Comparative Study of Focal Distance Measures in Theory of Belief Functions

Yi Yang
SKLSVMS,
School of Aerospace
Xi'an Jiaotong University
Xi'an, China 710049
Email: jiafeiyy@mail.xjtu.edu.cn

Deqiang Han
Inst of Integrated Automation
School of Elect & Info Eng
Xi'an Jiaotong University
Xi'an, China 710049
Email: deqhan@gmail.com

Jean Dezert
ONERA,
The French Aerospace Lab
Chemin de la Hunière,
F-91761 Palaiseau, France
Email: jean.dezert@onera.fr

Yan Lin
Inst of Integrated Automation
School of Elect & Info Eng
Xi'an Jiaotong University
Xi'an, China 710049
Email: yanlin@mail.xjtu.edu.cn

***Abstract*—The focal distance is used to describe the degree of dissimilarity between focal elements in a basic belief assignment (BBA) by jointly using the focal elements' composition and their corresponding mass assignments. It has already been used in applications such as the design of the uncertainty degree of a BBA, and has been used as a criterion to remove focal elements in BBA approximations. The related works on focal distance are relatively limited when compared with researches on the distance between two BBAs. In this paper, the available definitions of focal distance are analyzed and compared based on experiments. All the existing focal distance measures are compared based on the applications aforementioned. We verify that all the existing focal distance measures are not strict metrics. This paper aims to call attention of the researchers on focal distance, and to provide the support for the selection of the focal distance in belief functions based applications.**

***Index Terms*—Belief functions, focal elements, focal distance, uncertainty, BBA approximation.**

I. Introduction

The theory of belief functions was first proposed by Dempster and then further developed by Shafer [1], thus it is also called the Dempster-Shafer evidence theory (DST). DST is an important tool of the uncertainty modeling and reasoning. It has been widely used in many applications like information fusion [2], pattern recognition [3], multiple-attribute decision making [4], etc. DST has been criticized due to its limitations and flaws by researchers [5], [6], [7]. Some generalized or refined theories have been proposed including the transferable belief model (TBM) [8] and Dezert-Smarandache Theory (DSmT) [9].

In DST, the basic belief assignment (BBA) is a basic function, which includes focal elements and their corresponding mass assignment. The focal distance [10] is for representing the degree of dissimilarity between two focal elements in a BBA based on the focal elements' composition (i.e., the elements included) and their corresponding mass assignments. The existing focal distance measures include Erkmen's focal distance [10], Denœux's two focal distance measures [11] using the intersection and union operation, respectively, and Liu's focal distance [12]. Focal distance has already been used in some belief function-related fields. For example, Erkmen [10] et al use a focal distance to construct the degree of uncertainty (called focal divergence) in a BBA. Denœux [11] uses two types of focal distance as criteria for BBA approximations. In our previous work [13], we use the focal distance to define the degree of redundancy of BBA, which is used as criteria for BBA approximations. Although the focal distance is an important concept, however, the related works on focal distance are relatively limited. Comparatively, there have already emerged many researches on the distance or dissimilarity between two BBAs. Researchers and users working in the DST community, to some extent, lack of sufficient attention to the focal distance.

Since the focal distance is useful in belief function related applications like uncertainty representation and BBA approximations, to call attention of the researchers on the focal distance, the available works on the focal distance are briefly reviewed in this paper. Existing focal distance definitions' behaviors and properties are analyzed based on illustrative examples. All focal distance definitions are compared in terms of two types of applications including the construction of the uncertainty degree, and the BBA approximations. We aim to provide references for the practical use of focal distance based on our experimental results and related analyses.

II. Basics of The Theory of Belief Functions

In Dempster-Shafer evidence theory (DST) [1], the elements in a discrete frame of discernment (FOD) Θ are mutually exclusive and exhaustive. In DST, a basic belief assignment (BBA, also called a mass function) $m : 2^\Theta \to [0, 1]$ satisfies:

$$\sum\nolimits_{A\subseteq\Theta} m(A) = 1, \quad m(\emptyset) = 0 \tag{1}$$

where 2^Θ is the powerset of Θ and A is called a focal element, if $m(A) > 0$. The belief function (Bel) and plausibility function (Pl) are defined respectively as follows.

$$\begin{aligned} Bel(A) &= \textstyle\sum_{B\subseteq A, B\in 2^\Theta} m(B); \\ Pl(A) &= \textstyle\sum_{A\cap B\neq\emptyset, B\in 2^\Theta} m(B). \end{aligned} \tag{2}$$

The length of a proposition or focal element A's belief interval $[Bel(A), Pl(A)]$ can be used describe the imprecision or uncertainty degree of the focal element A [1].

Dempster's rule of combination, which is used to combine distinct BBAs, is as follows. $\forall A \in 2^{\Theta}$:

$$m(A) = \begin{cases} 0, & \text{if} \quad A = \emptyset \\ \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1-K}, & \text{if} \quad A \neq \emptyset \end{cases} \tag{3}$$

where

$$K = \sum\nolimits_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) \tag{4}$$

is called the conflict coefficient representing the total degree of conflict between two BBAs. Dempster's rule has been criticized due to normalization by removing the conflict. Many alternative rules were proposed to redistribute the conflict [9].

As we can see, in DST, focal elements and the related mass assignments are the basics for the uncertainty representation and reasoning, therefore, it is worth researching some properties or indices related to the focal elements. The focal distance [10], which is used to describe the dissimilarity between focal elements, is an important concept. The existing works on focal distance are recalled in the next section.

III. Distances between Focal Elements

In general, the distance between two focal elements should use the two aspects of information in focal elements including the focal elements (set compositions) and the corresponding mass assignments as shown in

$$\delta(A_i, A_j) \triangleq f(m(A_i), A_i, m(A_j), A_j) \tag{5}$$

The available distances between focal elements are introduced below.

