha}\mathbf{S_m} & = & \mathbf{M^{-1}}.{}^{\alpha}\mathbf{Q}.\mathbf{M} \\
 & = & \mathbf{M^{-1}}.\Big((\mathbf{1} - \alpha).\mathbf{Q} + \alpha.\mathbf{Q_\Omega}\Big).\mathbf{M} \\
 & = & (1 - \alpha).\mathbf{M^{-1}}.\mathbf{Q}.\mathbf{M} + \alpha.\mathbf{M^{-1}}\mathbf{Q_\Omega}.\mathbf{M}
\end{array}
$$

This result allows to write the following relationship:

$$ {}^{\alpha}\mathbf{S_m} = (1 - \alpha).\mathbf{S_m} + \alpha.\mathbf{S_{m_\Omega}} $$

where:$\mathbf{S_{m_\Omega}} = \mathbf{I}$

One can see that the Dempsterian specialisation matrix of the discounted bba is calculated analogically in the same way as the discounted bba itself. The analogy is realized by replacing the bba's by their related matrices. □

2.3.3. Intermediate result on Dempsterian specialization matrices of simple support bba's (sbba's)

Proposition 2. *Suppose a sbba on the frame Ω, $m = m_X^w$ focussed on the set $X \subsetneq \Omega$. If its Dempsterian specialization matrix is denoted by $S = S_{m_X^w}$, one has:*

$$S = \bar{w} S_{m_X} - wI, \tag{23}$$

with $\bar{w} = 1 - w$, S_{m_X} the Dempsterain specialization matrix of m_X, the categorical bba on the set X, and I the identity matrix.

Proof. By definition of specialization matrices, one has:

$$S(A, B) = m(A|B),$$

with $m(.|B)$ the sbba m after conditionning on B. Using the conditionning formula, one can write:

$$S(A, B) = \sum_{\substack{C \subseteq \Omega \\ C \cap B = A}} m(C).$$

Since m has only two focal elements, *i.e.* X and Ω, one has:

$$\begin{aligned} S(A, B) &= 1_{X \cap B = A} m(X) + 1_{\Omega \cap B = A} m(\Omega), \\ &= \bar{w} 1_{X \cap B = A} + w 1_{B = A}, \end{aligned} \tag{24}$$

with 1_Y the indicator function such that :

$$1_Y = \begin{cases} 1 & \text{if } Y \text{ is true,} \\ 0 & \text{otherwise .} \end{cases}$$

Let us denote by S_{m_X} the matrix such that:

$$S_{m_X}(A, B) = \begin{cases} 1 & \text{if } X \cap B = A, \\ 0 & \text{otherwise .} \end{cases}$$

Consequently, one can write:

$$S = \bar{w} S_{m_X} + wId.$$

Moreover, applying this result with $w = 0$ gives $S = S_{m_X}$. Consequently, S_{m_X} can also be interpreted as the specialization matrix of the categorical bba on X, which is denoted by m_X. □

3. Specialization matrix as a basis for a new belief function distance

3.1. A short review of bba distances

The distance measure is used in the evidence theory to describe the difference among distinct pieces of evidence. Since the introduction of the Dempster conflict degree, many distance measures are defined in the literature to deal with the disagreement between pieces of evidence while the conflict degree is not appropriate to measure dissimilarity between bba's. A survey about the most known distances defined so far in the belief functions theory is introduced in 2012 by Jousselme and Maupin [11].

A distance in the belief functions theory is a measure giving some scalar comparison information between two bba's. This measure is explicitly dependent on the structure of the compared bba's and must satisfy some properties.

Definition 5. *Given a frame of discernment $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$. A mapping $d : 2^\Omega * 2^\Omega \longrightarrow [0, 1]$ is called normalized distance between two bba's $m_1(.)$ and $m_2(.)$ defined in Ω if the following properties are satisfied:*

(P1)- Non-negativity and normality : $0 \leq d(m_1, m_2) \leq 1$.

(P2)- Symmetry : $d(m_1, m_2) = d(m_2, m_1)$.

(P3)- Definiteness : $d(m_1, m_2) = 0 \Leftrightarrow m_1 = m_2$.

(P3)- Triangle inequality : $d(m_1, m_2) \leq d(m_1, m_3) + d(m_3, m_2)$

In a general point of view, there exists two types of distance measures to deal with dissimilarity in the Dempster-Shafer theory. Classically, we name the direct measures those based on the own geometrical interpretation of the bba's. It is principally question of distance measures in a vector space generated by the focal elements. The well known measure of this kind is certainly the distance defined by A.L. Jousselme in 2001 [7]. It is based on a similarity measure between focal elements. This distance is given by:

$$d_J(m_1, m_2) = \sqrt{\frac{1}{2}(\mathbf{m_1} - \mathbf{m_2})' \, \mathbf{D} \, (\mathbf{m_1} - \mathbf{m_2})} \tag{25}$$

where $\mathbf{D}$ is the Jaccard's similarity matrix between the focal elements. Its components are:

$$D(A, B) = \begin{cases} 1 \;\; if \;\; A = B = \emptyset \\ \frac{|A \cap B|}{|A \cup B|} \;\; \forall A, B \in 2^{\Omega} \end{cases} \tag{26}$$

Diaz et al. [8] proposed also a distance based on a similarity measure between focal elements. They proposed to use a modified function of the similarity between focal elements. The resulting dissimilarity rewards the small cardinalities and penalises at the same time high cardinalities of the focal sets.

$$d_D(m_1, m_2) = \sqrt{\frac{1}{2}(\mathbf{m_1} - \mathbf{m_2})' \, \mathbf{F}(\mathbf{S}, R) \, (\mathbf{m_1} - \mathbf{m_2})} \tag{27}$$

where $\mathbf{S}$ is the similarity matrix, R is a ratio that evaluates the proximity to the total ignorance and $\mathbf{F}$ is a monotonically increasing function depending on $\mathbf{S}$ and R.

In the indirect measures, the bba is first transformed to a new space of representation dealing with uncertainty, then, a dissimilarity measure is after computed in this new representation space. In this aim, dissimilarity degree can be calculated by using the pignistic transformation $BetP$ of the bba into probability space for example (B. Tessem in 1993 [2], W .Liu in 2006[4]). The most used indirect degree of dissimilarity computed using the probability transformation is the Tessem's distance defined as follow:

$$d_T = \max_{A \subseteq \Omega} \{|BetP_1(A) - BetP_2(A)|\} \tag{28}$$

Zouhal and Denœux also defined a dissimilarity measure based on the pignistic transformation of bba's to measure the distance between some belief function and an indicator vector. This measure is so used to improve a classification algorithm based on the *k-NN* rule. It is defined as follows:

$$d_{ZD}(\mathbf{m_1}, \mathbf{m_2}) = \sqrt{\frac{1}{2} \sum_{\omega_i \in \Omega} \left(BetP_1(\omega_i) - BetP_2(\omega_i)\right)^2} \tag{29}$$

One can easily prove that the Tessem's measure and Zouhal and Denœux's measure are pseudo-distances since they don't respect the definiteness property of the definition 5.

Another transformation based in fuzzy set theory is also possible in the field of indirect dissimilarity measures. This latter is introduced in the belief function theory in 2011 by D. Han et al [9]. They also established a lossy transformation which converts the bba's into a membership functions and non-membership functions expressed in the framework of intuitionistic fuzzy set theory. The reason of such an operation is that, in the fuzzy set theory, there exists several well-defined measures of similarity between two bodies of knowledge.

