



## A belief function distance metric for orderable sets



Zachary Sunberg<sup>1</sup>, Jonathan Rogers<sup>\*,2</sup>

Texas A&M University, College Station, TX 77843, United States

### ARTICLE INFO

#### Article history:

Received 26 March 2012  
 Received in revised form 13 March 2013  
 Accepted 18 March 2013  
 Available online 29 March 2013

#### Keywords:

Dempster–Shafer theory  
 Distance metric  
 Orderable sets  
 Sensor fault detection

### ABSTRACT

This paper describes a new metric for characterizing conflict between belief assignments. The new metric, specifically designed to quantify conflict on orderable sets, uses a Hausdorff-based measure to account for the distance between focal elements. This results in a distance metric that can accurately measure conflict between belief assignments without saturating simply because two assignments do not have common focal elements. The proposed metric is particularly attractive in sensor fusion applications in which belief is distributed on a continuous measurement space. Several example cases demonstrate the proposed metric's performance, and comparisons with other common measures of conflict show the significant benefit of using the proposed metric in cases where a sensor's error and noise characteristics are not known precisely *a priori*.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

Belief function theory provides an attractive framework for sensor fusion primarily due to its flexibility in combining a variety of types of information. The type of flexibility provided by Dempster–Shafer Theory [1], Fuzzy Set Theory [2], Plausibility Theory [3], and other derivatives is particularly relevant in the autonomous systems community, where low-cost sensors of varying types are often integrated to perform real-time state estimation, artificial intelligence tasks, or health monitoring. In many applications, sensor fault detection and isolation (FDI) plays a key role in assessing sensor accuracy and maintaining robustness of fused outputs. Typically, FDI processing occurs prior to a filter's use of sensor data to guard against use of inaccurate or problematic sensor feedback. Other artificial intelligence (AI) applications such as object recognition, material characterization, or quality control require quantification of disagreement between sensors. A major challenge in both FDI and AI system design lies in accurate detection and characterization of sensor disagreement, failure, or unreliability.

Numerous investigators have explored methods of characterizing sensor failure or disagreement in real-time systems. An example is simple outlier rejection commonly used in Kalman filtering algorithms [4], in which sensor data that produces measurement innovations beyond a certain bound are discarded. However, sensor fusion performed in the context of belief function theory often entails imprecise collection of data (such as nonsingleton sets)

from a wide variety of sensor types. In such cases, simple outlier rejection schemes may not be straightforward to implement. Several alternatives have been proposed for measuring the amount of conflict between basic probability assignments (BPAs), which can be used for instance to eliminate sensors that consistently show disagreement with a majority of other data sources.

The original conception of the conflict between two belief assignments is the so-called “internal conflict” proposed by Shafer [5,6]. This value is determined by the amount of belief mass that would be assigned to the empty set if the orthogonal sum were not normalized and measures the support for conflicting evidence between belief assignments. Numerous authors [7,8] have shown that this measure of conflict is not a particularly useful measure of sensor disagreement in a variety of scenarios, especially since the internal conflict of a BPA with itself is generally nonzero. More recently, Jusselme et al. [9] proposed a *distance metric* between belief assignments and later argued that the distance is a suitable quantifier of belief function conflict or disagreement [10]. Although this concept of distance differs from the traditional notion of conflict, Jusselme's metric has proven highly attractive in that it satisfies the mathematical constraints of a metric distance while providing intuitive outputs (as opposed to the sometimes confusing internal conflict). Since the introduction of Jusselme's metric, a variety of other distance metrics have been proposed that use a similar quadratic form but different similarity functions between focal elements [11,12]. Other metrics measure conflict on a transformed domain. For instance, in [13–15] metrics are proposed in which BPA's are first transformed to pignistic probabilities instead of operating directly in the Dempster–Shafer belief assignment domain. A comprehensive survey of various belief function distances proposed to date is provided in Ref. [16].

\* Corresponding author.

E-mail address: [jrogers8@gmail.com](mailto:jrogers8@gmail.com) (J. Rogers).

<sup>1</sup> Graduate Research Assistant, Department of Aerospace Engineering.

<sup>2</sup> Assistant Professor, Department of Aerospace Engineering.

The underlying assumption behind Jousselme's metric and other measures of distance in the Dempster–Shafer domain using a traditional similarity function is that the measurement space on which BPAs are built are “indiscernible and unorderable” [9] and thus the only distinguishing feature between focal elements is the cardinality of their union or intersection. However, when belief function theory is used to process real-time sensor measurements, this assumption is usually not only inaccurate but can lead to significant problems. Sensor measurements typically occur on a continuous measurement space, and thus the physical distance between focal elements of two BPAs should play a part in determining the total disagreement between them. If this physical distance is not accounted for, an FDI system may incorrectly label sensors as malfunctioning in cases where a sensor error distribution is not known precisely *a priori*. An example of belief function theory applied to a sensor fusion problem on a continuous (orderable) measurement space in which this difficulty is encountered is provided in Ref. [7]. While this paper is not concerned with the FDI problem in itself, difficulty in performing FDI with inaccurate sensor conflict information provides a primary motivation behind this work.

This paper presents a new distance metric which mirrors the quadratic form proposed by Jousselme, but uses a similarity function derived from the Hausdorff distance. The metric is attractive in practical sensing applications due to its robustness to uncertainty in measurement noise and error parameters. Another attractive feature is the inclusion of a gain that can be tuned appropriately based on the nature of *a priori* knowledge of sensor errors. The paper begins with an overview of belief function theory and a description of the metric. The proposed measure's classification as a metric distance is addressed, and example cases are provided showing the performance of the new metric. Comparisons with other metrics are performed in both simulated and experimental settings, demonstrating that the proposed metric is superior in most practical problems where sensor error characteristics may not be known precisely. Note that while comparisons with distance metrics in transformed domains (for instance, pignistic probabilities) may also be performed, this paper considers only distance metrics that operate directly on BPAs within the original Dempster–Shafer framework. A primary drawback of pignistic probability-based distances is that such distances may produce a value of zero when two BPAs are different but have the same pignistic probabilities, and thus they may not yield an accurate answer in the Dempster–Shafer domain.

