On some properties of distances in evidence theory

Anne-Laure Jousselme
R & D Defence Canada Valcartier
2459, Pie XI North
Québec, QC, G3J 1X5,
Canada
Email: Anne-Laure.Jousselme@drdc-rddc.gc.ca

Patrick Maupin
R & D Defence Canada Valcartier
2459, Pie XI North
Québec, QC, G3J 1X5,
Canada
Email: Patrick.Maupin@drdc-rddc.gc.ca

***Abstract*— Dissimilarity measures have been widely studied in the field of probability theory and fuzzy set theory. The purpose of this paper is to provide an overview of the existing dissimilarity measures in evidence theory. We base our description on a geometrical interpretation of evidence theory. We show that most of the dissimilarity measures are based on inner products, in some cases degenerated. Experimental results outline some relationships between existing measures.**

Keywords: Dissimilarity, Distance, Similarity, Evidence theory, Belief functions.

I. Introduction

We are interested in this paper in surveying the main dissimilarity measures defined so far in the framework of evidence theory. The aim is to identify the theoretical foundations of these measures together with their properties. Experimental results illustrate some relationship between surveyed measures.

The vast body of literature on distances between probability distributions is of course a great source of inspiration for the definition of dissimilarities in evidence theory. Among the reference papers also aimed at studying dissimilarities in the broad sense, we can cite the works of M. Basseville [1] on distances and divergences in probability theory and of I. Bloch who proposed a detailed survey of distances between fuzzy sets in [2].

In recent years, many works on measuring the distance between belief functions have emerged. For a long time, Demspter's conflict factor has been the only way to quantify the interaction between belief functions (see for example [3], [4]). However, this factor may not be appropriate to quantify the dissimilarity between two belief functions as the conflict between two equal belief functions is not 0. Several ways for defining distances in evidence theory have been adopted: In [5], Perry and Stephanou extend the Kullback-Liebler divergence for probability distributions, Blackman and Popoli [6] and Ristic and Smets [7] define a distance based on Dempster's conflict factor. Other authors propose geometrical (Euclidean) distances: Fixsen and Mahler [8] define a classification miss-distance, Jousselme *et al.* [9] propose a geometric distance accounting for the interaction between sets, Cuzzolin [10] defines an Euclidean measure between belief values, Wen propose to quantify the similarity based as the angle between two mass vectors [11].

Two main needs may be identified regarding the use of dissimilarity measures between belief functions: (1) for algorithms evaluation or optimization, for example in classification algorithms [8], [9], or in belief functions approximation algorithms [10], [12]–[14], or for combination rules parameters estimation [15]–[17], (2) as a definition of agreement between sources of information, for example in clustering techniques [18], [19], or as a basis for discounting factors (for instance [11], [20]–[24]. In algorithms for evaluation or optimization, the distance is computed to a reference belief function Bel^r, while in the definition of agreement between sources of information, no such reference exists. Depending on the use, some formal properties are required while some other are superfluous. Our position is that none of the dissimilarity measures may be better than an another in the absolute, but rather that their choice should be directed by the practical use.

In Section II, we review some basics of evidence theory with an emphasis on the geometrical interpretation in Section II-B. The properties of similarities and dissimilarities are detailed in Section II-C. A classification of the distances is proposed in Section III based on different inner-products between belief functions. It is shown that most of the existing quantities can fit in this general formulation. We also mention other works of interest. A comparison of distances is proposed in Section IV through the use of correlation coefficients between distance functions, and Section V concludes on future works to be developed in upcoming papers.

II. Background

A. Basics on evidence theory

We denote by X the a frame of discernment containing N distinct objects x_i, $i = 1, \cdots, N$, and by x any element of X. 2^X is the power set of X. We denote by $\mathcal{F}$ the *set* of all the focal elements of a Basic Probability Assignment (BPA) m and by $\mathcal{C}$, the *core* of m (*i. e.* the *union* of the focal elements according to Shafer's definition [25]). A *body of evidence* is the couple $(m, \mathcal{F})$. Bel, Pl and q are the *belief*, *plausibility* and *commonality* of A respectively and since they are in one-to-one correspondence, they will be used interchangeably. The *pignistic probability* [26] is defined for all A of X by:

$$\mathrm{BetP}_m(A) = \sum_{B \subseteq X} m(B) \frac{|A \cap B|}{|B|} \tag{1}$$

where $|A|$ is the cardinality of set A. In particular, if A is a singleton $\{x\}$, we have $\mathrm{BetP}_m(\{x\}) = \sum_{x\in B} \frac{m(B)}{|B|}$.