A. Existing focal distance measures

1) Erkmen's distance.
   Erkmen et al [10] proposed a distance (denoted by d_0^F here) as

$$\begin{aligned}\delta_E(A_i, A_j) = &\frac{1}{|A_i \cap A_j| \, / \, |A_i \cup A_j|} \\ &\cdot [m(A_i) - m(A_j)] \log_2 \frac{m(A_i)}{m(A_j)}\end{aligned} \tag{6}$$

   It is easy to find that Erkmen's distance borrows the idea of cross-entropy for the probability. The value of δ_E could be very large according to Eq. (6).
2) Denœux's intersection distance.
   Denœux [11] proposed an intersection-operation based distance as

$$\begin{aligned}\delta_\cap(A_i, A_j) = m(A_i) \cdot |A_i| + m(A_j) \cdot |A_j| \\ - [m(A_i) + m(A_j)] \cdot |A_i \cap A_j|\end{aligned} \tag{7}$$

2) Denœux's union distance.
   Denœux [11] also proposed a union-operation based distance as

$$\begin{aligned}\delta_\cup(A_i, A_j) = [m(A_i) + m(A_j)] \cdot |A_i \cup A_j| \\ - m(A_i) \cdot |A_i| - m(A_j) \cdot |A_j|\end{aligned} \tag{8}$$

   Actually, both $\delta_\cup$ and $\delta_\cap$ can be considered as a weighted sum of the Hamming distance [11].
4) Liu's focal distance.
   In Liu et al's work [12], a new focal distance is proposed as follows

$$\delta_L(A_i, A_j) = \left(1 - \frac{|A_i \cap A_j|}{|A_i \cup A_j|}\right) \cdot \frac{e^{|m(A_i) - m(A_j)|}}{e} \tag{9}$$

where $e \approx 2.7183$. Obviously, $\delta_L \in [0, 1)$.

B. Analysis of these existing focal distance measures

A strict distance metric should satisfy four requirements including non-negativity, non-degeneracy, symmetry, and triangle inequality.

1) On non-degeneracy: $m(A_i) = m(A_j)$.

Here, $\delta_E(A_i, A_j) = 0$, where the focal distance value d_E is totally dominated by mass assignments, although the compositions of A_i and A_j are different. That is, d_E violates the non-degeneracy, therefore, δ_E is not a strict distance metric.

In this case, δ_L, $\delta_\cap$, and $\delta_\cup$ will not bring the above counter-intuitive behavior.

2) On triangle inequality

Suppose that the FOD is $\Theta = \{\theta_1, \theta_2, \theta_3\}$. A BBA defined on Θ is

$$\begin{aligned}&m(\{\theta_2\}) = 0.6058, m(\{\theta_1, \theta_2\}) = 0.1498, \\ &m(\{\theta_1, \theta_3\}) = 0.0382, m(\{\theta_2, \theta_3\}) = 0.0454, \\ &m(\Theta) = 0.1608.\end{aligned}$$

Here assume that $A_1 = \{\theta_2\}, A_2 = \{\theta_1, \theta_2\}, A_3 = \{\theta_1, \theta_3\}$, $A_4 = \{\theta_2, \theta_3\}, A_5 = \Theta$. One can get that

$$\begin{aligned}&\delta_\cap(A_1, A_2) + \delta_\cap(A_2, A_3) \\ &= 0.1498 + 0.1880 = 0.3378 < \delta_\cap(A_1, A_3) = 0.6822 \\ &\delta_\cup(A_1, A_2) + \delta_\cup(A_2, A_3) \\ &= 0.6085 + 0.1880 = 0.7965 < \delta_\cup(A_1, A_3) = 1.2498 \\ &\delta_L(A_1, A_2) + \delta_L(A_2, A_3) \\ &= 0.2902 + 0.2742 = 0.5644 < \delta_L(A_1, A_3) = 0.6489\end{aligned}$$

Note that d_E also cannot satisfy the triangle inequality, since

$$\begin{aligned}&\delta_E(A_1, A_2) + \delta_E(A_2, A_4) \\ &= 1.8384 + 0.5394 = 2.3778 < \delta_E(A_1, A_4) = 4.1896\end{aligned}$$

Therefore, all the existing focal distance definitions are not strict distance metrics.

3) On mathematical rigorousness of definition: $A_i \cap A_j = \emptyset$.

That is, $|A_i \cap A_j| = 0$, then $\delta_E(A_i, A_j)$ cannot be calculated (because of a division by zero). We can also say that it tends to infinity; however, this is not reasonable because the value of distance is dominated by the relationship between focal sets' composition.

In this case, δ_L, $\delta_\cap$, and $\delta_\cup$ all can be calculated, and will not be dominated by the relationship between focal elements' compositions.

We further analyze the properties of the different measures in terms of changing trend in some typical cases.

Suppose that m is a BBA defined on the FOD Θ where $|\Theta| = n$. To simplify the analysis, we assume that m only has two focal elements A_1 and A_2 with mass assignments

$m(A_1) = a$ and $m(A_2) = 1 - a$. We will analyze the four existing focal distance measures under different situations below.

1) Focal elements' relation: $A_1 \subset A_2$: In such a case, for δ_E, one gets

$$\begin{aligned}\delta_E(A_1, A_2) =& \frac{1}{|A_1 \cap A_2| \,/\, |A_1 \cup A_2|} \\ & \cdot [m(A_1) - m(A_2)] \log_2 \frac{m(A_1)}{m(A_2)} \\ =& \frac{|A_2|}{|A_1|} \cdot [m(A_1) - m(A_2)] \\ & \cdot [\log_2(m(A_1)) - \log_2(m(A_2))]\end{aligned} \tag{10}$$

As shown in Eq. (10), when the mass assignments are fixed, δ_E will become larger with the enlargement of the ratio between focal elements' cardinalities $|A_2|/|A_1|$. This is intuitive, because the difference between their focal elements' compositions is enlarged. If the ratio $|A_2|/|A_1|$ is fixed, δ_E will become smaller with the decrease of difference $|m(A_1) - m(A_2)|$. This also makes sense, since the two focal elements' difference becomes smaller in terms of mass assignments.