## 2. Background

### 2.1. Belief function theory

Belief function theory is a framework for data fusion given uncertain or imprecise pieces of data. The theory assigns so-called belief mass to different propositions based on evidential support, and, unlike in classical Bayesian probability, support can be assigned to both single propositions and general nonsingleton sets. Given a set of  $N$  mutually exclusive basic propositions, the *frame of discernment*,  $\Theta$  is given by

$$\Theta = \{1, 2, 3, \dots, N\}. \quad (1)$$

Each sensor can distribute its evidential weight over the power set  $2^\Theta$  that contains all possible subsets contained in  $\Theta$ , given by

$$2^\Theta = \{\emptyset, 1, \dots, N, \{1, 2\}, \{1, 3\}, \dots, \{N-1, N\}, \dots, \Theta\}. \quad (2)$$

Note that  $2^\Theta$  is an exhaustive set that contains  $2^N$  elements. A sensor distributes belief mass over  $2^\Theta$  to create a basic probability assignment (BPA)  $m$  such that

$$\sum_{A \in 2^\Theta} m(A) = 1, \quad (3)$$

where  $m(A)$  is the belief mass assigned to element  $A$ . Note that the elements of  $2^\Theta$  to which the sensor assigns nonzero belief are termed *focal elements*.

Belief function theory is distinguished from classical Bayesian probability theory in that belief (or probability in Bayesian terms) can be distributed not only to singleton sets, but to a nonsingleton group of propositions. Here inlies the flexibility of belief function theory – a sensor that can select a group of propositions, but cannot distinguish between them, can still provide meaningful quantitative data by assigning belief to a nonsingleton set. This belief can then be combined (using any of a number of combination rules) with other BPAs that may incorporate higher resolution data to generate a joint inference. Furthermore, belief function theory allows belief assignment to the ignorance element,  $\emptyset$ , representing a sensor's level of ignorance regarding the measurement. Bayesian theory does not offer a convenient representation for ignorance and can cause over-committal to individual propositions based on indistinct pieces of evidence. In applications, where FDI is important, the ability to assign belief to an ignorance element becomes crucial since, in the event of failure, a sensor's belief can be completely assigned to  $\emptyset$  with no disruptive effect on the fusion process. Such flexibility is necessary to avoid large-scale reconfiguration as sensors fluctuate between operational and non-operational modes.

The critical functionality of combining BPAs into a joint inference is performed by recursively applying a combination rule, allowing the fusion of data from an arbitrary number of contributing sensors. Dempster proposed the first rule of combination, the orthogonal rule, given by

$$(m_1 \oplus m_2)(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B)m_2(C)} \quad (4)$$

where  $m_1$  and  $m_2$  are BPAs. Numerous other authors have proposed additional combination rules that have proven more robust under various conditions compared to the orthogonal rule [17–19]. The fused BPA generally represents the optimal joint inference given all sensor data and can be used in a variety of ways within estimation or decision algorithms. Furthermore, in weighted combination rules such as that proposed in [19] conflict metrics can actually be used to compute weightings that may enable higher accuracy or robustness.

### 2.2. Belief function distance metric overview

The ability to quantify the amount of disagreement between belief functions has become important in many applications and extensions of belief function theory. Robust combination rules [19] as well as sensor diagnostic algorithms [7] have leveraged measures of sensor conflict to identify malfunctioning sensors or redistribute belief assignments. Jousselme et. al. proposed one of the most ubiquitous measures of disagreement in the form of a distance metric between two belief functions. The notion of disagreement as quantified by a “distance” differs significantly from the traditional notion of conflict proposed by Shafer. However this class of metrics offers a more intuitive and robust measure of disagreement for the class of FDI and AI sensor management problems alluded to earlier. Given BPAs  $m_1$  and  $m_2$ , Jousselme's distance is given by

$$d_J(m_1, m_2) = \sqrt{\frac{1}{2}(\underline{m}_1 - \underline{m}_2)^T D(\underline{m}_1 - \underline{m}_2)} \quad (5)$$

where  $\underline{m}_1$  and  $\underline{m}_2$  are “mass vectors” whose elements are the masses corresponding to each of the members of the combined

set of focal elements from both BPAs. The similarity matrix  $D$  is defined such that  $D_{ij} \in [0, 1]$  corresponds to the similarity between the  $i$ th and  $j$ th elements of the combined set of focal elements, with 1 meaning the elements are identical and 0 conveying a complete lack of similarity. Numerous alternatives for quantifying similarity between focal elements have been proposed [12]. In the case of Jousselme, focal element similarity is defined using the Jaccard coefficient according to

$$D_{J_{ij}} = \frac{|A_i \cap A_j|}{|A_i \cup A_j|}, \quad (6)$$

where  $|A|$  is the cardinality of  $A$ , and  $A_i$  is the  $i$ th member of the combined set of focal elements. As an example of its practical utility, Jousselme’s metric was used with some success in Ref. [7] to perform fault detection and isolation tasks on a set of sensors on-board an air vehicle.