We introduce the intersection index between two sets Int:

$$\mathrm{Int}(A,B) = 1 \text{ if } A\cap B \neq \varnothing \text{ and } 0 \text{ otherwise} \tag{2}$$

and the inclusion index Inc:

$$\mathrm{Inc}(A,B) = 1 \text{ if } A\subseteq B \text{ and } 0 \text{ otherwise} \tag{3}$$

We note that Inc is not symmetric ($\mathrm{Inc}(B,A) \neq \mathrm{Inc}(A,B)$ for all $A,B\subseteq X$) while Int is. The dual index of Int is $1-\mathrm{Int}$ which is such that $1-\mathrm{Int}(A,B)=1$ if $A\cap B=\varnothing$ and 1 otherwise. Introducing these indexes allows alternate notations for the belief, plausibility and commonality functions. For example, we have

$$\mathrm{Bel}(A) = \sum_{B\subseteq X} m(B)\mathrm{Inc}(B,A) \tag{4}$$

B. A geometrical interpretation of evidence theory

The geometrical interpretation of evidence theory can be traced back to the work of Ronald Mahler in 1996 [27], where the author set the bases in the random set theory framework. This interpretation has also been used in [9] to define a distance between two belief functions, and further developed in [10].

Let $\mathcal{E}_X$ be the 2^N-dimensional Cartesian space. The set of vectors $\{\mathbf{e_A}, A\subseteq X\}$ forms a basis of $\mathbb{R}^N$, so that any vector $\mathbf{v}$ of $\mathcal{E}_X$ can be written $\mathbf{v}=\sum_{A\subseteq X}\alpha_A\mathbf{e_A}$, where $\alpha_A\in\mathbb{R}$ is the coordinate of $\mathbf{v}$ along the dimension $\mathbf{e_A}$.

A BPA is a vector $\mathbf{m}$ of $\mathcal{E}_X$ such that $\sum_{A\subseteq X}\alpha_A=1$ with $\alpha_A\geq 0$ together with $\alpha_A=m(A)$. Using a vector-matrix notation, the belief, plausibility and commonality functions are written as:

$$\mathbf{Bel}=\mathbf{Inc}'.\mathbf{m} \quad \mathbf{Pl}=\mathbf{Int}.\mathbf{m} \quad \mathbf{q}=\mathbf{Inc}.\mathbf{m} \tag{5}$$

where $\mathbf{Inc}$ is the binary matrix whose elements are $\mathrm{Inc}(A,B)$, A in rows and B in columns, $\mathbf{Inc}'$ is the transpose matrix of $\mathbf{Inc}$ and $\mathbf{Int}$ is the symmetric matrix whose elements are $\mathrm{Int}(A,B)$.

C. Inner products and dissimilarities

Let assume that an inner product $\langle .,.\rangle$ exists over $\mathcal{E}_X$. Then, $\mathcal{E}_X$ together with $\langle .,.\rangle$ is an *inner product space*. An inner product must satisfy the 3 following axioms for all vectors $\mathbf{v_1},\mathbf{v_2},\mathbf{v_3}$ of $\mathcal{E}_X$ and all scalars $a\in\mathbb{R}$: (1) Symmetry: $\langle\mathbf{v_1},\mathbf{v_2}\rangle=\langle\mathbf{v_2},\mathbf{v_1}\rangle$, (2) Linearity in the first argument: $\langle a\mathbf{v_1}+b\mathbf{v_2},\mathbf{v_3}\rangle = a\langle\mathbf{v_1},\mathbf{v_3}\rangle + b\langle\mathbf{v_2},\mathbf{v_3}\rangle$, and (3) Positive-definiteness: $\langle\mathbf{v},\mathbf{v}\rangle\geq 0$ with equality only for $\mathbf{v}=0$. A general formulation for an inner product is:

$$\langle\mathbf{v_1},\mathbf{v_2}\rangle = \mathbf{v_1}'\mathbf{W}\mathbf{v_2} \tag{6}$$

where $\mathbf{W}$ is a matrix of weights required to be symmetric and positive-definite.

A norm in then naturally defined as $\|\mathbf{v}\|=\sqrt{\langle\mathbf{v},\mathbf{v}\rangle}$ and represents the "length" of $\mathbf{v}$. The angle between $\mathbf{v_1}$ and $\mathbf{v_2}$ is then:

$$\theta = \arccos\left(\frac{\langle\mathbf{v_1},\mathbf{v_2}\rangle}{\|\mathbf{v_1}\|.\|\mathbf{v_2}\|}\right) \tag{7}$$

The norm can be used to define a distance function on $\mathcal{E}_X$ by:

$$d(\mathbf{v_1},\mathbf{v_2}) = ||\mathbf{v_1}-\mathbf{v_2}|| = \sqrt{(\mathbf{v_1}-\mathbf{v_2})'\mathbf{W}(\mathbf{v_1}-\mathbf{v_2})} \tag{8}$$