For $\delta_\cap$, one gets

$$\begin{aligned}\delta_\cap(A_1, A_2) &= m(A_1) \cdot |A_1| + m(A_2) \cdot |A_2| \\ &\quad - [m(A_1) + m(A_2)] \cdot |A_1 \cap A_2| \\ &= m(A_1) \cdot |A_1| + m(A_2) \cdot |A_2| \\ &\quad - [m(A_1) + m(A_2)] \cdot |A_1| \\ &= (1 - m(A_1))(|A_2| - |A_1|)\end{aligned} \tag{11}$$

As shown in Eq. (11), if $m(A_1)$ is fixed, $\delta_\cap$ will become larger with the enlargement of the difference between focal elements' cardinalities $|A_2| - |A_1|$. This makes sense. If the difference of cardinalities i.e., $|A_2| - |A_1|$ is fixed, $\delta_\cap$ will become smaller with the increase of mass assignment of A_1 (which is contained by A_2).

For $\delta_\cup$, one gets

$$\begin{aligned}\delta_\cup(A_1, A_2) &= [m(A_1) + m(A_2)] \cdot |A_1 \cup A_2| \\ &\quad - m(A_1) \cdot |A_1| - m(A_2) \cdot |A_2| \\ &= [m(A_1) + m(A_2)] \cdot |A_2| \\ &\quad - m(A_1) \cdot |A_1| - m(A_2) \cdot |A_2| \\ &= m(A_1) \cdot (|A_2| - |A_1|)\end{aligned} \tag{12}$$

As shown in Eq. (12), if $m(A_1)$ is fixed, $\delta_\cup$ will become larger with the enlargement of the difference between focal elements' cardinalities $|A_2| - |A_1|$. This makes sense. If the difference of cardinalities i.e., $|A_2| - |A_1|$ is fixed, $\delta_\cup$ will become lager with the increase of mass assignment of A_1 (which is contained by A_2). That is, when $A_1 \subset A_2$ and $|A_2| - |A_1|$ is fixed, $\delta_\cup$ is positively correlated to the mass of focal element with smaller cardinality (A_1), while $\delta_\cap$ is positively correlated to the mass of focal element with larger cardinality (A_2).

For δ_L, one gets

$$\begin{aligned}\delta_L(A_1, A_2) &= \left(1 - \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|}\right) \cdot \frac{e^{|m(A_1) - m(A_2)|}}{e} \\ &= (1 - |A_1|/|A_2|) \cdot \frac{e^{|m(A_1) - m(A_2)|}}{e}\end{aligned} \tag{13}$$

As shown in Eq. (13), when the mass assignments are fixed, δ_L will become larger with the enlargement of the ratio between focal elements' cardinalities $|A_2|/|A_1|$. This makes sense, because the difference between their focal elements' compositions is enlarged. If the ratio $|A_2|/|A_1|$ is fixed, δ_L will become smaller with the decrease of difference $|m(A_1) - m(A_2)|$. This also makes sense, since the two focal elements' difference becomes smaller in terms of mass assignments.

The analyses above are supported by the Example 1 below.

Example 1

Suppose that the FOD is $\Theta = \{\theta_1, ..., \theta_5\}$. Four BBAs are defined and each has two focal elements as listed in Table I.

TABLE I
FOUR BBAS IN EXAMPLE 1.

BBA	A_1	A_2
m_1	$\{\theta_1\}$	Θ
m_2	$\{\theta_1, \theta_2\}$	Θ
m_3	$\{\theta_1, \theta_2, \theta_3\}$	Θ
m_4	$\{\theta_1, \theta_2, \theta_3, \theta_4\}$	Θ

For each BBA, the mass assignment of A_1 changes from 0.01 to 0.99 with an increase of 0.01 at each step. The values of different focal distance measures are shown in Fig .1 and Fig. 2, where the changing trend of different focal distance measures are accordant to the above analyses.

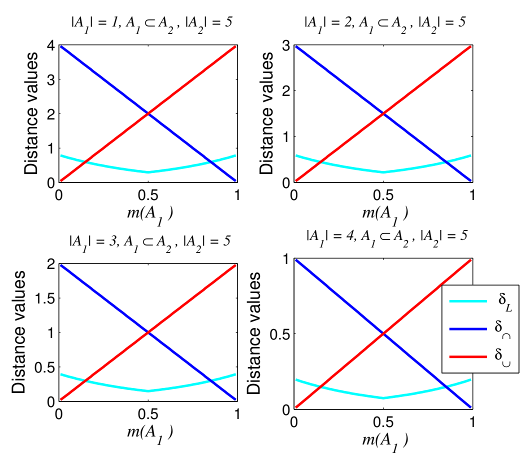

Fig. 1. Comparisons of $\delta_L, \delta_\cap, \delta_\cup$ in Example 1.

As shown in Figs. 1 and 2, $\delta_\cup$ is positively proportional to the mass of focal element with smaller cardinality ($|A_1|$), while $\delta_\cap$ is positively proportional to the mass of focal element with larger cardinality ($|A_2|$). Given a fixed $m(A_1)$, with the increase of $|A_1|$, i.e., the decrease of $|A_2| - |A_1|$, both $\delta_\cap$ and

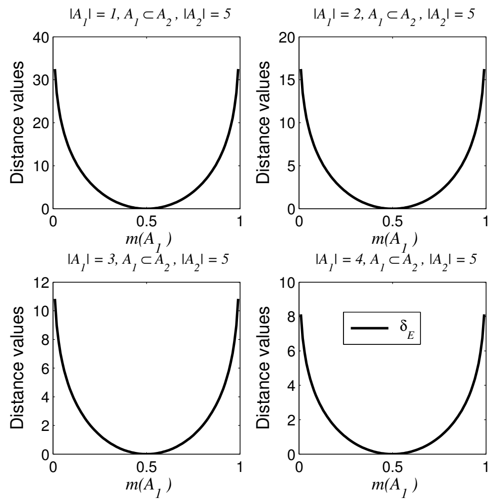

Fig. 2. δ_E in Example 1.

$\delta_\cup$ become smaller. Given a fixed $m(A_1)$, with the increase of A_1, i.e., the decrease of $|A_2|/|A_1|$, δ_E and δ_L become smaller. Given a fixed $|A_1|$, when $m(A_1)$ tends to $m(A_2)$, δ_E and δ_L become smaller. Such behaviors of δ_L and δ_E are more intuitive than that of $\delta_\cap$ and $\delta_\cup$.