Another similar metric of interest has been proposed by Fixsen and Mahler [11] and is termed a pseudo-distance due to its degeneracy. Fixsen and Mahler’s metric maintains the quadratic form structure of Jousselme’s distance but replaces the Jaccard coefficient with

$$D_{J_{ij}} = \frac{|A_i \cap A_j|}{|A_i||A_j|}. \quad (7)$$

Additional distances have been proposed that rely on pignistic probabilities associated with BPAs (see for instance Tessem’s distance [14]). However, to the best of the authors’ knowledge, all distance metrics proposed to date (that do not involve a prior transformation to pignistic probabilities) have been developed under the assumption that elements of the frame of discernment are both unorderable and indiscernable. As a result, the cardinality of unions and intersections between focal elements is the only measure available for judging similarity. However, when  $\Theta$  is comprised of orderable elements (such as when a continuous measurement space is discretized), similarity between elements needs not be judged exclusively on cardinality, and in such cases, distance metrics that measure similarity strictly through cardinality often produce counterintuitive results and may be of limited use in practical scenarios.

Consider the example of three sensors,  $A$ ,  $B_1$ , and  $B_2$  that generate BPAs corresponding to the value of a scalar  $\psi$  of arbitrary units. In this case, consider the frame of discernment to be a discretized set of numbers on  $\mathbb{R}$ . Sensor  $A$  may be viewed as a precision sensor, distributing belief over a set of singleton elements according to a Gaussian distribution. Sensors  $B_1$  and  $B_2$  may be viewed as nonprecision sensors, assigning most of their belief to one nonsingleton set. Now, consider the case as shown in Fig. 1, in which sensors  $A$  and  $B_1$  generally agree that  $\psi$  is located in the region  $2 \leq \psi \leq 5$

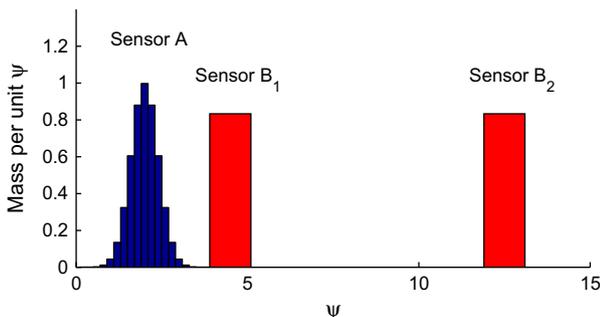


Fig. 1. A motivating example demonstrating counterintuitive behavior given measurements on an orderable space. The area of each block is proportional to the belief mass. Cardinality-based metrics will return the same value when comparing sensors  $A$  with  $B_1$  and  $A$  with  $B_2$ .

although strictly speaking their BPAs have no overlapping focal elements. Sensor  $B_2$  advocates a far different value closer to  $12 \leq \psi \leq 13$ , also with no focal elements overlapping focal elements in the other BPAs. In this case, use of Jousselme’s metric to compute the distance between BPAs from sensors  $A$  and  $B_1$  would yield the same value as that from sensors  $A$  and  $B_2$ , i.e.

$$d_j(A, B_1) = d_j(A, B_2). \quad (8)$$

While this is rigorously correct according to the traditional notion of conflict since none of the BPAs have common elements, it is not intuitive in that the physical proximity of the BPAs from  $A$  and  $B_1$  are not reflected, especially in comparison with their distance from  $B_2$ . Note that this nonintuitive property is true for any distance metric that bases similarity on set cardinality, since the cardinality operator cannot capture this physical distance relationship. In practice, this situation arises quite often when the width of BPA distributions may not be known precisely, so that noise or bias disturbances in sensors may cause BPAs to be in close proximity but non-overlapping. In such instances it is impossible to distinguish from sensors that are malfunctioning (i.e., in extreme disagreement from others) from sensors that simply exhibit higher noise or bias terms than was assumed during calibration. This motivational example highlights the need for alternative metrics designed specifically for orderable sets. The new metric proposed here deliberately moves away from a strict interpretation of conflict which holds that any two non-overlapping belief assignments are in complete and utter disagreement. Instead, it quantifies the disagreement in a more intuitive manner which assigns a gradually varying degree of disagreement based on the distance between the basic elements that make up the BPAs. This, as will be shown, results in a more robust way to measure disagreement for certain applications.

### 3. Description of proposed distance metric

#### 3.1. Metric definition

The new metric proposed here maintains the quadratic form structure of Jousselme but replaces the similarity function. The new metric  $d_H$  is defined as follows

$$d_H(m_1, m_2) = \sqrt{\frac{1}{2}(\underline{m}_1 - \underline{m}_2)^T D(\underline{m}_1 - \underline{m}_2)}, \quad (9)$$

$$\text{where } D_{ij} = S_H(A_i, A_j) = \frac{1}{1 + KH(A_i, A_j)}, \quad (10)$$

where  $H(A_i, A_j)$  is the Hausdorff distance between focal elements  $A_i$  and  $A_j$ .  $K > 0$  is a user-defined tuning parameter that adjusts metric response with respect to the orderable space discretization. The Hausdorff distance provides a simple method for quantifying the distance between sets  $A_i$  and  $A_j$ , and is one of the most widely used measures for quantifying such a distance. It is defined according to

$$H(A_i, A_j) = \max\{\sup_{b \in A_i} \inf_{c \in A_j} d(b, c), \sup_{c \in A_j} \inf_{b \in A_i} d(b, c)\}, \quad (11)$$

where  $d(x, y)$  is the distance between two elements of the sets and can be defined as any valid metric distance on the measurement space [20]. In the case where elements of the sets are real numbers, that is, the 1-dimensional Euclidean case, distance can be measured as the absolute value of the difference between the elements, and thus the Hausdorff distance may be defined as

$$H_{1-D}(A_i, A_j) = \max\{|\min(A_i) - \min(A_j)|, |\max(A_i) - \max(A_j)|\}. \quad (12)$$

Note that through the use of a general  $n$ -dimensional Euclidean norm the proposed distance metric can actually be defined for a frame of discernment whose elements are  $n$ -dimensional vectors.