The cosine measure is a natural measure of similarity: -1 means that $\mathbf{v_1}$ and $\mathbf{v_2}$ are opposite, 1 means that $\mathbf{v_1}$ and $\mathbf{v_2}$ are the same, 0 means that they are "independent" and in between values represent intermediate similarity or dissimilarity. Other formal definition of similarity exist. For instance, a function s on $\mathcal{E}_X$ is a *similarity* if and only if (1) s is symmetric $s(\mathbf{v_1},\mathbf{v_2})=s(\mathbf{v_2},\mathbf{v_1})$ and (2) $s(\mathbf{v},\mathbf{v})\geq s(\mathbf{v},\mathbf{v}')$. Furthermore if s satisfied $s(\mathbf{v},\mathbf{v})=1$, it is *normed*. Also, a function d on $\mathcal{E}_X$ is a *dissimilarity* if and only if (1) d is symmetric $d(\mathbf{v_1},\mathbf{v_2})=d(\mathbf{v_2},\mathbf{v_1})$, (2) non-negative $d(\mathbf{v_1},\mathbf{v_2})\geq 0$ and (3) reflexive $d(\mathbf{v},\mathbf{v})=0$. Furthermore, if d satisfies the triangle inequality (4) $d(\mathbf{v_1},\mathbf{v_2})\leq d(\mathbf{v_1},\mathbf{v_3})+d(\mathbf{v_2},\mathbf{v_3}) \quad \forall\mathbf{v_3}$, then d is a *semi-distance*. If d is definite (5) $d(\mathbf{v_1},\mathbf{v_2})=0 \Leftrightarrow \mathbf{v_1}=\mathbf{v_2}$, then d is a *distance*. A function in the form of Equation (8) is a distance if and only if the matrix $\mathbf{W}$ is symmetric and positive-definite. If it is not definite, then d is a semi-distance.

We distinguish thus two main types of measure of dissimilarity: (1) inner product type, which are like a covariance between $\mathbf{v}_1$ and $\mathbf{v}_2$ and (2) distance type, which are like a variance of the difference between $\mathbf{v}_1$ and $\mathbf{v}_2$. Figure 2 illustrates these two types of measure between belief functions.

III. Main distances in evidence theory

We will show in this section that in fact, most of the dissimilarity measures introduced in evidence theory are derived from inner products. We propose in the following to review the measures by first identifying the underlying inner products and discuss then the distances metrics they induce belonging to the Euclidean metrics family, known as L_2. We also present another family of metrics known as the Minkowski family, denoted as L_∞ and end this section with a divergence measure and a global conflict index.

A. Inner-products

1) The simplest inner product in $\mathcal{E}_X$ is:

$$s_E^{\otimes}(m_1,m_2) = \mathbf{m_1}'\mathbf{I}\,\mathbf{m_2} \tag{9}$$

where $\mathbf{I}$ is the identity matrix. This measure only accounts for the mass distribution over the focal elements, but not the interaction between the focal elements themselves.

2) In order to evaluate the performance of identification algorithms, Fixsen and Mahler proposed in [8] a "classification miss-distance metric" called BPAM, for Bayesian Percent Attribute Miss based on the following inner product:

$$s_F^{\otimes}(m_1,m_2) = \mathbf{m_1}'\mathbf{P}\,\mathbf{m_2} \tag{10}$$

where $\mathbf{P}$ is the matrix whose elements are:

$$P(A,B) = \frac{p(A\cap B)}{p(A)p(B)} \tag{11}$$

with p being a Bayesian *a priori* distribution on X.

3) In [9], we defined an inner product which accounts for the similarity between focal elements through the Jaccard index:

$$s_J^{\otimes}(m_1, m_2) = \mathbf{m_1}'\mathbf{Jac}\, \mathbf{m_2} \tag{12}$$

where $\mathbf{Jac}$ is the matrix whose elements are Jaccard indices:

$$\mathbf{Jac}(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{13}$$

This definition should be put together (10) since $P(A, B)$ can quantify the similarity between focal elements if p is a uniform distribution, *i. e.* $P(A, B) = \frac{N|A \cap B|}{|A||B|}$.

4) In [28], Diaz *et al.* propose to use any similarity measure between sets $S(A, B)$ for defining the matrix $\mathbf{W}$ and they suggest 17 possible measures. Furthermore, and it is the main purpose of their work, they propose to modify the similarity function between focal elements by a function F so that the resulting similarity measure has some interesting properties. Their idea is to "reward" small cardinalities while penalising high cardinalities of focal sets. The modification of inner products is then defined by:

$$s_{DI}^{\otimes}(m_1, m_2) = \mathbf{m_1}'F(\mathbf{S}, R)\mathbf{m_2} \tag{14}$$

where $\mathbf{S}$ is the matrix whose elements quantify similarity between focal elements.

5) Although not defined explicitly, Cuzzolin in [10] introduced an inner product through a natural Euclidean distance between belief values. Using the matrix notations introduced in (5), we have that:

$$s_{EB}^{\otimes}(m_1, m_2) = \mathbf{m_1}'\mathbf{Inc}\, \mathbf{Inc}'\mathbf{m_2} \tag{15}$$

$\mathbf{Inc}\, \mathbf{Inc}'$ is also a way to quantify the interaction between focal elements based on their inclusion rather than on their similarity.