2) Focal elements' relation: $A_1 \cap A_2 = \emptyset$: In such a case, for δ_E, one gets

$$\begin{aligned}\delta_E(A_1, A_2) =&\frac{1}{0} \cdot [m(A_1) - m(A_2)] \log_2 \frac{m(A_1)}{m(A_2)} \\ =&\infty\end{aligned} \tag{14}$$

For $\delta_\cap$, one gets

$$\begin{aligned}\delta_\cap(A_1, A_2) &= m(A_1) \cdot |A_1| + m(A_2) \cdot |A_2| \\ &- [m(A_1) + m(A_2)] \cdot |A_1 \cap A_2| \\ &= m(A_1) \cdot |A_1| + m(A_2) \cdot |A_2| \\ &= m(A_1) \cdot (|A_1| - |A_2|) + |A_2|\end{aligned} \tag{15}$$

For $\delta_\cup$, one gets

$$\begin{aligned}\delta_\cup(A_1, A_2) &= [m(A_1) + m(A_2)] \cdot |A_1 \cup A_2| \\ &- m(A_1) \cdot |A_1| - m(A_2) \cdot |A_2| \\ &= [m(A_1) + m(A_2)] \cdot (|A_1| + |A_2|) \\ &- m(A_1) \cdot |A_1| - m(A_2) \cdot |A_2| \\ &= m(A_1) \cdot |A_2| + m(A_2)\,|A_1| \\ &= m(A_1)\,(|A_2| - |A_1|) + |A_1|\end{aligned} \tag{16}$$

It can be seen that when $A_1 \cap A_2 = \emptyset$, if $|A_1|$ is closer to $|A_2|$, both $\delta_\cap$ and $\delta_\cup$ are smaller[1].

For δ_L, one gets

$$\begin{aligned}\delta_L(A_1, A_2) = (1-0) \cdot \frac{e^{|m(A_1)-m(A_2)|}}{e} \\ = \frac{e^{|m(A_1)-m(A_2)|}}{e}\end{aligned} \tag{17}$$

[1]This means that $\{\theta_1\}$ is farther from $\{\theta_2, \theta_3\}$ than from $\{\theta_2\}$, which makes some sense.

It can be seen that when $A_1 \cap A_2 = \emptyset$, δ_L is only determined by the difference between $m(A_1)$ and $m(A_2)$, which is not a satisfactory result.

The analyses above can be supported by the Example 2 below.

Example 2

Suppose that the FOD is $\Theta = \{\theta_1, ..., \theta_5\}$. Three BBAs are defined and each has two focal elements as listed in Table II.

TABLE II
THREE BBAS IN EXAMPLE 2.

BBA	A_1	A_2
m_1	$\{\theta_1\}$	$\{\theta_2\}$
m_2	$\{\theta_1\}$	$\{\theta_2, \theta_3\}$
m_3	$\{\theta_1\}$	$\{\theta_2, \theta_3, \theta_4\}$
m_4	$\{\theta_1\}$	$\{\theta_2, \theta_3, \theta_4, \theta_5\}$

In each BBA, the two focal elements have an empty intersection. For each BBA, the mass assignment of A_1 change from 0.01 to 0.99 with an increase of 0.01 at each step. The values of $\delta_\cap$ and $\delta_\cup$ are shown in Fig. 3. As we can see,

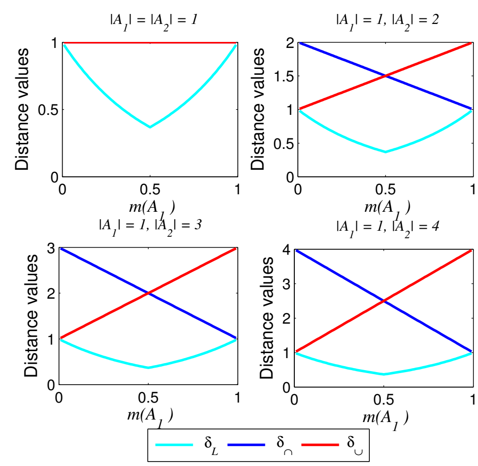

Fig. 3. Comparisons of $\delta_L, \delta_\cap, \delta_\cup$ in Example 2.

when $|A_1| = |A_2|$, both $\delta_\cap$ and $\delta_\cup$ remain unchanged. Given a fixed $m(A_1)$, when the difference $|A_2| - |A_1|$ becomes larger, both $\delta_\cap$ and $\delta_\cup$ become large. When the $|A_2| - |A_1|$ is fixed, $\delta_\cup$ is proportional to the mass of focal element with smaller cardinality (A_1), while $\delta_\cap$ is proportional to the mass of the focal element with larger cardinality (A_2), i.e., inversely proportional to the mass of the focal element with smaller cardinality ($|A_1|$). When $|A_2| - |A_1|$ is fixed and $m(A_1)$ tends to $m(A_2)$, δ_L becomes small and when the difference between $m(A_1)$ and $m(A_2)$ becomes larger, δ_L becomes larger, which is more intuitive than the behaviors of $\delta_\cap$ and $\delta_\cup$.

As we can see in Fig. 3, the changing trend of δ_L is not affected by the change of $|A_2|$, which is not a satisfactory behavior. This is due to the Jaccard's coefficient [14] used in δ_L. Since in Example 2, $|A_1 \cap A_2| \,/\, |A_1 \cup A_2| = 0$, no matter how A_2 changes.

3) Focal elements' relation: $A_1 \cap A_2 \neq \emptyset$: Here A_1 cannot be contained by A_2, and A_2 cannot be contained by A_1.

We provide an example to show $\delta_\cap$ and $\delta_\cup$'s behaviors in this situation.