The membership and non-membership functions to a some intuitionist fuzzy set provided from a given bba are given by:

$$\begin{cases} \mu(\omega_i) = Bel(\omega_i) \\ \nu(\omega_i) = 1 - Pl(\omega_i) \end{cases} \tag{30}$$

The hesitation degree of the element ω_i is then given by:

$$\pi(\omega_i) = 1 - \mu(\omega_i) - \nu(\omega_i) \tag{31}$$

Three dissimilarity measures are so established by using the intuitionist fuzzy set transformation:

$$\begin{aligned} d_{IF1}(\mathbf{m_1}, \mathbf{m_2}) = \tfrac{1}{2n} \textstyle\sum_{i=1}^{n} \Big(&|\mu_1(\omega_i) - \mu_2(\omega_i)| \\ +|\nu_1(\omega_i) - \nu_2(\omega_i)| + |\pi_1(\omega_i) - \pi_2(\omega_i)|\Big) \end{aligned} \tag{32}$$

$$d_{IF2}(\mathbf{m_1}, \mathbf{m_2}) = \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} \Big(\varphi_1(\omega_i) - \varphi_2(\omega_i)\Big)^p} \tag{33}$$

where $\varphi_k(\omega_i) = \frac{\mu_k(\omega_i)+1-\nu_k(\omega_i)}{2}$

$$d_{IF3}(\mathbf{m_1}, \mathbf{m_2}) = \frac{1}{\sqrt[p]{n}} \sqrt[p]{\sum_{i=1}^{n} \Big(\varphi^{\mu}_{1,2}(\omega_i) + \varphi^{\nu}_{1,2}(\omega_i)\Big)^p} \tag{34}$$

where: $\varphi^{\mu}_{1,2}(\omega_i) = \frac{|\mu_1(\omega_i)-\mu_2(\omega_i)|}{2}$
and $\varphi^{\nu}_{1,2}(\omega_i) = |\frac{1-\nu_1(\omega_i)}{2} - \frac{1-\nu_2(\omega_i)}{2}|$

Our aim is not to establish a comparative study between these different measures. But we propose, in this contribution, a new distance between belief functions defined over the power set and able to discriminate efficiently the differences between bba's without explicitly using the structural similarity between the focal sets.

3.2. Motivation of the work

The use of distance definition to deal with the dissimilarity measure in the belief functions theory is an interesting way to define useful measures to compare the bba's. But not all the distances we define in the belief functions theory are appropriate to quantify dissimilarity between belief functions because of the complexity related to the uncertainty. The simplest Euclidean distance, for example, don't take care of the similarity between the focal elements. It compares only the mass values, so the structural interactions due to intersection properties between the focal elements are not considered when using the simple Euclidean distance.

The measures based on the geometrical interpretation use all the information available in the belief function representation concept since they are defined over the power set 2^{Ω}. Whereas such a representation lakes of solid justifications [9] since the basis vectors of the geometrical space are spanned from the focal subsets of the frame of discernment which are not independent because they are linked by the inclusion properties.

In the probability field and also the fuzzy set theory, there is a well established measures of dissimilarity which are exploited in the belief functions theory. The bba's are first transformed into pignistic probabilities or fuzzy membership and non-membership functions by some lossy information transformations, and then, the dissimilarities are calculated in the new representation space. Since there is no bijection between the bba's and the pignistic probabilities, also the fuzzy membership functions, the resulting indirect dissimilarity measures might not use all the information provided from the bba's to perform dissimilarity measures between them.

Lets analyse the behaviour of the most used dissimilarity measures in the belief functions theory by the following simple example :

Example 1. *Given a frame of discernment $\Omega = \{\omega_1, \omega_2\}$. Three bba's are so defined as follow :*

$$\begin{aligned} &m_1(\{\omega_1\}) = m_1(\{\omega_2\}) = 0.5 \\ &m_2(\Omega) = 1 \\ &m_3(\{\omega_1\}) = 1 \end{aligned}$$

Table 3: Different dissimilarity measures between $m_1(.)$, $m_2(.)$ and $m_3(.)$ - Example 1

Distances	$d(m_1, m_2)$	$d(m_1, m_3)$	$d(m_2, m_3)$
d_J	**0.5**	**0.5**	0.7071
d_T	**0**	**0.5**	**0.5**
d_Z	**0**	**0.5**	**0.5**
d_{IF1}	**1**	0.5	**1**
d_{IF2}	**0**	**0.5**	**0.5**
d_{IF3}	**0.5**	**0.5**	**0.5**

The results obtained from the main dissimilarity measures are given in table 3

In this example, one can see the m_1 is a uniform Bayesian distribution, m_2 is a vacuous bba whereas m_3 is a certain and categoric bba. The bba's m_1 and m_2 are very different in term of mass assignment and specificity but they produce the same lakes of indeterminate choice in decision making. In table 3, one can see that the measures d_T, d_Z and d_{IF2} provide a null result when comparing the two different bba's m_1 and m_2. Such a result is not satisfactory because the property $(P3)$ of a distance measure is not satisfied (see definition 5). These measures are pseudo-distances. Using the Jousselme's distance, the results show that the distance between m_1 and m_2 is the same as between m_1 and m_3. So Jousselme's distance cannot make difference between the two different situations presented in the example above dealing with specificity and decision making. The same problem is observed when analysing the results of the measures d_{IF1} and d_{IF3}. These two measures cannot clearly discriminate the three compared situations.

According to the results provided from the simple example presented above, one can see that all the presented measures suffer from lakes dealing with the specificity and the commitment of the three compared bba's. Intuitively, the obtained results are not satisfactory. These results are so questionable and not very persuasive.

Until now, there is no well established distance measure in the belief functions theory. Many recent works have treated the problem but the search for a new distance between bba's is not yet elucidated and remains still a challenge for the belief functions community.

3.3. New distance based on the Dempsterian specialisation matrices

In this work, we introduce a new measure to compute the distance between bodies of evidence taking the maximum advantage of the information available in the belief functions framework. Our approach is to compare bba's through their Dempsterian specialisation matrices, since there is a one to one correspondence between each bba and its Dempsterian specialisation matrix.

When the conjunctive combination rule is used, the way that some bba interact with each other is provided from its Dempsterian specialisation matrix taking into account the structural interactions between all the subsets of the frame of discernment. Our idea is then to compare the bba's using their Dempsterian specialisation matrices.

In the Dempster-Shafer theory, a dissimilarity measure can be defined using the distance between the compared basic belief assignments. Given an increasing monotone function $f(x)$ and a normalized distance measure $d(m_1, m_2) \in [0, 1]$ between the bba's m_1 and m_2, one can write:

$$f(0) \leq f(d(m_1, m_2)) \leq f(1)$$

$$\Rightarrow 0 \leq \frac{f(d(m_1, m_2)) - f(0)}{f(1) - f(0)} \leq 1 \tag{35}$$

The normalisation of the function $f(x)$ given by the equation (35) is used to define a dissimilarity measure in the framework of Dempster-Shafer theory. The relationship between the dissimilarity and the distance between two basic belief assignment is so given through the function $f(x)$ by the formula:

$$D(m_1, m_2) = \frac{f(d(m_1, m_2)) - f(0)}{f(1) - f(0)} \tag{36}$$

Many choices of the function $f(x)$ are possible, note for instance :

- $f(x) = 1 - exp(-x) \quad \Rightarrow D(m_1, m_2) = \frac{1-exp(-d(m_1,m_2)}{1-exp(-1)}$
- $f(x) = \frac{x}{1+x} \quad \Rightarrow D(m_1, m_2) = \frac{2.d(m_1,m_2)}{1+d(m_1,m_2)}$

For the rest of this work we use the simplest choice of the function $f(x) = x$. In other words, we define the dissimilarity between two basic belief assignments as the distance between them:

$$D(m_1, m_2) = d(m_1, m_2) \tag{37}$$

In the aim of defining new distance between bba's, we use a matrix norm to compare between their two Dempsterian specialisation matrices.

Definition 6. *In the K-vector space $\mathcal{M}_n$ of square matrices of size $n * n$, a matrix norm is a mapping defined on $\mathcal{M}_n \longrightarrow \mathbb{R}_+$ and satisfying the following conditions:*

1. $\| \mathbf{A} \| = \mathbf{0} \Leftrightarrow \mathbf{A} = \mathbf{0}$
2. $\| \lambda\mathbf{A} \| = |\lambda| . \| \mathbf{A} \|$
3. $\| \mathbf{A} + \mathbf{B} \| \leq \| \mathbf{A} \| + \| \mathbf{B} \|$
4. $\| \mathbf{A}.\mathbf{B} \| \leq \| \mathbf{A} \| . \| \mathbf{B} \|$

A norm $\mathcal{N}(x)$ is a structure of a metric space allowing to define a distance between the elements of this space. In a so normed space, a distance is defined by:

$$d(x, y) = \mathcal{N}(x - y) \tag{38}$$

To calculate the distance between bba's, we first calculate the Dempsterian specialisation matrices related to the compared bba's. Then, the Frobenius metric is used to evaluate the difference between these specialisation matrices.