6) Let us introduce $\mathbf{Bet}$ a matrix whose elements are $Bet(A, B) = \frac{|A \cap B|}{|B|}$. Then we have that $\mathbf{betp} = \mathbf{Bet}\, \mathbf{m}$, the pignistic distribution over $\mathcal{E}_X$, and we can define the inner product:

$$s_{Bet}^{\otimes}(m_1, m_2) = \mathbf{m_1}'\, \mathbf{Bet}'\mathbf{Bet}\, \mathbf{m_2} \tag{16}$$

7) The Bhattacharyya coefficient [29] can also define an inner product if we take the square root of the vectors $\mathbf{m_i}$:

$$s_B^{\otimes}(m_1, m_2) = \sqrt{\mathbf{m_1}}'\mathbf{I}\, \sqrt{\mathbf{m_2}} \tag{17}$$

8) The conflict factor defined by Dempster [30] is probably the first quantification of the interaction between two belief functions. It can also be under the form of an inner product:

$$d_C^{\otimes}(m_1, m_2) = \mathbf{m_1}'(1 - \mathbf{Int})\, \mathbf{m_2} \tag{18}$$

where $\mathbf{Int}$ is the matrix of intersections between two subsets of X introduced in Section II-A in Equation (2).

It is easy to show that all the inner products introduced above are symmetric and linear regarding the first component (axioms (1) and (2) of an inner product). However, some of them are degenerate as their matrix $\mathbf{W}$ is not positive-definite.

B. Dissimilarities based on inner products

1) Direct use: The most classical case of a direct use of an inner product is Dempster's conflict factor $d_C^{\otimes}$ (Equation (18)) that has been widely used for quantifying a notion of dissimilarity or distance between belief functions (see for example [19]). However in some cases, $d_C^{\otimes}$ may be not a suitable measure of dissimilarity as the internal conflict $d_C^{\otimes}(m, m)$ is not 0.

In [11], Wen *et al.* use $s_E^{\otimes}$ to define a cosine measure which is obtained by normalising the inner product (9)

$$s_{CS} = \cos\theta = \frac{s_E^{\otimes}(m_1, m_2)}{||\mathbf{m_1}|| \cdot ||\mathbf{m_2}||} \tag{19}$$

This measure defines a similarity between m_1 and m_2.

2) Through a function f: An extension of the Bhattacharyya distance for probability theory to evidence theory has been proposed in [7] and modified in [31], as is based on $s_B^{\otimes}$:

$$d_B(m_1, m_2) = \left(1 - s_B^{\otimes}(m_1, m_2)\right)^p \tag{20}$$

with $p \in \mathbb{R}^{+*}$.

Mahler suggests in [27] that a quantity of the form:

$$d_M = -\log(\langle \mathbf{m}, \mathbf{e_A} \rangle) \tag{21}$$

can be used as a basis for information measures on bodies of evidence. Based on $d_C^{\otimes}$, Ristic and Smets [7] define an additive global dissimilarity measure as:

$$d_R(m_1, m_2) = -\log\left(1 - d_C^{\otimes}(m_1, m_2)\right) \tag{22}$$

3) As a basis for Euclidean metrics: Each of the inner products introduced above potentially defines a distance in $\mathcal{E}_X$ of the form of (8) with vectors $\mathbf{v}$ satisfying the constraint of BPAs:

$$d_i(m_1, m_2) = \sqrt{(\mathbf{m_1} - \mathbf{m_2})'\mathbf{W}(\mathbf{m_1} - \mathbf{m_2})} \tag{23}$$

Among the possible distances, some have already been defined while others have not. The resulting distance is either a (weighted) Euclidean distance if $\mathbf{W}$ is positive-definite or a semi-distance if it does not satisfy the separability property. Moreover, not all the distances defined this way are appropriate for quantifying the dissimilarity between two belief functions. Indeed, the simplest Euclidean distance (denoted by d_E hereafter) between BPAs ($\mathbf{W} = \mathbf{I}$, in item 1) is not fully appropriate since it does not account for the interaction between focal elements of m_1 and m_2. Indeed, for instance $m_1(\{1, 2, 3\}) = 0.8$, $m_1(\{2, 3\}) = 0.2$ and $m_2(\{1, 2, 3, 4\}) = 1$ are very far from each other according to d_I, while intuitively they are not. The cosine measure of Wen *et al.* (19) can be questioned for the same reasons. The measure d_E has been used in [32] and in [7] in an association algorithm.[1]

A more intuitive definition requires that the distance accounts for the dissimilarity between (1) the sets of focal elements $\mathcal{F}_1$ and $\mathcal{F}_2$ and (2) the distribution of the masses

[1]We assume that the authors of [7] referred to Equation (9) instead of their expression (34) which is always equal to 1, as noticed in [31].