Example 3

Suppose that the FOD is $\Theta = \{\theta_1, ..., \theta_6\}$. Four BBAs are defined and each has two focal element as listed in Table III. For each BBA, the mass assignment of A_1 changes from 0.01 to 0.99 with an increase of 0.01 at each step. The values of $\delta_\cap$, $\delta_\cup$ and δ_L are shown in Fig. 4. The changing trend of δ_E is as shown in Fig. 5. As shown in Fig. 4, when $|A_1| = |A_2|$, both $\delta_\cap$ and $\delta_\cup$ equal to 1, and remain unchanged. This is because when $|A_1| = |A_2| = 2$, according to Eq. (7) and Eq. (8), one has

TABLE III
FOUR BBAS IN EXAMPLE 3.

BBA	A_1	A_2
m_1	$\{\theta_1, \theta_2\}$	$\{\theta_2, \theta_3\}$
m_2	$\{\theta_1, \theta_2\}$	$\{\theta_2, \theta_3, \theta_4\}$
m_3	$\{\theta_1, \theta_2\}$	$\{\theta_2, \theta_3, \theta_4, \theta_5\}$
m_4	$\{\theta_1, \theta_2\}$	$\{\theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$

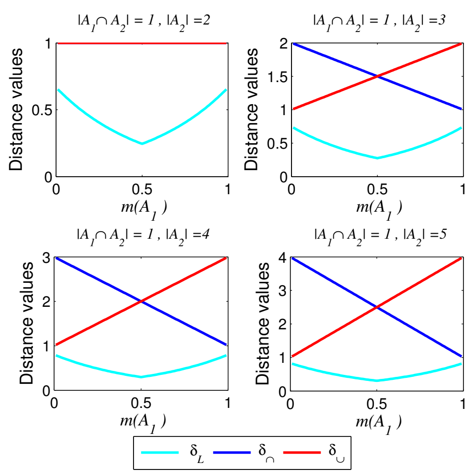

Fig. 4. Comparisons of $\delta_L, \delta_\cap, \delta_\cup$ in Example 3.

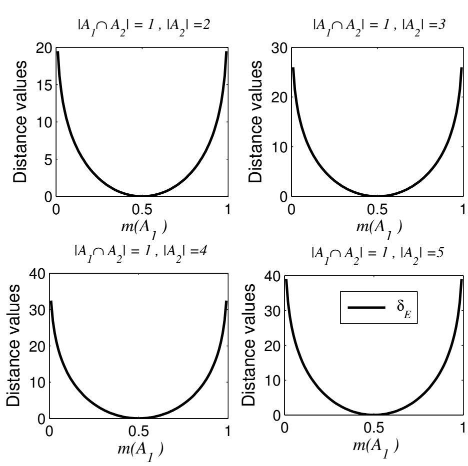

Fig. 5. Comparisons of δ_E in Example 3.

$$\delta_\cup(A_1, A_2) = |A_1 \cup A_2| - |A_2|$$

and

$$\delta_\cap(A_1, A_2) = |A_2| - |A_1 \cap A_2|.$$

Therefore, $\delta_\cap$ and $\delta_\cup$ = 1.

Given a fixed $m(A_1)$, when the difference $|A_2| - |A_1|$ becomes larger, both δ_L, $\delta_\cap$ and $\delta_\cup$ become larger as shown in Fig. 4. This makes sense, because the uncommon part of A_1 and A_2 becomes large. When the difference $|A_2| - |A_1|$ is fixed, $\delta_\cup$ is proportional to the mass of focal element with smaller cardinality (A_1), while $\delta_\cap$ is proportional to the mass of the focal element with larger cardinality (A_2), i.e., inversely proportional to the mass of the focal element with smaller cardinality (A_1). When $|A_2| - |A_1|$ is fixed, δ_L becomes smaller when $m(A_1)$ tends to $m(A_2)$ and becomes larger when the difference between $m(A_1)$ and $m(A_2)$ becomes larger. Such a behavior of δ_L is more intuitive than that of $\delta_\cap$ and $\delta_\cup$.

As shown in Fig. 5, when the difference $|A_2|-|A_1|$ becomes larger, δ_E becomes larger. When the difference $|A_2| - |A_1|$ is fixed, δ_E becomes smaller when $m(A_1)$ tends to $m(A_2)$ and becomes larger when the difference between $m(A_1)$ and $m(A_2)$ becomes larger.

In summary, δ_E is the usually not a good choice for measuring the distance between two focal elements. δ_L behaves more rational than the other available measures, although it still has some limitations caused by the Jaccard coefficient used (i.e., it can not well reflect[2] the difference about focal compositions as shown in Example 2).

In the next section, we further compare the available focal distance measures based on applications.

IV. APPLICATION-ORIENTED COMPARISONS OF FOCAL DISTANCE MEASURES

The focal distance can be used in many belief functions related applications such as the uncertainty representation and the BBA approximations. In this section, we introduce and analyze these two types of applications of focal distance and compare different focal distance measures based on the two types of applications.

A. Uncertainty degree based on focal distance

In Erkmen's work [10], the focal divergence of a BBA is defined by using the focal distance δ_E as

$$div_E(m) = \sum_{i=1}^{M}\sum_{j=1}^{M} \delta_E(A_i, A_j) \qquad (18)$$

[2]This can be suppressed by replacing Jaccard's coefficient with other similarity coefficient for focal element composition [14]. This is because that Jaccard coefficient only uses the $|A \cap B|$ and $|A \cup B|$, while other more informative type of coefficient also consider the factor like $|\Theta - (A \cup B)|$. When we consider this factor, it is possible to discriminate those indiscriminate cases when using δ_L as shown in Example 2.

where M is the number of focal elements in a BBA. The focal divergence of a BBA can be considered as a measure of the degree of uncertainty.