Several examples depicting the Hausdorff distance for sets of integers are illustrated in Fig. 2 (note that this figure demonstrates the Hausdorff distance on sets of the real line, *not* the proposed belief function metric). Finally, note that when  $A_i$  and  $A_j$  are both singleton sets on the real line,  $H(A_i, A_j)$  is simply the absolute value of their difference.

Use of the Hausdorff distance rather than cardinality to characterize the similarity of focal elements is the primary feature that separates the new distance metric from previous distance definitions. Because of the use of the Hausdorff distance, the metric does not reach a saturated value when the two BPAs have no overlap. This is the primary distinguishing property of the new metric.

3.2. Metric properties

In order for the proposed metric  $d_H$  to be properly defined as a metric distance, the following criteria must be satisfied:

1. Non-negativity ( $d_H(L, M) \geq 0$ )
2. Non-degeneracy ( $d_H(L, M) = 0 \Leftrightarrow L = M$ )
3. Symmetry ( $d_H(L, M) = d_H(M, L)$ )
4. Subadditivity ( $d_H(L, M) \leq d_H(L, N) + d_H(N, M)$ )

where  $L, M$ , and  $N$  are any possible BPAs defined on an appropriate frame of discernment. Actually, the first criterion (non-negativity) is implied if the latter three are proven. However, discussion of the non-negativity of the metric is included here because a complete proof of the subadditivity criterion is not offered.

*Non-negativity.* The quadratic form of the proposed metric in (8) leads to the conclusion that non-negativity is satisfied as long as  $D$  is positive definite. Positive definiteness of the Jaccard similarity matrix originally proposed by Jusselme has only recently been proven [22]. The similarity matrix proposed in (9) shares certain properties with Jusselme’s similarity matrix, namely

$$D_{i,i} = 1 \quad \forall i, \tag{13}$$

$$0 < D_{i,j} < 1 \quad \forall i, j; i \neq j. \tag{14}$$

These conditions alone are not sufficient to prove  $D$  positive definite, and thus it is important to recognize the specific compatibility relationships between the elements of  $D$ . Consider the following  $n \times n$  similarity matrix, where  $n$  is the cardinality of the union of the two sets of focal elements of the BPAs. This matrix is computed from two valid BPAs according to Eq. (10)

$$D = \begin{bmatrix} 1 & D_{1,2} & \dots & \dots & D_{1,n} \\ D_{2,1} & 1 & & & \vdots \\ \vdots & & \ddots & & \\ & & & \boxed{D_{k,j}} & \\ & & & \vdots & 1 \\ \vdots & & D_{i,k} & \dots & \underline{D_{i,j}} & \ddots & \vdots \\ D_{n,1} & \dots & & & \dots & \dots & 1 \end{bmatrix}. \tag{15}$$

Here the frame of discernment is considered to be exhaustive, and composed of a discrete set of values from  $\mathbb{R}$  (real numbers). Each off-diagonal element,  $D_{ij}$ , (underlined) is constrained by  $n - 2$  specific pairs of other members (boxed) in row  $i$  and column  $j$ . Member  $D_{k,j}$  quantifies the (Hausdorff-based) closeness of the  $k$ th and  $j$ th elements, and  $D_{i,k}$  quantifies the closeness of the  $i$ th and  $k$ th elements. If the  $k$ th and  $j$ th elements are a given distance apart and the  $i$ th and  $k$ th elements are another given distance apart, then the  $i$ th and  $j$ th elements cannot physically be arbitrarily close to or far from one another. Thus, for each  $k \neq i, j$ , there is an implied compatibility constraint on  $D_{ij}$ . These geometric constraints can be written as follows:

$$\frac{1 - D_{ij}}{D_{ij}} \in \left[ \frac{|D_{i,k} - D_{k,j}|}{D_{i,k}D_{k,j}}, \frac{D_{i,k} + D_{k,j} - 2D_{i,k}D_{k,j}}{D_{i,k}D_{k,j}} \right] \forall k \in [1, n], k \neq j, k \neq i. \tag{16}$$

A  $3 \times 3$  similarity matrix constructed from (10)

$$A = \begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}, \tag{17}$$

therefore is constrained according to

$$\frac{1 - c}{c} \in \left[ \frac{|a - b|}{ab}, \frac{a + b - 2ab}{ab} \right]. \tag{18}$$

This reduces to the following inequality constraints:

$$ab - abc - c|a - b| \geq 0, \tag{19}$$

$$ac + bc - abc - ab \geq 0. \tag{20}$$

A proof of  $D$  positive definite for this  $3 \times 3$  case is offered in Appendix A. Positive definiteness of  $D$  for dimension higher than 3 is left as conjecture, and proving this remains open challenge that is beyond the scope of this work. In order to offer further evidence that the similarity matrix is positive definite, 100,000 similarity matrices were randomly generated according to the constraints in (16) and all were found to be positive definite.

*Non-degeneracy.* Given two identical BPAs  $A_1$  and  $A_2$  with mass distributions  $\underline{m}_1 = \underline{m}_2$ , the vector  $\underline{m}_1 - \underline{m}_2 = 0$  and thus  $d_H(A_1, A_2) = 0$ . Conversely, suppose  $d_H(A_1, A_2) = 0$ . This means that either  $\underline{m}_1 - \underline{m}_2 = 0$  and thus the two BPAs are identical, or  $\underline{m}_1 - \underline{m}_2$  lies in the null space of  $D$ . However, since  $D$  is positive definite, then the matrix is full rank and has an empty null space.