TABLE I
INNER PRODUCTS $\mathbf{m_1}'\mathbf{W}\mathbf{m_2}$ AND ASSOCIATED DISTANCES $\sqrt{(m_1 - m_2)'\mathbf{W}(m_1 - m_2)}$.

		$\mathbf{W}$	
Euclidean m	$s_E^{\otimes}$	$\mathbf{I}$	d_E
Euclidean Bel	$s_{EB}^{\otimes}$	$\mathbf{Inc}\ \mathbf{Inc}'$	d_{EB}
Conjunctive	$d_C^{\otimes}$	$(1 - \mathbf{Int})$	-
Jaccard	$s_J^{\otimes}$	$\mathbf{Jac}$	d_J
Fixsen and Malher	$s_F^{\otimes}$	$\mathbf{P}$	d_F
Pignistic	$s_{Bet}^{\otimes}$	$\mathbf{Bet}'\ \mathbf{Bet}$	d_{Bet}
Diaz *et al.*	$s_{Di}^{\otimes}$	$F(\mathbf{S}, R)$	d_{Di}
Bhattachayra coefficient	$s_B^{\otimes}$	$\mathbf{I}$	-

among them $\mathbf{m_1}$ and $\mathbf{m_2}$. This is the case for the distances proposed in [8]–[10], [28].

Another way to compare belief functions is through their betting ability: Two belief functions are close if their betting functions are close, *i. e.* if their pignistic transformations are close. Then, any distance between probability distributions can be used. The most natural one is the L_2 measure proposed in [17] to measure the distance between a belief function and an indicator vector. The distance is based on $s_{Bet}^{\otimes}(m_1, m_2)$, which also accounts for the similarity between focal elements. Table I summarises the distances built upon inner products.

A function of d_J is proposed in [23] as a measure of coherence between belief functions:

$$s_C(m_1, m_2) = \frac{1}{2}[\cos(\pi d_J) + 1] \tag{24}$$

In [6], Blackman and Popoli proposed a distance based the internal and external conflict of two belief functions:

$$d_{BP}(m_1, m_2) = ||\mathbf{m_1}||_C + ||\mathbf{m_2}||_C - d_C^{\otimes}(m_1, m_2) \tag{25}$$

where $||\mathbf{m}||$ is the square norm of m.

C. Chebyshev L_∞

In [12], with the aim of assessing the quality of Bayesian approximation algorithms of belief functions, Tessem proposed three error measures between two belief functions based on the Minkowski family, say L_∞ of Chebyshev. One measure between pignistic probabilities:

$$d_T(m_1, m_2) = \max_{A \subseteq X} \{|\mathrm{BetP}_1(A) - \mathrm{BetP}_2(A)|\} \tag{26}$$

The equivalent measures between belief values and between plausibilities of singletons are also defined in [12]. In [33], Cuzzolin discusses the problem of consistent approximations of belief function based on L_p measures between belief functions.

D. Divergence types

Arguing that some distances measure "the difference between in the amount of information available when they are considered separately and when they are combined", Perry and Stephanou proposed in [5] an extension of the symmetric version of Kullback-Liebler divergence for probability distributions based on the fact that the updating rule is Dempster's combination rule rather than Bayes' rule:

$$d_D(m_1, m_2) = \ (|\mathcal{F}_1 \cup \mathcal{F}_2| - |\mathcal{F}_1 \cap \mathcal{F}_2|) + \dots \sum_{A \in \mathcal{F}_1 \cup \mathcal{F}_2} m_1(A) m_2(A)(1 - (m_1 \oplus m_2)(A)) \tag{27}$$

where $\mathcal{F}_i$ is the set of focal elements of m_i and $\oplus$ is Dempster's rule of combination.

E. Two-dimensional measures

In [34], Liu defines a two-dimensional measure to better quantify the conflict between belief functions.

$$\mathrm{Ind}_L = \left(d_C^{\otimes}(m_1, m_2); d_T(m_1, m_2)\right) \tag{28}$$