Definition 7. *We call **Frobenius** norm of a square matrix* $\mathbf{A}$ *the quantity defined by:*

$$\| \mathbf{A} \|_F = \left(\sum_{1 \leq i,j \leq 2^{|\Omega|}} A_{ij}^2 \right)^{\frac{1}{2}} \tag{39}$$

Remark 1. *One can easily prove that :*

- $\| \mathbf{A} \|_F = \left(trace(\mathbf{A}^t \mathbf{A}) \right)^{\frac{1}{2}}$
- $\| \mathbf{I}_n \|_F = \sqrt{n}$*, where* $\mathbf{I_n}$ *is the identity matrix of size n.*
- *given a Dempsterian specialisation matrix* $\mathbf{S_m}$*,*
  $\| \mathbf{S_m} \|_F \leq \| \mathbf{S_{m_\Omega}} \|_F = \sqrt{n}$.

This leads us to the following distance definition between bba's.

Definition 8. *given two basic belief assignments m_1 and m_2, their dempsterian specialisation matrices are respectively* $\mathbf{S_{m_1}}$ *and* $\mathbf{S_{m_2}}$*. The distance between m_1 and m_2 is computed as:*

$$d_S(m_1, m_2) = \frac{1}{\rho} \| [\mathbf{S_{m_1}} - \mathbf{S_{m_2}} \|_F \tag{40}$$

where $\rho = 2.(2^{|\Omega|} - 1)$ *is a normalisation coefficient.*

The calculus of the normalisation coefficient ρ is performed in the annexe 6.1, item 2.

3.3.1. Main properties

In this paper we introduced a full metric to compute the distance between bba's. It satisfies all the basic formal properties provided from the definition 5. Further, it satisfies also some specific properties related to the belief functions theory concept. These specific properties, also the main properties of a full metric, are noted hereafter. The proofs of these properties are given in the annexe 6.1.

Proposition 3. *Given m_1, m_2 and m_3 three basic belief assignments defined on Ω and $\alpha \in [0,1]$ a discounting factor, then the following properties hold for d_S:*

1. *Definiteness :* $d_S(m_1, m_2) = 0 \;\Leftrightarrow\; m_1 = m_2$.
2. *Symmetry :* $d_S(m_1, m_2) = d_S(m_2, m_1)$.
3. *Non-negativity and normality :* $0 \leq d_S(m_1, m_2) \leq 1$.
4. *Triangle inequality :* $d_S(m_1, m_2) \leq d_S(m_1, m_3) + d_S(m_3, m_2)$.
5. *Same rate discounting :* $d_S({}^{\alpha}m_1, {}^{\alpha}m_2) = (1-\alpha).d_S(m_1, m_2) \leq d_S(m_1, m_2)$
6. *Conjunctive rule and distance :* $d_S(m_{1 \circledcirc 2}, m_1) \leq \| \mathbf{S_{m_1}} \|_F \, .d_S(m_2, m_\Omega)$
7. *Discounting based triangle inequality:* $d_S({}^{\alpha}m_1, m_2) \leq (1-\alpha).d_S(m_1, m_2) + \alpha.d_S(m_2, m_\Omega)$
8. $(1-\alpha).d_S(m_1, m_2) - \alpha.d_S(m_2, m_\Omega) \leq d_S({}^{\alpha}m_1, m_2)$

3.3.2. Belief ratio property

Proposition 4. *Suppose m_1 a bba on the frame Ω. If $m_2 = m_A^{w_2}$ and $m_3 = m_A^{1-a+aw_2}$ are two simple belief basic assignments (sbba) on Ω focussed on the set A, then the following relation holds for the distance d_S:*

$$\frac{d_S\left(m_1, m_1 \circledcirc m_2\right)}{d_S\left(m_1, m_1 \circledcirc m_3\right)} = \frac{m_2\left(A\right)}{m_3\left(A\right)} = a, \tag{41}$$

with $a \in [0,1]$.

Proof. Let S_i denote the Dempsterian specialization matrix of the sbba m_i. It is known that if S_{12} denotes the specialization matrix of bba $m_1 \circledcirc m_2$, one has $S_{12} = S_1.S_2$. Consequently, one can write:

$$\begin{aligned} S_1 - S_{12} &= S_1 - S_1 S_2, \\ S_1 - S_{12} &= S_1\left(I - S_2\right), \end{aligned} \tag{42}$$

with I the identity matrix. In addition, proposition 2 states that if m_2 is a sbba, then one has $S_2 = \bar{w}_2 S_{A_2} + w_2 I$ with S_{A_2} a the specialization matrix of the categorical bba on A_2. Using this result in equation (42) yields:

$$\begin{aligned} S_1 - S_{12} &= S_1\left(I - \bar{w}_2 S_{A_2} - w_2 I\right), \\ S_1 - S_{12} &= S_1\left(\bar{w}_2 I - \bar{w}_2 S_{A_2}\right) \\ S_1 - S_{12} &= \bar{w}_2\left(S_1 - S_1 S_{A_2}\right). \end{aligned} \tag{43}$$

An immediate consequence of equation (43) is that :

$$d_S\left(m_1, m_1 \circledcirc m_2\right) = \bar{w}_2 d_S\left(m_1, m_1 \circledcirc m_{A_2}\right). \tag{44}$$

This result can be applied as well to m_3:

$$\begin{aligned} d_S\left(m_1, m_1 \circledcirc m_3\right) &= m_3\left(A_2\right) d_S\left(m_1, m_1 \circledcirc m_{A_2}\right), \\ d_S\left(m_1, m_1 \circledcirc m_3\right) &= \left(1 - \left(1 - a + a w_2\right)\right) d_S\left(m_1, m_1 \circledcirc m_{A_2}\right), \\ d_S\left(m_1, m_1 \circledcirc m_3\right) &= a \bar{w}_2 d_S\left(m_1, m_1 \circledcirc m_{A_2}\right). \end{aligned} \tag{45}$$

Finally, dividing each term of equation (45) by those of (44) gives:

$$\frac{d_S\left(m_1, m_1 \circledcirc m_2\right)}{d_S\left(m_1, m_1 \circledcirc m_3\right)} = \frac{m_2\left(A_2\right)}{m_3\left(A_2\right)} = a.$$

□

3.3.3. Maximal distance interpretation

Proposition 5. *Suppose a frame of discernment Ω. The following relation holds for the distance d_S:*

$$d_S(m_1, m_2) = 1 \Leftrightarrow m_1 = m_{A_1}, m_2 = m_{A_2} \text{ with } A_1 = \Omega \setminus A_2. \tag{46}$$

Proof. The two implications will be proved separately:

- Suppose m_1 and m_2 are two bbas on Ω such that $d_S(m_1, m_2) = 1$. By definition of Loudahi distance, one has:

  $$\sum_{A,B \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 = 2\left(2^n - 1\right),$$

  with n the cardinal of Ω. Since for all bba m_i, one has $S_{m_i}(\emptyset, \emptyset) = 1$, the above expression writes as follows:

  $$\sum_{\substack{B = \subseteq \Omega \\ B \neq \emptyset}} \sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 = 2\left(2^n - 1\right).$$

  As there is exactly $2^n - 1$ possible choices left for B, we have:

  $$\sum_{\substack{B = \subseteq \Omega \\ B \neq \emptyset}} \left(2 - \sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2\right) = 0. \tag{47}$$

  Now, let us expand the following expression:

  $$\sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 = \sum_{A \subseteq \Omega} S_{m_1}(A, B)^2 - 2 S_{m_1}(A, B) S_{m_2}(A, B) + S_{m_2}(A, B)^2.$$

  Since all elements of specialization matrices are positive, one has:

  $$\sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 \leq \sum_{A \subseteq \Omega} S_{m_1}(A, B)^2 + S_{m_2}(A, B)^2.$$