According to the definition format in Eq. (18), one can also use δ_L, $\delta_\cap$ and $\delta_\cup$ to define the focal divergence as follows.

$$div_L(m) = \sum_{i=1}^{M}\sum_{j=1}^{M}\delta_L(A_i, A_j) \tag{19}$$

$$div_\cap(m) = \sum_{i=1}^{M}\sum_{j=1}^{M}\delta_\cap(A_i, A_j) \tag{20}$$

$$div_\cup(m) = \sum_{i=1}^{M}\sum_{j=1}^{M}\delta_\cup(A_i, A_j) \tag{21}$$

Suppose that the FOD $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Here, the following seven BBAs defined on Θ cited from [12] are used to compare different divergence definitions.

a) $m_1(\{\theta_1\}) = 1$.
As a result, one gets

$$div_E(m_1) = 0, \quad div_L(m_1) = 0.$$
$$div_\cap(m_1) = 0, \quad div_\cup(m_1) = 0.$$

This is a categorical BBA, i.e., there is only one focal element. All the divergence measures equal to 0.

b) $m_2(\{\theta_1\}) = 0.2, m_2(\{\theta_2\}) = 0.3, m_2(\{\theta_3\}) = 0.5$.
As a result, one gets

$$div_E(m_2) = \infty, \quad div_L(m_2) = 2.7050.$$
$$div_\cap(m_2) = 4.0, \quad div_\cup(m_2) = 4.0.$$

Since m_2 has only singleton focal elements, Jaccard coefficient used in Eq. (10) is zero. Therefore, there exists $1/0$ when calculating δ_E and thus $div_E(m_2) = \infty$.

c) $m_3(\{\theta_1\}) = 1/3, m_3(\{\theta_2\}) = 1/3, m_3(\{\theta_3\}) = 1/3$.
As a result, one gets

$$div_E(m_3) = NaN, \quad div_L(m_3) = 2.2073.$$
$$div_\cap(m_3) = 4.0, \quad div_\cup(m_3) = 4.0.$$

Since m_3 has only singleton focal elements, Jaccard coefficient used in Eq. (10) is zero. Furthermore, all the mass assignments are equal for all singletons. Therefore, there exists $0/0$ when calculating δ_E and thus $div_E(m_3) = NaN$, where NaN means "Not a Number". Although m_2 and m_3 have the same focal elements, their corresponding mass assignments are different. However, $div_\cap$ and $div_\cup$ are the same for m_2 and m_3. div_L can discriminate m_2 and m_3.

d) $m_4(\{\theta_1\}) = 0.2, m_4(\{\theta_1, \theta_2\}) = 0.3, m_4(\Theta) = 0.5$.
As a result, one gets

$$div_E(m_4) = 3.0556, \quad div_L(m_4) = 1.3682.$$
$$div_\cap(m_4) = 3.6, \quad div_\cup(m_4) = 1.8.$$

e) $m_5(\{\theta_1\}) = 1/3, m_5(\{\theta_1, \theta_2\}) = 1/3, m_5(\Theta) = 1/3$.
As a result, one gets

$$div_E(m_5) = 0, \quad div_L(m_5) = 1.1036.$$
$$div_\cap(m_5) = 2.6667, \quad div_\cup(m_5) = 2.6667.$$

Since all the focal elements of m_5 have equal mass assignments and the Jaccard coefficients are not 0, $div_E(m_5)$ is 0. Although the focal elements compositions are different, the focal distance is totally determined by the mass assignments in this case. This is unsatisfactory.

f) $m_6(\{\theta_1\}) = 0.9, m_6(\{\theta_2\}) = 0.1$.
As a result, one gets

$$div_E(m_6) = \infty, \quad div_L(m_6) = 1.6375.$$
$$div_\cap(m_6) = 2.0, \quad div_\cup(m_6) = 2.0.$$

Since m_6 only has singleton focal elements, Jaccard's coefficient used in Eq. (10) is zero. Therefore, there exists $1/0$ when calculating δ_E and thus $div_E(m_6) = \infty$.

g) $m_7(\{\theta_1\}) = 0.5, m_7(\{\theta_2\}) = 0.5$.
As a result, one gets

$$div_E(m_7) = NaN, \quad div_L(m_7) = 0.7358.$$
$$div_\cap(m_7) = 2.0, \quad div_\cup(m_7) = 2.0.$$

Since m_7 only has singleton focal elements, Jaccard's coefficient used in Eq. (10) is zero. Furthermore, all the mass assignments are equal for all singletons. Therefore, there exists $0/0$ when calculating δ_E and thus $div_E(m_7) = NaN$. Although m_6 and m_7 have the same focal elements, their corresponding mass assignments are different. However, $div_\cap$ and $div_\cup$ are the same for m_6 and m_7. div_L can discriminate m_6 and m_7.

In summary, δ_L is relatively more rational when compared with other measures, since it has better discrimination capability and can avoid the value of ∞ and NaN.

B. Application of focal distance in BBA approximation

In our previous work [13], Denœux's focal distance measures [11] including $\delta_\cap$ and $\delta_\cup$ are used to define the degree of non-redundancy of focal elements as follows. Suppose that there are M focal elements in a BBA.

$$\mathrm{nRd}_\cap(A_i) \triangleq \frac{1}{M-1}\sum_{j=1, j\neq i}^{M}\delta_\cap(A_i, A_j) \tag{22}$$

$$\mathrm{nRd}_\cup(A_i) \triangleq \frac{1}{M-1}\sum_{j=1, j\neq i}^{M}\delta_\cup(A_i, A_j) \tag{23}$$

Based on the degree non-redundancy, more redundant focal elements can be removed at first and then the BBA approximation can be accomplished.

According to the definition format in Eqs. (22) and (23), one can also use δ_L and δ_E, respectively, to define the degree of non-redundancy as follows.

$$\mathrm{nRd}_\mathrm{L}(A_i) \triangleq \frac{1}{M-1}\sum_{j=1, j\neq i}^{M}\delta_L(A_i, A_j) \tag{24}$$

$$\text{nRd}_{\text{E}}\left(A_i\right) \triangleq \frac{1}{M-1} \sum_{j=1,,j \neq i}^{M} \delta_E\left(A_i, A_j\right) \tag{25}$$

By using any non-redundancy definition aforementioned, we can implement the BBA approximation. Let m denote the original BBA to approximate with M focal elements. In the approximation, $k < M$ denotes the number of focal elements kept. The BBA approximation procedure is as follows.