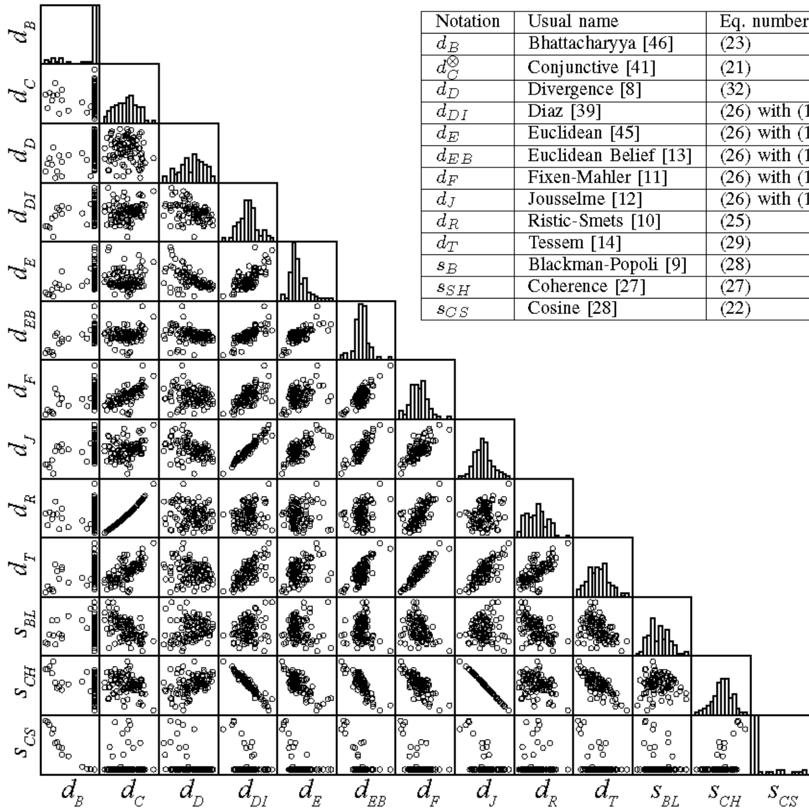
As we will see in Section IV, many measures of this type can be built based for example on correlation coefficient between measures.

IV. EXPERIMENTAL COMPARISON

We follow the technique described in [35] for analysing the semantic similarities between distances measures between belief functions. A number N_s of belief functions is randomly generated $\{\mathrm{Bel}^n\}_{n=1}^{N_s}$ (as described by **Algorithm 1**). The distances previously introduced in Section III are then computed for each pair (m^r, m^n), where m^r is a unique belief function of reference also randomly generated.

Input: X: Frame of discernment;
$N_{\max}$: Maximum number of focal elements
Output: Bel: Belief function (under the form of a BPA, m)
Generate the power set of $X \to P(X)$;
Generate a random permutation of $P(X) \to R(X)$;
Generate a integer between 1 and $N_{\max} \to k$;
foreach *First k elements of $R(X)$* **do**
| Generate a value within $[0, 1] \to m_k$;
end
Normalize the vector $\mathbf{m} = [m_1 \dots m_k] \to m'$;
$m(A_k) = m_k$;
Algorithm 1: Random generation of a belief function

Figures 1 and 2 show the results of simulation for $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $N_{\max} = 5$, $N_s = 100$ and $|\mathcal{D}| = 13$. In d_F, the prior probability distribution has been assumed uniform over X so that $P(A, B) = \frac{N|A \cap B|}{|A|.|B|}$. In d_{DI}, F has been chosen as in [28]. Although in the original paper [5], the authors of d_D restricted the m_i to be simple support functions, we removed this restriction in our simulations. Figure 1 shows the scatter plots for each pair $(d_i(m^r, m^n), d_j(m^r, m^n))$, $i, j \in \mathcal{D}$. We observe for instance that d_J and d_R are strongly correlated, in accordance to their respective definitions. On the contrary, d_D and d_T are very little correlated and could have been chosen for a two-dimensional measure as an alternative of Liu's measure [34] (see Equation (28)). The boxes on



Notation	Usual name	Eq. number
d_B	Bhattacharyya [46]	(23)
$d_C^{\otimes}$	Conjunctive [41]	(21)
d_D	Divergence [8]	(32)
d_{DI}	Diaz [39]	(26) with (17)
d_E	Euclidean [45]	(26) with (12)
d_{EB}	Euclidean Belief [13]	(26) with (18)
d_F	Fixen-Mahler [11]	(26) with (13)
d_J	Jousselme [12]	(26) with (15)
d_R	Ristic-Smets [10]	(25)
d_T	Tessem [14]	(29)
s_B	Blackman-Popoli [9]	(28)
s_{SH}	Coherence [27]	(27)
s_{CS}	Cosine [28]	(22)

Fig. 1. Scatter plots based on replications obtained using Algorithm 1.

TABLE II

STATISTICS CALCULATED ON 100 REPLICATIONS GENERATED BY ALGORITHM 1.

Statistic	d_B	d_C	d_D	d_{DI}	d_E	d_{EB}	d_F	d_J	d_R	d_T	s_B	s_{SH}	s_{CS}
Minimum	0,480	0	6,578	0,250	0,389	1,459	0,128	0,252	0	0,224	0,340	0,284	0
Maximum	1	0,549	13,089	0,667	0,762	5,081	0,553	0,643	0,346	0,721	0,853	0,851	0,354
Mean	0,952	0,241	10,302	0,464	0,514	2,952	0,287	0,444	0,125	0,428	0,569	0,586	0,025
Standard Deviation	0,132	0,120	1,603	0,075	0,067	0,498	0,071	0,069	0,071	0,100	0,108	0,105	0,077
Skewness	-2,620	-0,008	-0,340	0,252	1,326	0,293	0,415	0,285	0,341	0,265	0,351	-0,398	3,180
Kurtosis	5,332	-0,688	-0,641	0,484	2,383	4,172	1,322	0,516	-0,272	-0,092	-0,241	0,398	9,276

the diagonal of the scatter plot show the distributions of the measures, and table II completes the statistics.