  In addition, all elements of specialization are less or equal to 1, which implies that for all A, B one has $S_{m_i}(A, B)^2 \leq S_{m_i}(A, B)$. This allows us to write:

  $$\sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 \leq \sum_{A \subseteq \Omega} S_{m_1}(A, B) + S_{m_2}(A, B).$$

  Moreover, any specialization matrix S_{m_i} is such that for all B, $\sum_{A \subseteq \Omega} S_{m_i}(A, B) = 1$, hence the following expression:

  $$\begin{aligned} \sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 &\leq 2, \\ 2 - \sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 &\geq 0 \end{aligned} \tag{48}$$

  Using equation (48), expression (47) turns out to be a sum of positive terms. As this sum equals zero, it means that all of its terms equal zero. Consequently, one has:

  $$\forall B \subseteq \Omega \text{ and } B \neq \emptyset, \sum_{A \subseteq \Omega} \left(S_{m_1}(A, B) - S_{m_2}(A, B)\right)^2 = 2. \tag{49}$$

Let us use equation (49) with $B = \Omega$. In particular, one can write:

$$\begin{aligned}
2 &= \sum_{A\subseteq\Omega} \left(S_{m_1}(A,\Omega) - S_{m_2}(A,\Omega)\right)^2,\\
2 &= \sum_{A\subseteq\Omega} \left(S_{m_1}(A,\Omega)^2 - S_{m_1}(A,\Omega) S_{m_2}(A,\Omega) + S_{m_2}(A,\Omega)\right)^2,\\
2\left(1 + \sum_{A\subseteq\Omega} S_{m_1}(A,\Omega) S_{m_2}(A,\Omega)\right) &= \sum_{A\subseteq\Omega} \left(S_{m_1}(A,\Omega)^2 + S_{m_2}(A,\Omega)^2\right).
\end{aligned}$$

As any specialization matrix is such that $\forall A$, $S_{m_i}(A,\Omega) = m_i(A)$, one can write:

$$2\left(1 + \sum_{A\subseteq\Omega} m_1(A) m_2(A)\right) = \sum_{A\subseteq\Omega} \left(m_1(A)^2 + m_2(A)^2\right).$$

The fact that for any bba m_i, $\sum_{A\subseteq\Omega} m_i(A)^2 \leq 1$ has already been used earlier in this proof. Let us now make the following additional assumption: m_1 is not a categorical bbas. This assumption implies that $\forall A \subsetneq \Omega$, $m_1(A) < 1$ and consequently $\sum_{A\subseteq\Omega} m_1(A)^2 < 1$. This would imply:

$$2\left(1 + \sum_{A\subseteq\Omega} m_1(A) m_2(A)\right) < 2,$$

which is impossible. Consequently, m_1 is a categorical bba on a set denoted by A_1. The same reasonning holds for m_2, which is a categorical bba on a set denoted by A_2.
Finally, let us make the following additional assumption: $\exists A_3 \neq \emptyset$ such that $A_3 = \Omega \setminus (A_1 \cup A_2)$. One can apply relation (49) with $B = A_3$. In particular, one can write:

$$\sum_{A\subseteq\Omega} \left(S_{m_1}(A,A_3) - S_{m_2}(A,A_3)\right)^2 = 2. \tag{50}$$

By definition of specialization matrices, one has: $S_{m_i}(A,A_3) = m_i(A|A_3)$ with $m_i(.|A_3)$ the bba m_i after conditionning by A_3. But since m_1 and m_2 are categorical bbas and $A_1 \cap A_3 = \emptyset$ and $A_2 \cap A_3 = \emptyset$, we have $m_1(.|A_3) = m_2(.|A_3) = m_\emptyset$, the total conflict bba. This result implies that $\sum_{A\subseteq\Omega} \left(S_{m_1}(A,A_3) - S_{m_2}(A,A_3)\right)^2 = 0$ which is in contradiction with equation (50). Consequently, the assumption is wrong, meaning that $A_1 = \Omega \setminus A_2$.

- Suppose m_1 is a categorical bba on the set A_1, m_2 is a categorical bba on the set A_2 and $A_1 \cap A_2 = \emptyset$. By definition of Loudahi distance, one has:

$$d_S(m_1,m_2) = \frac{1}{2(2^n - 1)} \sum_{A,B\subseteq\Omega} \left(S_{m_1}(A,B) - S_{m_2}(A,B)\right)^2,$$

with n the cardinal of Ω. Applying equation (24) of proposition (2) for both m_1 and m_2 yields:

$$d_S(m_1,m_2) = \frac{1}{2(2^n - 1)} \sum_{A,B\subseteq\Omega} \left(1_{A_1\cap B=A} - 1_{A_2\cap B=A}\right)^2.$$

Since $A_1 \cap A_2 = \emptyset$, A cannot be a subset of both A_1 and A_2 unless $A = \emptyset$. One can thus write:

$$d_S(m_1,m_2) = \frac{1}{2(2^n - 1)} \left(\sum_{A\neq\emptyset,B\subseteq\Omega} 1_{A_1\cap B=A} + \sum_{A\neq\emptyset,B\subseteq\Omega} 1_{A_2\cap B=A} + 4\sum_{B\subseteq\Omega} 1_{A_1\cap B=A_2\cap B=\emptyset}\right).$$

As $A_1 = \Omega \setminus A_2$, there is no such B with $A_1 \cap B = A_2 \cap B = \emptyset$. The expression reduces thus to:

$$d_S(m_1,m_2) = \frac{1}{2(2^n - 1)} \left(\sum_{A\neq\emptyset,B\subseteq\Omega} 1_{A_1\cap B=A} + \sum_{A\neq\emptyset,B\subseteq\Omega} 1_{A_2\cap B=A}\right).$$

Computing the two remaining sums boils down to a counting problem. For each of these sums, there are $2^n - 1$ choices for set A and choosing A sets B. Consequently, one has:

$$\begin{aligned} d_S(m_1, m_2) &= \frac{1}{2(2^n-1)}(2^n - 1 + 2^n - 1), \\ d_S(m_1, m_2) &= 1. \end{aligned}$$

□

From the property 3, one can easily prove that $d_S(m_1, m_2) = 1$ only when the bba m_1 is the negation of m_2 and both are categoric; $\exists A \subseteq \Omega\ :\ \bar{m}_1(A) = m_2(\bar{A}) = 1$. Jousselme's distance, also Tessem's distance, is equal to 1 when the two compared bba's are categoric and completely conflicting; $m_1(A_1) = m_2(A_2) = 1$ and $A_1 \cap A_2 = \emptyset$. So, the specificity of the compared focal elements is not questioned in this case. This behaviour seems to be weaker than the above achieved by d_S because the negation condition is not respected.

In the next section, we will discuss the behaviour of our distance compared to the well known classically established distances in the belief functions theory.

4. Belief function distance comparison

The distances defined so far in the belief functions theory behave differently when they compare two bba's. Each of them has some particular features needed to achieve the requested objectives. In the sequel, we show the behaviour of the main established distances in the belief functions theory ,d_J, d_T, d_{ZD} and d_{IF2}, then we compare them to the distance we introduced in this paper d_S.

Example 2. *In this example, the convergence of the different compared distances is questioned when comparing non conflicting bba's. A similar example is proposed in [7] and reused in [9]. Given two bba's defined in an ordered frame of discernment of cardinality* $|\Omega| = 8$ *such as:*

$$\begin{aligned} m_1(\{\omega_1, \omega_2, \omega_3\}) &= 1 \\ m_2(A_t) &= 1 \end{aligned}$$

The subset A_t varies by successive inclusion of a singleton at each step from $\{\omega_1\}$ to reach Ω at the last step of computation. The results are shown in the figure (1).