*Symmetry.* The order of the arguments is irrelevant because  $\underline{m}_2 - \underline{m}_1$  is simply  $-(\underline{m}_1 - \underline{m}_2)$ , so

$$d_H(m_1, m_2) = \sqrt{\frac{1}{2}(\underline{m}_1 - \underline{m}_2)^T D(\underline{m}_1 - \underline{m}_2)} = \sqrt{\frac{1}{2}(\underline{m}_2 - \underline{m}_1)^T D(\underline{m}_2 - \underline{m}_1)} = d_H(m_2, m_1) \tag{21}$$

*Subadditivity.* As with the non-negativity criterion, we conjecture that  $d_H$  satisfies the triangle inequality without offering a formal proof, which is nontrivial since the property would need to be shown for any combination of singleton or nonsingleton sets. How-

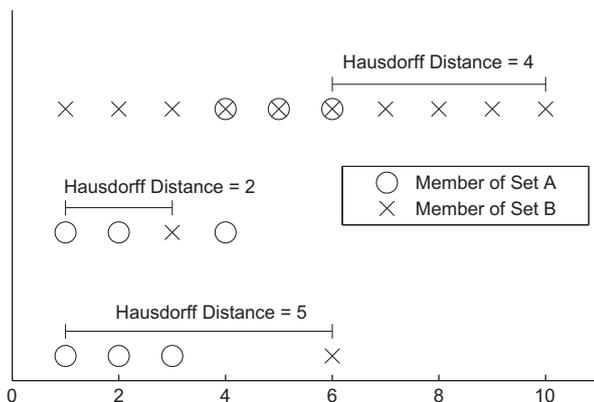


Fig. 2. Illustrations of the Hausdorff distance between sets A and B made up of integers. Note that these are *not* BPAs.

ever, subadditivity can be demonstrated for the particular case illustrated in Fig. 1, in which the (Euclidean) distance between the focal elements of two disjoint BPAs is increased as the focal elements of each are kept in the same position relative to each other, maintaining each BPA’s “shape”. As two BPAs are moved apart in this manner,  $d_H$  increases at a monotonically decreasing rate (as will be demonstrated in Section 4). This ensures subadditivity over the space that contains two BPAs which maintain their “shape” but can be moved relative to one another.

### 3.3. Maximum distance

Consider again a situation similar to that illustrated in Fig. 1. The BPA corresponding to sensor A (denoted by  $A$  in this section) remains stationary, while a second BPA (denoted by  $B$  in this section) begins at the location labeled “Sensor  $B_1$ ” and moves towards the “Sensor  $B_2$ ” location while maintaining its shape, that is, the focal elements of  $B$  remain fixed relative to one another. As the BPAs move away from each other, the value of the proposed metric will monotonically increase but will asymptotically approach a maximum value. Similarly, any metric that incorporates the similarity matrix (such as Jousselme’s) will saturate at a specific value. This maximum value of  $d_H$  can be easily calculated and depends only on the intrinsic structure of each of the two BPAs. First, when two BPAs are infinitely far away from one another, the Hausdorff distance between any focal element of  $A$  and any focal element of  $B$  is

$$H(A_i, B_j) \rightarrow \infty \quad \forall i \in A, j \in B, \quad (22)$$

and thus in the limit

$$D_{i,j} = 0 \quad \forall i \in A, j \in B, \quad (23)$$

Thus, the similarity matrix for the maximum metric value,  $D_\infty$ , will have the following structure:

$$D_\infty = \begin{bmatrix} D_{1,\text{self}} & \mathbf{0} \\ \mathbf{0} & D_{2,\text{self}} \end{bmatrix}, \quad (24)$$

where  $D_{A,\text{self}}$  and  $D_{B,\text{self}}$  are similarity matrices constructed for BPAs  $A$  and  $B$  using only the focal elements in the respective individual BPA (hence the “self” designation). That is,  $D_{A,\text{self}}$  is constructed using only the focal elements of BPA  $A$ , and  $D_{B,\text{self}}$  is constructed using only the focal elements of BPA  $B$ . When the two BPAs are an infinite distance apart, none of their focal elements overlap. Because of this, the values in BPA  $A$ ’s mass vector ( $\underline{m}_A$ ) corresponding to focal elements of  $B$  are zero and vice versa. Thus, the mass vectors have the following structure:

$$\underline{m}_A = \begin{bmatrix} \underline{m}_{A,\text{self}}^T & \mathbf{0} \end{bmatrix}^T, \quad \underline{m}_B = \begin{bmatrix} \mathbf{0} & \underline{m}_{B,\text{self}}^T \end{bmatrix}^T, \quad (25)$$

where  $\underline{m}_{A,\text{self}}$  and  $\underline{m}_{B,\text{self}}$  are the mass vectors of BPAs  $A$  and  $B$  corresponding only to their own respective focal elements. We can therefore reduce the expression of the maximum distance to

$$\begin{aligned} d_{H_{\text{max}}} &= \sqrt{\frac{1}{2} (\underline{m}_A - \underline{m}_B)^T D_\infty (\underline{m}_A - \underline{m}_B)} \\ &= \sqrt{\frac{1}{2} \underline{m}_{A,\text{self}}^T D_{A,\text{self}} \underline{m}_{A,\text{self}} + \frac{1}{2} \underline{m}_{B,\text{self}}^T D_{B,\text{self}} \underline{m}_{B,\text{self}}} \\ &= \sqrt{d_{A,\text{self}}^2 + d_{B,\text{self}}^2}. \end{aligned} \quad (26)$$

Thus, the maximum distance is simply the 2-norm of what could be loosely called the “intrinsic distance” of each BPA, defined as

$$d_{i,\text{self}} = \sqrt{\frac{1}{2} \underline{m}_{i,\text{self}}^T D_{i,\text{self}} \underline{m}_{i,\text{self}}} \quad \text{for } i = 1, 2. \quad (27)$$

It is also important to recognize that the theoretical maximum value of the metric over the domain of all possible BPAs is 1. This is achieved only in several special cases when comparing two categorical BPAs that focus on elements that are an infinite distance apart (see Section 3.5).