In Figure 2, we built a dendogram (additive tree) computed on a Pearson correlation coefficient matrix between 100 replications of Algorithm 1, whose elements are:

$$c(d_i, d_j) = \frac{\sum_{n=1}^{N}(d_i^n - \bar{d}_i)(d_j^n - \bar{d}_j)}{\sqrt{\sum_{n=1}^{N}(d_i^n - \bar{d}_i)^2}\sqrt{\sum_{n=1}^{N}(d_j^n - \bar{d}_j)^2}} \qquad (29)$$

with $\bar{d}_i = \frac{1}{N}\sum_{n=1}^{N} d_i$. If $c(d_i, d_j) = 0$, the two distances are uncorrelated while the distances are all the more correlated than $c(d_i, d_j)$ is close to 1 (or -1).

We distinguish four (4) groups of measures: (I) Measures involving autovariance and $\mathbf{W}$ is a similarity matrix, (II) measures involving covariance only and $\mathbf{W}$ is a dissimilarity matrix, (III) measures involving covariance and the identity matrix and (IV) measures involving covariance and $\mathbf{W}$ is a similarity matrix. It is interesting to notice that d_T belongs to Group (I) while it is a priori a L_∞ measure and not a L_2. A deeper analysis of group (IV) is required to explain the non-trivial relationships between the measures of this group and justify our classification.

V. CONCLUSIONS

In this paper we have outlined the existence of a formal link between the existing distances (either dissimilarities and similarities) defined in the framework of belief functions and the broad domain of inner products. Experimental comparisons also showed that the surveyed measures could fit into four categories. These categories can be related to general properties such as the type of weighting matrix used, and whether covariances or variances terms are used for the definition

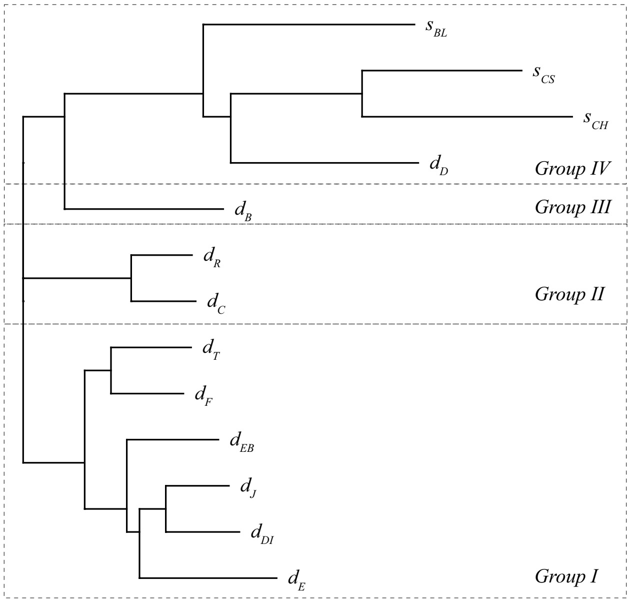


Fig. 2. Additive tree computed on a Pearson correlation matrix between replications based on Algorithm 1.

of the various measures. Future works will include (1) add other reference belief functions in the experimentation such as vacuous belief function and categorical belief functions, (2) use real data for the experimental comparison, and (3) detail furthermore the formal properties of the surveyed measures.

References

[1] M. Basseville, "Distance measures for signal processing and pattern recognition," *European Journal Signal Processing*, vol. 18, pp. 349–369, 1989.

[2] I. Bloch, "On fuzzy distances and their use in image processing under imprecision," *Pattern Recognition*, vol. 32, pp. 1873–1895, 1999.

[3] P. Smets, "Analyzing the combination of conflicting belief functions," *Information Fusion*, vol. 8, pp. 387–412, 2007.

[4] T. Denœux and M. Masson, "EVCLUS: Evidential Clustering of Proximity Data," *IEEE Transactions on Systems, Man and Cybernetics Part B*, vol. 34, pp. 95–109, 2004.

[5] W. L. Perry and H. E. Stephanou, "Belief Function Divergence as a Classifier," in *Proceedings of the 1991 IEEE International Symposium on Intelligent Control*, Arlington, VA (USA), 1991, pp. 280–285.

[6] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[7] B. Ristic and P. Smets, "The TBM global distance measure for the association of uncertain combat ID declarations," *Information Fusion*, vol. 7, pp. 276–284, 2006.

[8] D. Fixsen and R. P. S. Mahler, "The Modified Dempster-Shafer Approach to Classification," *IEEE Trans. on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, no. 1, pp. 96–104, Jan. 1997.