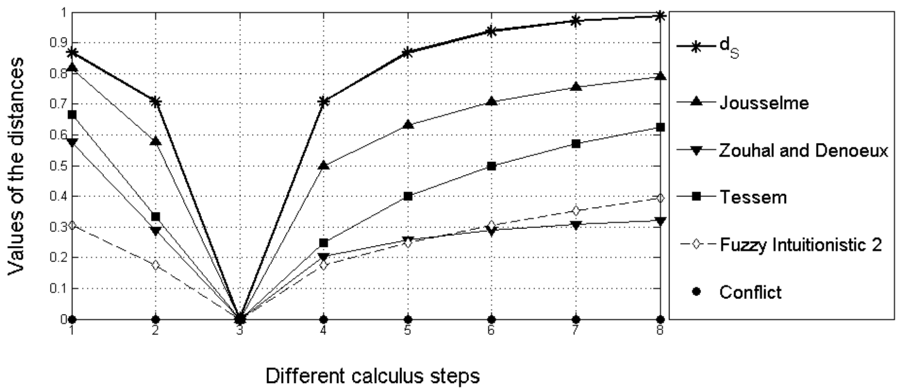

Figure 1: Dissimilarities between m_1 and m_2 - Example 2

In the example 2, the total mass assignment is attributed by the bba $m_1(.)$ to the subset $\{\omega_1, \omega_2, \omega_3\}$. Whereas, the bba $m_2(.)$ attributes its total mass assignment to the variable proposition A_t. Then, all the dissimilarity measures shown in this example present a similar behaviour and converge to their minimum value at the third step, when the two bba's believe that the truth is supported by the same focal element when $A_t = \{\omega_1, \omega_2, \omega_3\}$. The conflict degree is null at all the experience steps. One can also remark that the distances studied take into account the cardinality of the subset A_t in their variation. This means that the structural dissimilarity is taken into account by all the compared measures.

Example 3. *In this example, two different situations are observed, (1) totally conflicting bba's and (2) non-conflicting bba's. The maximal distance interpretation property of the proposed distance d_S is highlighted in this example. The behaviour of the studied distances varies from the maximum value (total divergence) to the minimum value (total convergence). The specificity of the compared bba's is questioned in the two observed situations. A similar example is proposed in [7]. Given two bba's defined in an ordered frame of discernment of cardinality $|\Omega| = 8$ such as:*

$$\begin{aligned} m_1(\{\omega_8\}) &= 1 \\ m_2(A_t) &= 1 \end{aligned}$$

The subset A_t variates by successive inclusion from $\{\omega_1\}$ to Ω, then its first element is subtracted at each step varying from Ω to ω_8.

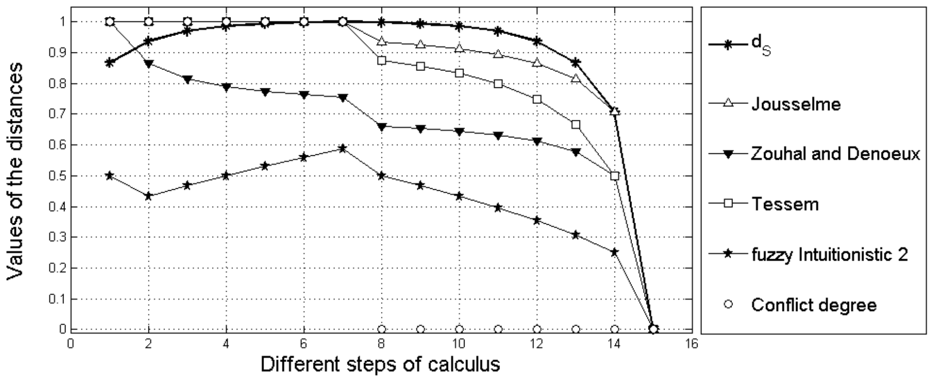

Figure 2: dissimilarities between m_1 and m_2 - Example 3

In the example 3, m_1 supports $\{\omega_8\}$ while m_2 supports the varying subset A_t. One can see that from step 1 to 7, $|A_t|$ is increasing and $A_t \cap \{\omega_8\} = \emptyset$. At step 7, A_t is the complementary set of $\{\omega_8\}$ in Ω. From step 8 to 15, $|A_t|$ is decreasing and $A_t \supseteq \{\omega_8\}$. At step 15 $A_t = \{\omega_8\}$. Analysing the results provided from the different distances summarized in figure (2), one can see that all the measures are null at the 15^{th} step when $A_t = \{\omega_8\}$. Zouhal and Denœux's distance, d_{ZD}, is persistently decreasing from 1 to 0 during all the computation steps. This result does not take care of the different parts of the simulation. The Jousselme's distance, d_j, is equal to 1 along the steps 1 to 7 ($A_t \cap \{\omega_8\} = \emptyset$). So, d_J does not take into account the variation of the specificity of A_t during these steps. From step 7 to 15, d_J is decreasing and reaches zero at step 15, when the two compared bba's support the same focal element. Tessem's distance behaves similarly as Jousselme's distance during all the simulation steps. The new proposed distance in this paper, d_S, presents more suitable results since it is still increasing from step 1 to 7 $A_t \cap \{\omega_8\} = \emptyset$ taking into account the specificity variations of A_t. It reaches the maximum value when A_t and $\{\omega_8\}$ are complementary in Ω. It is also decreasing as soon as the two subsets supported by the compared bba's are closer. $d_S = 0$ when $A_t = \{\omega_8\}$.

Example 4. *This example was proposed in [9]. The behaviour of the different distances with the increase of the cardinality of the frame of discernment is observed. The specificity and the commitment of the compared bba's are jointly questioned. Let's be $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$ a frame of discernment satisfying the Shafer's model. Three bba's are so defined in Ω as :*

$m_1(\{\omega_i\}) = \frac{1}{n}, \ \ \forall i \in \{1, \cdots, n\}$.
$m_2(\Omega) = 1$.
$m_3(\{\omega_k\}) = 1$, *for some* $k \in \{1, \cdots, n\}$.

The behaviour of the different dissimilarity measures studied in the example are visualized in the figure (3).

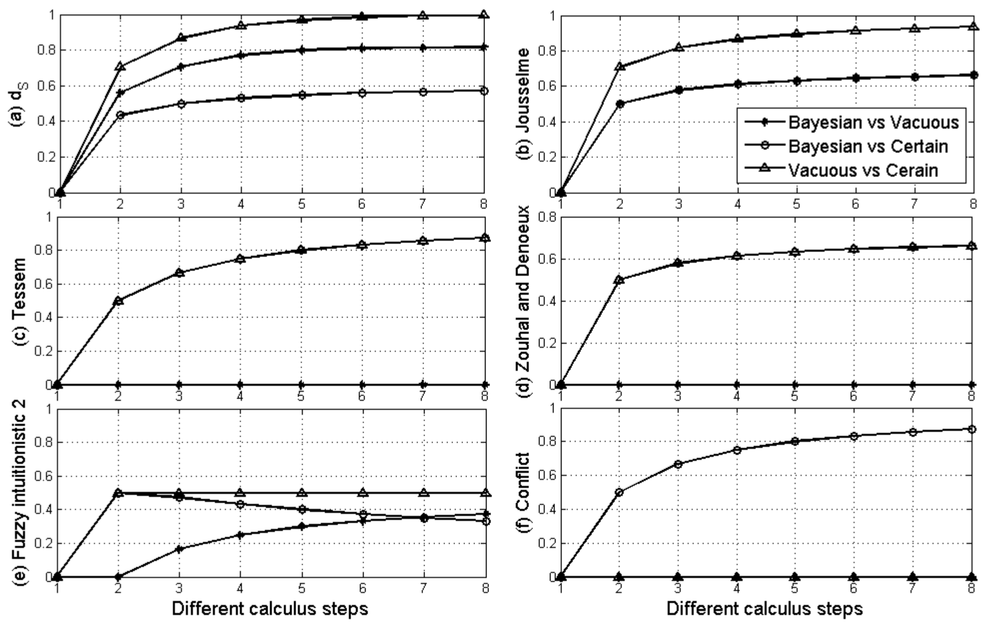

Figure 3: dissimilarities between m_1, m_2 and m_3 - Example 4

The behaviour of the dissimilarity measures studied in this paper with the increase of the cardinality of Ω is shown in the figure (3). The horizontal axis of the graphics represent the values of n. One can see, in this example, that $m_1(.)$ and $m_2(.)$ are different. They represent both two different uncertain sources while $m_3(.)$ is absolutely confident in ω_k. $m_1(.)$ is a uniform Bayesian Belief assignment whereas the bba $m_2(.)$ is less specific since it is a vacuous belief assignment. In figure (3b), Jousselme's distance cannot discriminate the difference between two different situations dealing with the specificity of the informations and the commitment of the sources. This result is observed while $d_J(m_1, m_2) = d_J(m_1, m_3)$. Zouhal's distance shows a completely different behaviour. In figure (3d), one can see that $d_{ZD}((m_1, m_2) = 0$. The same result is provided with Tessem's distance in figure (3c). This behaviour can not make a difference between $m_1(.)$ and $m_2(.)$ despite of the difference in the structure of the two bba's and the specificity of the informations provided by the two sources. One can also see that the two compared bba's present the same problem dealing with the lake of commitment. The measure d_{IF2},shown in figure (3e), can explicitly discriminate the three compared bba's in the present example when $n \neq 2$ but the behaviour of each measure provided from d_{IF2} is different from each other when n is still increasing. In figure (3a), one can see that the new established distance in this work, d_S, is increasing with the increase of the frame of discernment size n. It is also able to discriminate clearly the three situations presented in this example.