### 3.4. Minimum distance

Calculating the minimum distance as two BPAs move closer together is less straightforward. No simple universal method for calculating the minimum distance between two arbitrarily-shaped BPAs can be found. For two BPAs with identical shapes, the minimum distance is zero because, when the two BPAs are aligned,  $\underline{m}_1 - \underline{m}_0 = 0$ . However, if the BPAs do not have the same shape, then the relative position at which the distance between them is minimized must be determined before the minimum distance can be calculated. This relative position varies depending on the structure of the BPAs. It should also be noted that the minimum distance will change based on the gain value  $K$ .

### 3.5. Special cases

Several special cases also merit discussion. First, consider the calculation of the distance between two categorical BPAs. A categorical belief assignment is a belief function completely focused on a single focal element  $A^* \subseteq \Theta$  so that

$$m_{\text{cat}}(A_i) = \begin{cases} 1 & \text{if } A_i = A^* \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

The distance metric applied to two categorical belief functions is therefore

$$d_H(m_{1,\text{cat}}, m_{2,\text{cat}}) = \sqrt{\frac{KH^*}{1 + KH^*}} \quad (29)$$

where  $H^* = H(A^{1*}, A^{2*})$  is the Hausdorff distance between the single focal elements of  $m_{1,\text{cat}}$  and  $m_{2,\text{cat}}$ ,  $A^{1*}$  and  $A^{2*}$ . Of course, if the two categorical BPAs focus on the same element, that is  $A^{1*} = A^{2*}$ , the distance between the BPAs is  $d_H = 0$  because  $H^* = 0$ . Also, if the categorical BPAs focus on elements that are an infinite distance apart then the distance between the assignments will approach the limit  $d_H = 1$ . Furthermore, in the case where the focal elements  $A^{1*}$  and  $A^{2*}$  are both singleton elements, the metric returns the value

$$d_H(m_{1,\text{cat}}, m_{2,\text{cat}}) = \sqrt{\frac{Kd^*}{1 + Kd^*}} \quad (30)$$

where  $d^* = d(A^{1*}, A^{2*})$  is simply the distance defined between the two singleton elements  $A^{1*}$  and  $A^{2*}$  (where “distance” here is that used in Eq. (11)).

Next, consider the case in which the vacuous or ignorant BPA,  $m_{\text{vac}}$ , is compared to a second arbitrary BPA,  $m_2$ . All of the mass of  $m_{\text{vac}}$  is assigned to the ignorance element,  $\Theta$ , the set of all elements on the frame of discernment, that is

$$m_{\text{vac}}(A_i) = \begin{cases} 1 & \text{if } A_i = \Theta, \\ 0 & \text{otherwise.} \end{cases} \quad (31)$$

Suppose that the frame of discernment is infinite. An example of this is the discretized real line that could be used to represent precision sensor measurements. Assuming that  $m_2$  contains only finite focal elements, the Hausdorff distance between any one of its focal elements and  $\Theta$  is infinite. Thus the elements of the similarity matrix  $D$  relating the elements of  $m_2$  to  $\Theta$  approach zero. The metric equation therefore assumes the same form as in Eq. (26). Since the vacuous BPA contains only a single focal element, the intrinsic

distance for it defined in Eq. (27) is  $d_{m_{vac},self} = \sqrt{1/2}$ . Thus, the distance between an arbitrary BPA made up of finite focal elements,  $m_2$ , and an ignorant BPA,  $m_{vac}$  is

$$d_H(m_{vac}, m_2) = \sqrt{\frac{1}{2} + d_{m_2,self}^2}, \tag{32}$$

where  $d_{m_2,self}^2$  is defined in Eq. (27). Note that if  $m_2$  is a categorical BPA, then  $d_{m_2,self}^2 = 1/2$  and  $d_H(m_{vac}, m_2) = 1$ . Also note that the form in Eq. (32) does not hold for BPAs defined on finite frames of discernment.

### 4. Examples

This section contains examples illustrating several properties of the proposed metric and performance comparisons with other metrics. First, the proposed metric's performance is compared to others that are based on cardinality for an example similar to that discussed in Section 3. The effect of the gain parameter is then analyzed. A third example demonstrates the proposed metric's ability to differentiate malfunctioning sensors from those that are miscalibrated, and highlights the benefits of using the proposed metric for FDI and AI scenarios in comparison with Josselme's distance and others based on cardinality. A final example uses experimental data from a small unmanned air vehicle autopilot to confirm the conclusions of the previous example.

#### 4.1. Comparison with cardinality-based metrics

The motivating example illustrated in Fig. 1 can be used to compare the proposed distance metric with cardinality-based metrics such as that proposed by Josselme or Fixsen and Mahler. In this comparison, BPA A is once again comprised of a Gaussian belief distribution over singleton sets. BPA B is comprised of a single belief assignment to one nonsingleton set. The frame of discernment is once again created by discretizing the infinite set of real numbers. BPA A is held stationary while the position of BPA B is shifted left or right along the real line, such that the absolute value of the midpoints of each BPA take on different values.