- **Step 1:** For each $A_i, i = 1, ..., M$, compute its non-redundancy value $\text{nRd}(A_i)$ (could be $\text{nRd}_{\text{L}}(A_i)$, $\text{nRd}_{\text{E}}(A_i)$, $\text{nRd}_{\cap}(A_i)$, or $\text{nRd}_{\cup}(A_i)$);
- **Step 2:** Sort all the elements in descending order according to the values of $\text{nRd}(A_i)$;
- **Step 3:** Remove the $M - k$ bottom focal elements;
- **Setp 4:** Normalize the mass values of the remaining k focal elements and output the approximated BBA.

Here, we use $\text{nRd}_{\text{L}}(A_i)$, $\text{nRd}_{\text{E}}(A_i)$, $\text{nRd}_{\cap}(A_i)$, and $\text{nRd}_{\cup}(A_i)$, respectively, to do the BBA approximation. We compare and evaluate different BBA approximations in terms of the averaging computational cost of BBAs' combination and the closeness to the original one. A BBA transformation with less computational cost of BBAs' combination and more closeness is preferred. To measure the closeness or the dissimilarity between different BBAs, Jousselme's distance measure [15] between BBA is used as follows:

$$d_J(m_1, m_2) \triangleq \sqrt{\frac{1}{2} \cdot (m_1 - m_2)^T \mathbf{Jac}\ (m_1 - m_2)} \tag{26}$$

where $\mathbf{Jac}$ is the so-called Jaccard's weighting matrix, whose elements are the Jaccard coefficients of focal element pairs.

The belief interval based distance [16] between BBAs is defined as

$$d_{BI}^E(m_1, m_2) \triangleq \sqrt{N_c \cdot \sum\nolimits_{i=1}^{2^n-1} \left[d_{BI}(BI_1(A_i), BI_2(A_i))\right]^2} \tag{27}$$

Here $N_c = 1/2^{n-1}$ is the normalization factor, and $BI_1(A_i) = [Bel_1(A_i), Pl_1(A_i)]$ and $BI_2(A_i) = [Bel_2(A_i), Pl_2(A_i)]$ are the belief interval of the focal element A_i for m_1 and m_2, respectively. The $d_{BI}(\cdot, \cdot)$ in Eq. (27) is the Wasserstein's distance between belief intervals $BI_1(A_i) = [a_1, b_1]$ and $BI_2(A_i) = [a_2, b_2]$ as shown in Eq. (28)

$$\begin{aligned} &d_{BI}\left([a_1, b_1], [a_2, b_2]\right) \\ &\triangleq \sqrt{\left[\tfrac{a_1+b_1}{2} - \tfrac{a_2+b_2}{2}\right]^2 + \tfrac{1}{3}\left[\tfrac{b_1-a_1}{2} - \tfrac{b_2-a_2}{2}\right]^2} \end{aligned} \tag{28}$$

Other types of distance of evidence [17] could also be used here to evaluate the closeness between BBAs. Here we choose Jousselme's distance and our proposed belief interval based distance since they are both strict distance metrics [17], [16].

Our comparative analysis is based on a Monte Carlo simulation using $M = 200$ random runs. In j-th simulation run, the BBA to approximate m^j is randomly generated and the different approximation results are obtained using the approximation with different non-redundancy definitions. We calculate the computational time of the original evidence combination of $m^j \oplus m^j$ with Dempster's rule, and the computation time of Dempster's combination of each approximated BBA.

In our simulations, the cardinality of the FoD Θ is chosen to 4. In each random generation, there are 15 focal elements (emptyset $\emptyset$ is excluded) in the original BBA. The remaining number of focal elements for all the approaches used here is from 14 down to 2 (decreased by 1 for each time). Random generation of BBA is based on Algorithm 1 [18] in Table IV.

TABLE IV
ALGORITHM 1: RANDOM BBA GENERATION - UNIFORM SAMPLING FROM ALL FOCAL ELEMENTS.

Input: Θ: Frame of discernment;
N_{max}: Maximum number of focal elements
Output: **Output**: m: BBA
Generate $\mathcal{P}(\Theta)$, which is the power set of Θ;
FOReach $1 \leq i \leq |\mathcal{P}(\Theta)|$ do
Generate a value according to the Gamma distribution $\mathcal{G}(1, 1) \rightarrow m_i$,
END
Normalize the vector $m = [m_1, ..., m_{|\mathcal{P}(\Theta)|}] \rightarrow m'$;
$m(A_i) = m'_i$;

The averaging computational time of combination and the averaging distance values over 200 runs between the original BBA and the approximated BBAs obtained using different non-redundancy definitions given different remaining focal elements' numbers are shown in Figs. 6 and 7, respectively.

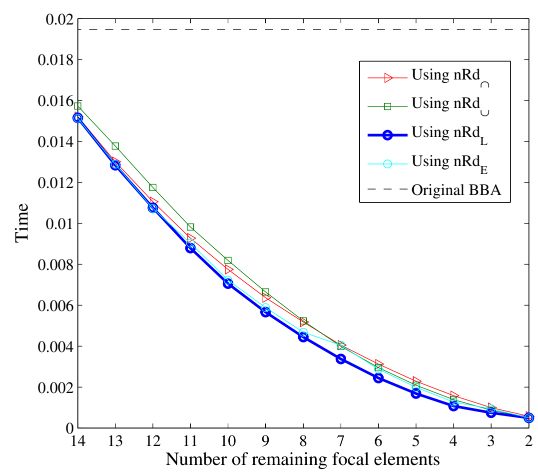

Fig. 6. Computation time comparisons.

As we can see in Fig. 6, in terms of combination time, nRd_L and nRd_E perform better than $\text{nRd}_{\cap}$ and $\text{nRd}_{\cup}$. When using nRd_L, the combination time is the smallest, which means the most reduction in complexity.

As we can see in Figs. 7 and 8, in terms of the closeness (loss of information), $\text{nRd}_{\cap}$ and $\text{nRd}_{\cup}$ perform better than nRd_L and nRd_E. When using nRd_E, the distance is the largest; when using $\text{nRd}_{\cap}$, the distance is the smallest, which means the least information loss in the BBA approximation.