Example 5. *This example was proposed in [9]. It shows the principal difference between distances ans pseudo distances and also the incapacity of certain distances to discriminate a well different comparison situations between bba's. Given three bba's defined over a frame of discernment* $\Omega = \{\omega_1, \omega_2, \omega_3\}$ *as shown in table 4.*

The different results of the dissimilarity measures are given in table 5.

Table 4: Different values of the bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$ - Example 5

focal element	$\emptyset$	ω_1	ω_2	$\{\omega_1, \omega_2\}$	ω_3	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	Ω
$m_1(.)$	0	0.3333	0.3333	0	0.3333	0	0	0
$m_2(.)$	0	0.1	0.1	0	0.1	0	0	0.7
$m_3(.)$	0	0.1	0.1	0	0.8	0	0	0

Table 5: Different dissimilarity measures between $m_1(.)$, $m_2(.)$ and $m_3(.)$ - Example 5

Distances	$d(m_1, m_2)$	$d(m_1, m_3)$	$d(m_2, m_3)$
d_J	**0.4041**	**0.4041**	0.5715
d_T	**0**	<u>0.4667</u>	<u>0.4667</u>
d_{ZD}	**0**	<u>0.4041</u>	<u>0.4041</u>
d_{IF1}	**0.7000**	0.3111	**0.7000**
d_{IF2}	0.1167	0.3300	0.3500
d_{IF3}	**0.3500**	0.3300	**0.3500**
d	**0.4950**	**0.3500**	**0.6062**

One can see that $m_1(.)$ is equivalent to a uniform Bayesian distribution. The bba $m_2(.)$ attributes the large part of its mass assignment to the total ignorance whereas $m_3(.)$ believes mostly in the singleton ω_3. In [9], The authors have only observed the two first columns of table 5. In our work we make a more complete analysis by observing all the results of the distances provided from the example 5. A preliminary analysis shows that it is impossible to make a rational decision from $m_1(.)$ and $m_2(.)$ because of the uniform Bayesian distribution for the first and the large mass assignment given to the total ignorance while the other part of assignment is equally divided between the singletons for the farmer. These two bba's yield the same problem in the viewpoint of decision making despite the fact that they are very different in term of specificity of their informational content. Conversely, $m_3(.)$ assigns its large mass assignment to the hypothesis ω_3.

The results obtained using Jousselm's distance show that $d_J(m_1, m_2) = d_J(m_1, m_3)$. This is not satisfactory because one can see that $m_1(.)$ is more similar to $m_2(.)$ than $m_3(.)$ in the viewpoint of commitment and decision making. Otherwise, $m_1(.)$ is more similar to $m_3(.)$ than $m_2(.)$ in the viewpoint of specificity. Analysing the results of d_{ZD} one can see that $d_T(m_1, m_2) = 0$ where $m_1 \neq m_2$. This implies that d_T does not take care of the property *(P3)* provided from the definition 5. the same behaviour is also observed from the results of d_{ZD}. The results $d_{IF1}(m_1, m_2) = d_{IF1}(m_2, m_3)$, also $d_{IF3}(m_1, m_2) = d_{IF3}(m_2, m_3)$, are not satisfactory for the same reason as given for the results of d_J. Only d_{IF2} and d_S can provide acceptable results discriminating clearly the different compared bba's regarding their particular structures in this example.

Example 6. *We especially develop this example to prove that the intuitionistic measure d_{IF2} is a pseudo-distance and to show that the distance introduced in this paper provides suitable results in more general cases. Given three bba's defined in a frame of discernment $\Omega = \{\omega_1, \omega_2, \omega_3\}$ as shown in table 6.*

Table 6: Different values of bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$ - Example 6

focal element	$\emptyset$	ω_1	ω_2	$\{\omega_1, \omega_2\}$	ω_3	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	Ω
$m_1(.)$	0	0.3	0.5	0	0	0.1	0	0.1
$m_2(.)$	0	0.2	0.4	0.2	0	0.1	0	0.1
$m_3(.)$	0	0.2	0.4	0	0	0.1	0.2	0.1

The resulting dissimilarity measures are summarized in the table 7.

Table 7: Diffrent dissimilarity measures between bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$ - Example 6

Distances	$d(m_1,m_2)$	$d(m_1,m_3)$	$d(m_2,m_3)$
d_J	0.1000	0.1414	0.1633
d_T	**0**	<u>0.1000</u>	<u>0.1000</u>
d_{ZD}	**0**	<u>0.1000</u>	<u>0.1000</u>
d_{IF1}	**0.1333**	0.1667	**0.1333**
d_{IF2}	**0**	<u>0.0816</u>	<u>0.0816</u>
d_{IF3}	**0.0816**	0.1000	**0.0816**
d	**0.1118**	**0.1414**	**0.1732**

Analyzing the three bba's in the viewpoint of decision making, one can see that they are in agreement and support the same hypothesis ω_2. One can also observe that the measures d_T, d_{ZD} and d_{IF2} provide a zero value when comparing the two structurally different bba's m_1 and m_2. These same dissimilarity measures provide also the same value when comparing separately each of the bba's m_1 and m_2 to another one (m_3 for example). These two results are not satisfactory and not convenient because the property *(P3)* of definition 5 is not satisfied. Conversely to the above results, only d_J and d_s can clearly discriminate the three compared bba's.

In the light of the results given by the above simulation examples, some critical conclusions are made. One can observe that the studied measures are satisfactory in general use but they suffer from some drawbacks and provide a not satisfactory results in some special situations. For example, the measures d_T, d_{ZD} and d_{IF2} do not satisfy the property *(P3)* of definition 5.

5. Conclusions

6. Appendices

6.1. Proofs of the main properties

Given three basic belief assignments m_1, m_2 and m_3, which Dempsterian specialisation matrices are respectively $\mathbf{S_{m_1}}$, $\mathbf{S_{m_2}}$ and $\mathbf{S_{m_3}}$, and $\forall\ \alpha \in [0,1]$ a discounting coefficient.

1. Definiteness:
   If $\mathbf{m_1} = \mathbf{m_2}$, then $\mathbf{S_{m_1}} = \mathbf{S_{m_2}}$ and also :

$$\begin{aligned} d(m_1,m_2) &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \\ &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_1}} \parallel_F \\ \Rightarrow\ d(m_1,m_2) &= 0. \end{aligned}$$

   because $\parallel \mathbf{0}_n \parallel_F = 0$, and $\mathbf{0}_n$ a null matrix of size $n \geq 1$.