Fig. 3 shows the evolution of the proposed orderable set metric, Josselme's metric, and Fixsen and Mahler's pseudo-distance when BPA A is held stationary and B is shifted right to different locations. Note that even when the midpoints of the distributions coincide, the metric values are greater than zero due to the difference in

structure between A and B. More importantly, note that Josselme's and Fixsen and Mahler's metrics immediately saturate once the belief distributions no longer overlap. In contrast, the proposed metric distance does not saturate, but continues to increase (albeit at a decreasing rate). This continued variability in metric output makes the proposed distance attractive in numerous applications. Note that the gain value in this case is  $K = 1$ .

#### 4.2. Gain tuning

One feature of the proposed orderable set metric is that it can be tuned to provide a desired level of variability in a certain region of interest. This is accomplished by adjusting the gain parameter  $K$  so that the slope of the metric with respect to the approximate Hausdorff distance between BPAs is substantial in the measurement domain under consideration. A large variation of  $d_H$  over the domain of expected BPAs in the application will make easier the task of distinguishing between small disturbances and larger, more significant ones. For instance, if one were picking distance metric thresholds above and below which decisions are made based on sensor conflict, one would want reasonable dynamic range in the distance measure so that thresholds are not crossed erratically or unexpectedly due to noise in the system or environmental disturbances. Fig. 4 shows the behavior of the proposed metric with varying values of  $K$  for the example case used in Section 4.1. Note that as  $K$  increases from 0.01, the variation in the gain over the given window first increases, and then rapidly decreases for gain values greater than 1. Thus, maximum variability over the given window occurs at  $K = 0.1$ . Also note that the minimum and maximum values of the metric are nontrivial functions of  $K$ . Minimum values consistently increase with increasing  $K$ , while the maximum value decreases as  $K$  increases.

It is clear from Fig. 4 that if the maximum Hausdorff distance between the midpoints of the two BPAs is on the order of 10, a gain value of 0.1 provides suitable variability to differentiate between BPAs in general agreement and those that are in strong disagreement. However, if the maximum Hausdorff distance is anticipated to be on the order of 100, a better choice might be  $K = 0.01$  since this will provide suitable variation over a larger window. For smaller maximum Hausdorff distances on the order of 1, a gain value of  $K = 1$  seems to show suitable variation. For the general case involving more complex BPAs, where a maximum Hausdorff distance cannot be easily estimated, simulation will likely be required in order to appropriately tune the metric.

It should also be noted that the value of a suitable gain will depend on the complexity and shape of the BPAs being considered.

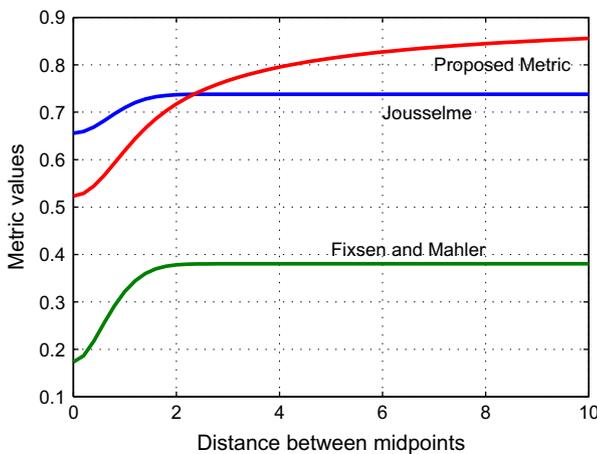


Fig. 3. Values for three distance metrics as two BPAs move further away from each other. Note that minimum distances are nonzero due to the difference in structure between the BPAs.

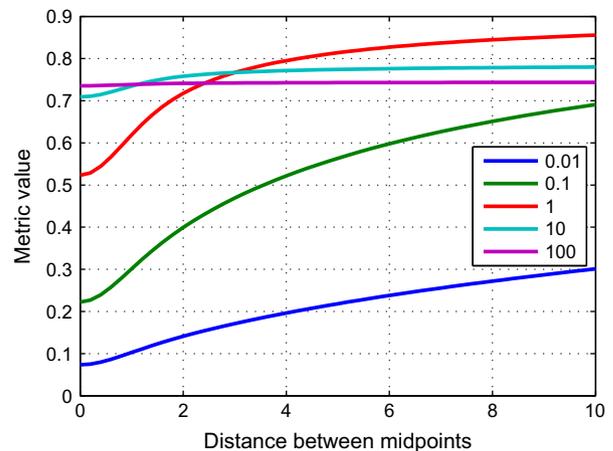


Fig. 4. Values for the New Metric with different gains,  $K$ , on the Hausdorff distance.

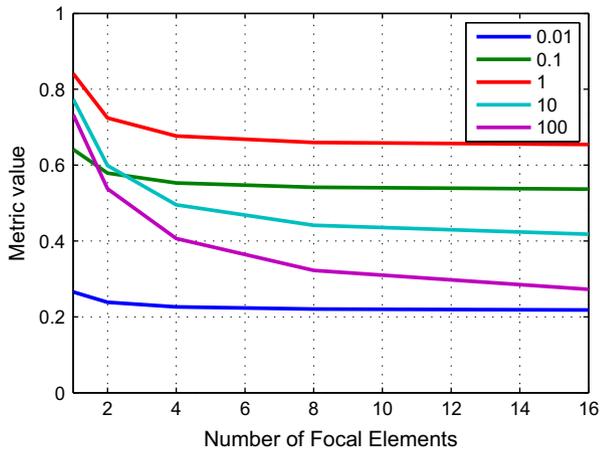


Fig. 5. Values for the New Metric with different gains,  $K$ , on the Hausdorff distance with varying BPA structure.