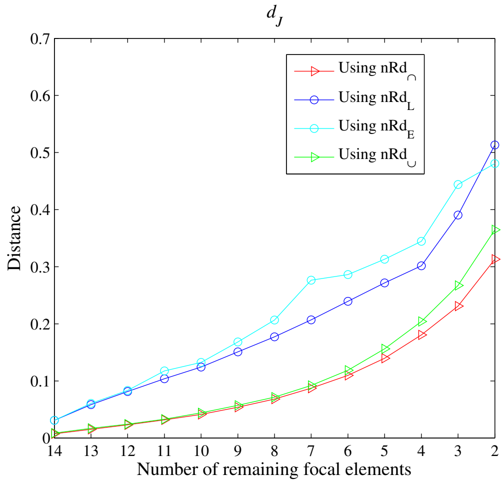

Fig. 7. Closeness comparisons using d_J.

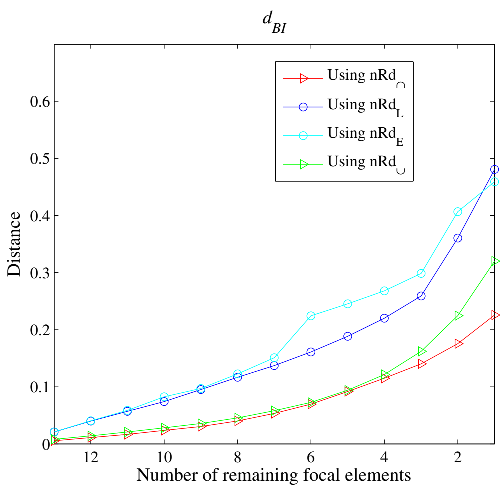

Fig. 8. Closeness comparisons using d_{BI}.

In summary, different focal distance measures can be applied to the applications like BBA approximations. δ_L performs well in terms of complexity reduction while $\delta_\cap$ performs well in terms of the loss of information.

V. Conclusion

In this paper, an important concept, i.e., the focal distance has been reviewed. Existing focal distance measures are compared based on experiments and related analyses. Comparisons based on applications including the uncertainty representation and BBA approximations are also provided. In conclusion, all existing focal distance measures are not strict distance metrics and have their own pros and cons, where δ_L is relatively a good compromise.

In future work, we will try to propose more rational, especially strict focal distance definitions, e.g., using more comprehensive similarity coefficient to replace the Jaccard coefficient to define focal distance. Furthermore, we aim to find more related applications for the focal distance.

Acknowledgment

This work was supported by the Grant for State Key Program for Basic Research of China (973) (No. 2013CB329405), National Natural Science Foundation (No. 61573275), Science and technology project of Shaanxi Province (No. 2013KJXX-46), and Fundamental Research Funds for the Central Universities (Nos. xjj2016066, xjj2014122).

References

[1] G. Shafer, *A Mathematical Theory of Evidence*. Princeton: Princeton university press, 1976.
[2] G. Lin, J. Liang, and Y. Qian, "An information fusion approach by combining multigranulation rough sets and evidence theory," *Information Sciences*, vol. 314, pp. 184–199, 2015.
[3] Z. Liu, Q. Pan, and J. Dezert, "Evidential classifier for imprecise data based on belief functions," *Knowledge-Based Systems*, vol. 52, pp. 246–257, Nov. 2013.
[4] D. Han, J. Dezert, J.-M. Tacnet, and C. Han, "A fuzzy-cautious OWA approach with evidential reasoning," in *Proceedings of the 15th International Conference on Information Fusion*, Singapore, July 2012, pp. 278–285.
[5] F. Voorbraak, "On the justification of Dempster's rule of combination," *Artificial Intelligence*, vol. 48, no. 2, pp. 171–197, 1991.
[6] P. Wang, "A defect in Dempster-Shafer theory," in *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence*, Seattle, USA, July 1994, pp. 560–566.
[7] J. Dezert and A. Tchamova, "On the validity of Dempster's fusion rule and its interpretation as a generalization of bayesian fusion rule," *International Journal of Intelligent Systems*, vol. 29, no. 3, pp. 223–252, 2014.
[8] P. Smets and R. Kennes, "The transferable belief model," *Artificial intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
[9] F. Smarandache and J. Dezert, Eds., *Applications and Advances of DSmT for Information Fusion (Vol III)*. Rehoboth, DE, USA: Amer. Res. Press, 2009.
[10] A. M. Erkmen and H. E. Stephanou, "Information fractals for evidential pattern classification," *IEEE Transactions on Systems Man and Cybernetics*, vol. 20, no. 5, pp. 1103–1114, 1990.
[11] T. Denœux, "Inner and outer approximation of belief structures using a hierarchical clustering approach," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 04, pp. 437–460, 2001.
[12] J. Liu, S. Huang, and Y. Deng, "A new focal divergence measure in belief function under dempster-shafer theory," *Journal of Information & Computational Science*, vol. 11, no. 3, pp. 971–977, 2014.
[13] D. Han, Y. Yang, and J. Dezert, "Two novel methods for BBA approximation based on focal element redundancy," in *Proceedings of the 18th International Conference on Information Fusion*, Washington DC, July 2015, pp. 428–434.
[14] J. Diaz, M. Rifqi, and B. Bouchon-Meunier, "A similarity measure between basic belief assignments," in *Proceedings of the 9th International Conference on Information Fusion*, Florence, Italy, July 2006, pp. 1–6.
[15] A.-L. Jousselme, D. Grenier, and É. Bossé, "A new distance between two bodies of evidence," *Information fusion*, vol. 2, no. 2, pp. 91–101, 2001.
[16] D. Han, J. Dezert, and Y. Yang, "New distance measures of evidence based on belief intervals," in *Belief Functions: Theory and Applications*. Springer, 2014, pp. 432–441.
[17] A.-L. Jousselme and P. Maupin, "Distances in evidence theory: Comprehensive survey and generalizations," *International Journal of Approximate Reasoning*, vol. 53, no. 2, pp. 118–145, 2012.
[18] T. Burger and S. Destercke, "Random generation of mass functions: A short howto," in *Proceedings of the 2nd International Conference on Belief functions*, Compiègne, France, May 2012, pp. 145–152.