   If $d(m_1,m_2) = 0$ then $\parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F = 0$.
   this implies that:

$$\sum_{i=1}^{2^{|\Omega|}} \sum_{j=1}^{2^{|\Omega|}} \left[(S_{m_1})_{ij} - (S_{m_2})_{ij} \right]^2 = 0$$

   and this is only possible if: $(S_{m_1})_{ij} = (S_{m_2})_{ij}\ \ \forall\ \ i,j \in \{1,2,...,2^{|\Omega|}\}$.
   so we conclude that

$$\mathbf{S_{m_1}} = \mathbf{S_{m_2}}$$

   and finally :

$$\mathbf{m_1} = \mathbf{m_2}$$

2. Symmetry :
$\forall\ m_1$ and m_2, the following holds:

$$\begin{aligned} d(m_1, m_2) &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \\ &= \frac{1}{\rho} \parallel (-1)(\mathbf{S_{m_2}} - \mathbf{S_{m_1}}) \parallel_F \\ &= \frac{|-1|}{\rho} \parallel \mathbf{S_{m_2}} - \mathbf{S_{m_1}} \parallel_F \end{aligned}$$

And finally the property of symmetry is proved :

$$d(m_1, m_2) = d(m_2, m_1)$$

3. Non-negativity and normalisation:

   (a) Non-negativity

   $$d(m_1, m_2) = \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F$$

   One can obviously say that:

   $$\begin{cases} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \geq 0, \ \forall\ A \subseteq \Omega \\ \frac{1}{\rho} > 0, \ \forall\ \rho > 0. \end{cases}$$

   so, the following holds:

   $$\frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \geq 0$$

   in other words :

   $$d(m_1, m_2) \geq 0 \ \ \forall m_1, m_2 \text{ defined on } \Omega \text{ and } \rho > 0.$$

   (b) Normalisation

   $$\begin{aligned} d(m_1, m_2) &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \\ &= \frac{1}{\rho}\left(\sum_{i=1}^{2^{|\Omega|}} \sum_{j=1}^{2^{|\Omega|}} \Big[(S_{m_1})_{ij} - (S_{m_2})_{ij} \Big]^2 \right)^{\frac{1}{2}} \end{aligned}$$

   It is obvious that $\forall\ m_1$ and m_2 defined on Ω, we can obtain for $j = 1$:

   $$\sum_{i=1}^{2^{|\Omega|}} \Big[(S_{m_1})_{i1} - (S_{m_2})_{i1} \Big]^2 = 0$$

   the dissimilarity measure can be rewritten as:

   $$d(m_1, m_2) = \frac{1}{\rho}\left(\sum_{i=1}^{2^{|\Omega|}} \sum_{j=2}^{2^{|\Omega|}} \Big[(S_{m_1})_{ij} - (S_{m_2})_{ij} \Big]^2 \right)^{\frac{1}{2}}$$

   The distribution of the power of 2 results:

   $$\begin{aligned} d(m_1, m_2) &= \frac{1}{\rho}\left(\sum_{i=1}^{2^{|\Omega|}} \sum_{j=2}^{2^{|\Omega|}} \Big[(S_{m_1})_{ij}^2 + (S_{m_2})_{ij}^2 - 2.(S_{m_1})_{ij}.(S_{m_2})_{ij} \Big] \right)^{\frac{1}{2}} \\ &\leq \frac{1}{\rho}\left(\sum_{i=1}^{2^{|\Omega|}} \sum_{j=2}^{2^{|\Omega|}} \Big[(S_{m_1})_{ij}^2 + (S_{m_2})_{ij}^2 \Big] \right)^{\frac{1}{2}} \end{aligned}$$

because $2.(S_{m_1})_{ij}.(S_{m_2})_{ij} \geq 0 \;\; \forall \;\; m_1, m_2$.
and since $(S_{m_1})_{ij} \in [0, 1]$, one can obtain $(S_{m_1})^2_{ij} \leq (S_{m_1})_{ij} \;\; \forall m_1$ (it is the same case for $(S_{m_2})_{ij}$).
This result allows:

$$\begin{aligned} d(m_1, m_2) &\leq \frac{1}{\rho}\Big(\sum_{j=2}^{2^{|\Omega|}} \Big[\sum_{i=1}^{2^{|\Omega|}} (S_{m_1})_{ij} + \sum_{i=1}^{2^{|\Omega|}} (S_{m_2})_{ij} \Big]\Big)^{\frac{1}{2}} \\ &\leq \frac{1}{\rho}\Big(\sum_{j=2}^{2^{|\Omega|}} \Big[1 + 1 \Big]\Big)^{\frac{1}{2}} \\ &\leq \frac{1}{\rho}\Big(\Big[(2^{|\Omega|} - 1) + (2^{|\Omega|} - 1) \Big]\Big)^{\frac{1}{2}} \end{aligned}$$

We can so conclude that:

$$0 \leq d(m_1, m_2) \leq \frac{1}{\rho}.2.(2^{|\Omega|} - 1), \;\; \forall \;\; m_1, m_2$$

For normalizing the distance, the coefficient ρ must be defined as follow:

$$\rho = 2.(2^{|\Omega|} - 1). \tag{51}$$

Consequently, one can write:

$$0 \leq d(m_1, m_2) \leq 1, \;\; \forall \;\; m_1, m_2$$

4. The triangle inequality:

$$\begin{aligned} d(m_1, m_2) &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \\ &= \frac{1}{\rho} \parallel \Big(\mathbf{S_{m_1}} - \mathbf{S_{m_3}}\Big) + (\mathbf{S_{m_3}} - \mathbf{S_{m_2}}) \parallel_F \\ &\leq \frac{1}{\rho}\Big[\parallel \mathbf{S_{m_1}} - \mathbf{S_{m_3}} \parallel_F + \parallel \mathbf{S_{m_3}} - \mathbf{S_{m_2}} \parallel_F \Big] \\ \Rightarrow d(m_1, m_2) &\leq d(m_1, m_3) + d(m_3, m_2) \end{aligned}$$

5. The distance between two bba's discounted with the same rate factor is calculated as follow:

$$\begin{aligned} d(^{\alpha}m_1, ^{\alpha}m_2) &= \frac{1}{\rho} \parallel {}^{\alpha}\mathbf{S_{m_1}} - ^{\alpha}\mathbf{S_{m_2}} \parallel_F \\ &= \frac{1}{\rho} \parallel \Big((1-\alpha)\mathbf{S_{m_1}} + \alpha\mathbf{I}\Big) - \Big((1-\alpha)\mathbf{S_{m_2}} + \alpha\mathbf{I}\Big) \parallel_F \\ &= \frac{1}{\rho} \parallel (1-\alpha)\Big(\mathbf{S_{m_1}} - \mathbf{S_{m_2}}\Big) \parallel_F \\ &= (1-\alpha) d(m_1, m_2) \\ \Rightarrow d(^{\alpha}m_1, ^{\alpha}m_2) &\leq d(m_1, m_2) \end{aligned}$$

One can see that the distance between the two same rate discounted bba's is smaller than the distance between the two original bba's.

6. Conjunctive rule and distance.

Given:

$$\begin{aligned}
d(m_{1\circledcirc 2}, m_1) &= \frac{1}{\rho} \parallel \mathbf{S_{m_1}}.\mathbf{S_{m_2}} - \mathbf{S_{m_1}} \parallel_F \\
&= \frac{1}{\rho} \parallel \mathbf{S_{m_1}}.(\mathbf{S_{m_2}} - \mathbf{I}) \parallel_F \\
&\leq \frac{1}{\rho} \parallel \mathbf{S_{m_2}} - \mathbf{I} \parallel_F . \parallel \mathbf{S_{m_1}} \parallel_F \\
\Rightarrow d(m_{1\circledcirc 2}, m_1) &\leq \parallel \mathbf{S_{m_1}} \parallel_F .d(m_2, m_\Omega)
\end{aligned}$$

Note that if m_2 moves to ignorance, the bba resulting from the combination of m_1 and m_2 is moving to m_1 with a ratio $\parallel \mathbf{S_{m_1}} \parallel_F$. In other words, the variation of the distance between the bba m_1 and the result of its conjunctive combination with the bba m_2 is at most $\parallel \mathbf{S_{m_1}} \parallel_F$ times faster than than the variation of the distance between m_2 and the vacuous bba which is the neutral element of the conjunctive combination.