[9] A.-L. Jousselme, D. Grenier, and E. Bossé, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91–101, June 2001.

[10] F. Cuzzolin, "A Geometric Approach to the Theory of Evidence," *IEEE Transactions on Systems, Man, and Cybernetics - Part C : Applications and Reviews*, vol. 38, no. 4, pp. 522–534, 2008.

[11] C. Wen, Y. Wang, and X. Xu, "Fuzzy Information Fusion Algorithm of Fault Diagnosis Based on Similarity Measure of Evidence," in *Advances in Neural Networks*, ser. Lecture Notes in Computer Science, vol. 5264. Springer Berlin / Heidelberg, 2008, pp. 506–515.

[12] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61, no. 2, pp. 315–329, June 1993.

[13] M. Bauer, "Approximation algorithms and decision making in the Dempster-Shafer theory of evidence - An empirical study," *International Journal of Approximate Reasoning*, vol. 17, no. 2-3, pp. 217–237, 1997.

[14] A.-L. Jousselme, E. Bossé, and A. Jouan, "Analysing an identity information fusion algorithm based on evidence theory," in *IST-040/RSY-012, RTO Symposium on Military Data and Information Fusion*, Prague, Czech Republic, 20-22 October 2003.

[15] T. Denoeux, "Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence," *Artificial Intelligence*, vol. 172, pp. 234–264, 2008.

[16] F. Pichon and T. Denoeux, "T-norm and uninorm-based combination of belief functions," in *Annual Meeting of the North American Fuzzy Information Processing Society*, 2008.

[17] L. M. Zouhal and T. Denœux, "An evidence-theoretic *k*-NN rule with parameter optimization," *IEEE Transactions on Systems, Man and Cybernetics - Part C*, vol. 28, no. 2, pp. 263–271, 1998.

[18] S. Ben-Hariz, Z. Elouedi, and K. Mellouli, "Clustering Approach Using Belief Function Theory," in *Artificial Intelligence: Methodology, Systems, and Applications*, ser. Lecture Notes in Computer Science, J. Euzenat and J. Domingue, Eds., vol. 4183, 2006, pp. 162–171.

[19] J. Schubert, "Clustering decomposed belief functions using generalized weights of conflict," *International Journal of Approximate Reasoning*, vol. 48, pp. 466–480, 2008.

[20] M. C. Florea, E. Bossé, and A.-L. Jousselme, "Metrics, Distances and Dissimilarity Measures within Dempster-Shafer Theory to Characterize Sources' Reliability," in *Cognitive Systems with Interactive Sensors (COGIS '09)*, 2009.

[21] A. Martin, A.-L. Jousselme, and C. Osswald, "Conflict measure for the discounting operation on belief functions," in *Proceedings of the 11^{th} Annual Conference on Information Fusion*, Cologne, Germany, 2008.

[22] Y. Deng, D. Wang, and Q. Li, "An Improved Combination Rule in Fault Diagnosis based on Dempster-Shafer Theory," in *Proceedings of the Seventh International Conference on Machine Learning and Cybernetics*, 2008.

[23] L.-Z. Chen, W.-K. Shi, Y. Deng, and Z.-F. Zhu, "A new fusion approach based on distance of evidences," *Journal of Zhejiang University SCIENCE*, vol. 6A, no. 5, pp. 476–482, 2005.

[24] X. Guan, X. Yi, and Y. He, "Research on conflicting evidences combination strategies," in *Proceedings of the Seventh International Conference on Machine Learning and Cybernetics*, 2008.

[25] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[26] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[27] R. P. S. Mahler, "Combining ambiguous evidence with respect to ambiguous a priori knowledge, I: Boolean logic," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 26, no. 1, pp. 27–41, January 1996.

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

[29] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.

[30] A. Dempster, "Upper and lower probabilities induced by multivalued mapping," *The Annals of Mathematical Statistics*, vol. 38, pp. 325–339, April 1967.

[31] M. C. Florea and E. Bossé, "Crisis Management Using Dempster Shafer Theory: Using Dissimilarity Measures to Characterize Sources' Reliability," in *C3I in Crisis, Emergency and Consequence Management, RTO-MP-IST-086*, Bucharest, Romania, 2009.

[32] B. Quost, T. Denoeux, and M.-H. Masson, "Pairwise classifier combination in the transferable belief model," in *Proceedings of the 7th International Conference on Information Fusion*, Philadelphia, USA, July 2005.

[33] F. Cuzzolin, "Consistent approximations of belief functions," in *6th International Symposium on Imprecise Probability: Theories and Applications*, Durham, United Kingdom, 2009.

[34] W. Liu, "Analyzing the degree of conflict among belief functions," *Artificial Intelligence*, vol. 170, pp. 909–924, 2006.

[35] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 1, no. 4, pp. 300–307, 2007.