7. Discounting based triangle inequality:

$$\begin{aligned}
d({}^\alpha m_1, m_2) &= \frac{1}{\rho} \parallel {}^\alpha\mathbf{S_{m_1}} - \mathbf{S_{m_2}} \parallel_F \\
&= \frac{1}{\rho} \parallel \Big((1-\alpha)\mathbf{S_{m_1}} + \alpha\mathbf{I}\Big) - \mathbf{S_{m_2}} \parallel_F \\
&= \frac{1}{\rho} \parallel (1-\alpha)\Big(\mathbf{S_{m_1}} - \mathbf{S_{m_2}}\Big) + (\alpha)\Big(\mathbf{I} - \mathbf{S_{m_2}}\Big) \parallel_F \\
\Rightarrow d({}^\alpha m_1, m_2) &\leq (1-\alpha)d(m_1, m_2) + \alpha d(m_2, m_\Omega)
\end{aligned}$$

This property leads to another form of specific triangle inequality which is closely linked to the discounting operator.
One can note that the equality holds for the extreme cases, otherwise, the cases where $\alpha = 1$, $\mathbf{m_1} = \mathbf{m_2}$ and/or $\mathbf{m_2} = \mathbf{m_\Omega}$.

(a) If $\alpha = 1$, then ${}^\alpha\mathbf{m_1} = \mathbf{m_\Omega}$.
One obtains

$$d({}^1 m_1, m_2) = d(m_\Omega, m_2)$$

(b) if $\mathbf{m_1} = \mathbf{m_2}$:

$$\begin{aligned}
d({}^\alpha m_1, m_1) &= \frac{1}{\rho} \parallel^\alpha \mathbf{S_{m_1}} - \mathbf{S_{m_1}} \parallel_F \\
&= \frac{1}{\rho} \parallel \Big((1-\alpha)\mathbf{S_{m_1}} + \alpha\mathbf{I}\Big) - \mathbf{S_{m_1}} \parallel_F \\
&= \frac{1}{\rho} \parallel \alpha\Big(\mathbf{I} - \mathbf{S_{m_1}}\Big) \parallel_F \\
\Rightarrow d({}^\alpha m_1, m_1) &= \alpha d(m_1, m_\Omega)
\end{aligned}$$

(c) If $\mathbf{m_2} = \mathbf{m_\Omega}$:

$$
\begin{aligned}
d(^{\alpha}m_1, m_\Omega) &= \frac{1}{\rho} \parallel {}^{\alpha}\mathbf{S}_{\mathbf{m_1}} - \mathbf{I} \parallel_F \\
&= \frac{1}{\rho} \parallel \Big((1-\alpha)\mathbf{S}_{\mathbf{m_1}} + \alpha\mathbf{I}\Big) - \mathbf{I} \parallel_F \\
&= \frac{1}{\rho} \parallel (1-\alpha)\Big(\mathbf{S}_{\mathbf{m_1}} - \mathbf{I}\Big) \parallel_F \\
\Rightarrow d(^{\alpha}m_1, m_\Omega) &= (1-\alpha) d(m_1, m_\Omega)
\end{aligned}
$$

This property is graphically interpreted in the figure (7)

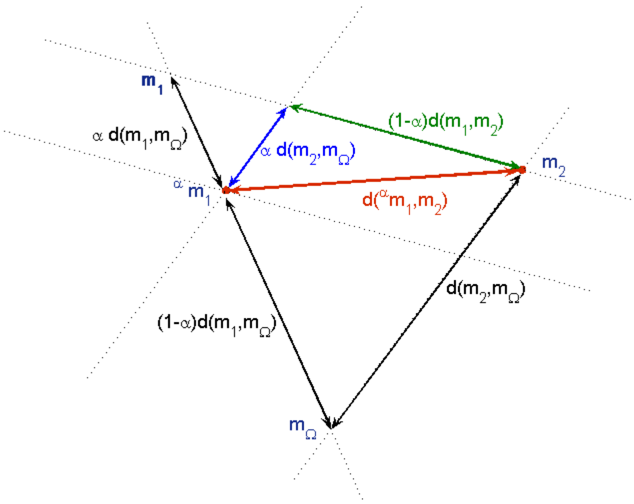

Figure 4: Geometrical interpretation of the 7^{th} property

8. Given :

$$
\begin{aligned}
(1-\alpha)d(m_1, m_2) &= \frac{1-\alpha}{\rho} \parallel \mathbf{S}_{\mathbf{m_1}} - \mathbf{S}_{\mathbf{m_2}} \parallel_F \\
&= \frac{1}{\rho} \parallel (1-\alpha)\Big(\mathbf{S}_{\mathbf{m_1}} - \mathbf{S}_{\mathbf{m_2}}\Big) \parallel_F \\
&= \frac{1}{\rho} \parallel (1-\alpha)\mathbf{S}_{\mathbf{m_1}} + \alpha\mathbf{I} - \mathbf{S}_{\mathbf{m_2}} - \alpha\Big(\mathbf{I} - \mathbf{S}_{\mathbf{m_2}}\Big) \parallel_F \\
&= \frac{1}{\rho} \parallel \Big({}^{\alpha}\mathbf{S}_{\mathbf{m_1}} - \mathbf{S}_{\mathbf{m_2}}\Big) + \alpha\Big(\mathbf{S}_{\mathbf{m_2}} - \mathbf{I}\Big) \parallel_F \\
&\leq d(^{\alpha}m_1, m_2) + \alpha d(m_2, m_\Omega) m_1, m_2)
\end{aligned}
$$

then

$$\Rightarrow (1-\alpha)d(m_1, m_2) - \alpha d(m_2, m_\Omega) \leq d(^{\alpha}m_1, m_2)$$

The two last properties can be summarized in the following :

$$|d(^{\alpha}m_1, m_2) - (1-\alpha)d(m_1, m_2)| \leq \alpha d(m_2, m_\Omega)$$

References

[1] G. Shafer, A mathematical theory of evidence., 1976.

[2] B. Tessem, Approximations for efficient computation in the theory of evidence, Artificial Intelligence 61 (1993) 315 – 329.

[3] M. Bauer, approximation algorithms and dicision making in the dempster-shafeer theory of evidence - an empirical study, International Journal of Approximate Reasoning 17 (1997) 217 – 237.

[4] W. Liu, Analysing the degree of conflict among belief functions, Artificial Intelligence 170 -11 (2006) 909–924.

[5] L. M. Zouhal, T. D. ux, An evidence-theoric k-nn rule with parameter optimisation, IEEE Transactions on Systems, Man and Cybernetics Part C28 (1998) 263 – 271.

[6] D. Fixen, R. Mahler, The modified dempster-shafer approach to classification, IEEE Transactions on Systems, Man and Cybernetics PART A 27 (1997) 96 – 104.

[7] A.-L. Jousselme, D. Grenier, E. Bossé, A new distance between two bodies of evidence, Information Fusion 2 (2001) 91–101.

[8] J. Diaz, M. Rifqi, B. Bouchon-Meunier, A similarity measure between basic belief assignments, in: Proceedings of the 9th International Conference Information Fusion, Firenze, Italy., 2006.

[9] D. Han, J. Dezert, chongzhao Han ahd Yi Yang, New dissimilarity measure in evidence theory., in: Fusion 2011 : 14th International Conference on Information Fusion, 2011.

[10] V. Khatibi, G. A. Montazer, A new evidential distance measure besed on belief intervals, Computer Science and engeneering and Electrical engeneering 17 - 2 (2010) 119 – 132.

[11] A.-L. Jousselme, P. Maupin, Distances in evidence theory : Comprehensive survey and generalizations, International Journal of Approximate Reasoning 53 (2012) 118 – 145.

[12] P. Smets, The application of the matrix calculus to belief functions, International Journal of Approximate Reasoning 31 (2002) 1–30.

[13] F. Klawonn, P. Smets, The dynamic of belief in the transferable belief model and specialisation-generalisation matrices, in: Proceedings of the eighth international conference on uncertainty in artificial inelligence. pp 130 - 137., 1992.

[14] P.-A. Monney, Dempster specialization matrices and the combination of belief functions, Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Lecture Notes in Computer Science. 2143 316-327 (2001) 316–327.

[15] F. Cuzzolin, Geometry of dempster's rule of combination, IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics, 34, NO. 2 (APRIL 2004) 961 – 977.

[16] F. Cuzzolin, A geometric approach to the theory of evidence, IEEE Transactions on Systems, Man, and Cybernetics. Part C, Applications and Reviews 38, NO. 4 (2008) 522–534.

[17] F. Pichon, Fonctions de croyance : Décompositions canoniques et régles de combinaison., Ph.D. thesis, Université de Technologie de Compiègne – UTC (2009).