Fig. 5 shows the values of  $d_H$  for a case similar to that shown in Section 4.1 and Fig. 1. In this case, however, the distance between the midpoints of the two BPAs is held constant while the structure is changed. The focal elements of BPA  $B$  consist of the region between 4 and 8 divided into a number of equally sized focal elements, each with an equal mass. The number of focal elements is increased beginning with 1 ( $B$  is a categorical BPA), and ending with 16 smaller focal elements. It is clear that a higher gain value results in more variation of the distance with respect to the number of focal elements. This particular example is meant to illustrate one way that the structure or complexity of BPA could impact the desired gain value for a given application.

The gain value is of course heavily dependent on the relationship between the discretization of the measurement space and the units used to measure distance. If the unit of distance is much smaller than the width of the singleton members of the frame of discernment (created by the discretization of the space), then a very small gain will be required to produce desirable results. Thus, the gain can be interpreted simply as a scale factor for relating the values of physical distances to the more abstract domain of elements of the frame of discernment.

#### 4.3. Simulated faulty magnetometer rejection example

An example case is presented motivated by autonomous vehicle guidance and control. Specifically, the guided projectile application discussed here is similar to that outlined in Ref. [7]. Consider a sensor that provides a vector measurement whose output is the inner product between a sensitive axis and a uniform vector field. A common example is a single-axis magnetometer, which is used extensively in smart weapons guidance packages. For the notional case considered here, suppose sensor outputs are normalized between  $-1$  and  $1$ , and that the vehicle roll angle can be derived from the sensor signal  $m$  according to

$$\phi = \arcsin(m). \quad (33)$$

Note that, in order to provide a straightforward example, the ambiguity that results when taking the arcsine of a vector measurement (producing two possible roll angles) is not considered here as it is in Ref. [7].

Suppose a vehicle is equipped with three such roll sensors producing  $m_j$ ,  $j = 1, 2, 3$ . The frame of discernment  $\Theta_\phi$  is created by discretizing the range of roll angles between  $0$  and  $2\pi$  into 600 equally-spaced elements. At each timestep, the sensor maps measurement  $m_j$  to a BPA on  $\Theta_\phi$ . Recognizing that there is noise asso-

ciated with each measurement, belief is assigned according to a Gaussian distribution to singleton elements only around a mean location determined by Eq. (33). In order to ease the computational burden, the Gaussian distribution is truncated at some distance from the mean (where the belief mass becomes insignificant) and the mass vector is re-normalized to satisfy Eq. (3). Note that while this example is concerned only with Gaussian-distributed belief among singleton sets (and thus Bayesian methods such as Bhattacharyya's distance [21] may be used to measure distance), it is meant for illustrative purposes only and similar results could be generated for more complicated BPAs involving nonsingleton focal elements, such as those found in Ref. [7].

Assume that the three example sensors are coaligned, but not aligned with the roll axis of the projectile. Then all three sensors output a sinusoidally-varying signal  $m_j$  at the roll frequency. Simulated data is produced using a projectile flight simulation, and Gaussian white noise with a standard deviation of  $\sigma_s$  is added to this simulated sensor data. Then BPAs are created at each timestep using the sensor model described above. A 2.5 s window of the projectile flight trajectory is used here for analysis purposes.

The nonlinear mapping between  $m_j$  and roll angle  $\phi$  in Eq. (33), and the Gaussian noise associated with  $m_j$ , lead to a probability distribution on  $\phi$  given by the probability density function

$$f_{\phi_j}(\phi) = \frac{\cos(\phi)}{\sqrt{(2\pi)\sigma_s}} e^{-\frac{(\sin(\phi)-m_j)^2}{2\sigma_s^2}}. \quad (34)$$

The shape of this distribution is not static as roll angle varies, but it can be approximated for sensor belief assignment purposes by a Gaussian with an appropriate constant standard deviation. In this case, the BPA is chosen to be a set of singleton elements with a Gaussian mass distribution with standard deviation  $\sigma_{bpa} = \sigma_s$ . The justification for equating  $\sigma_{bpa}$  and  $\sigma_s$  is that the nonlinear transformation between measurements  $m_j$  and the roll angle  $\phi$  is the arcsine function, which with reasonable accuracy can be approximated as a line with unity slope and zero intercept. Thus, using this linear approximation the noise characteristics of  $\phi$  and  $m_j$  are identical. Fig. 6 shows example BPAs from each sensor at a single timestep in the simulated trajectory using this Gaussian distribution. For all of the magnetometer examples, the gain value,  $K$ , is 0.1.

The focus of these examples is to examine how sensor disagreement can be captured with the proposed metric, especially in comparison to cardinality-based distances. At each timestep, the distances between all three possible pairs of sensors are calculated using the new distance proposed here and Jousselme's distance. A moving average filter is applied to each metric with a averaging

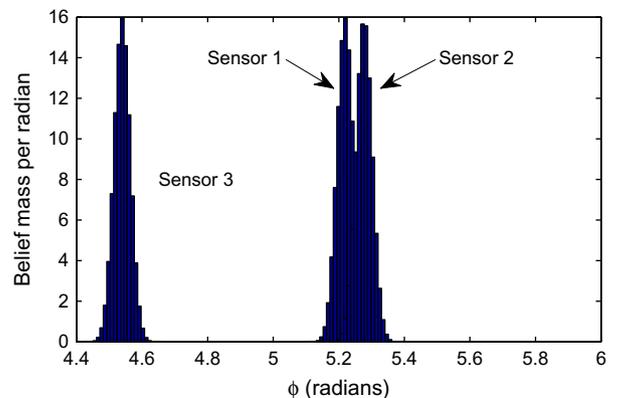


Fig. 6. Typical belief assignment output from three sensors. Sensors 1 and 2 are functioning correctly, while sensor 3 is malfunctioning and returning an inaccurate value.

length of 50 samples. In an ideal sense, if one sensor is in repeated disagreement with the majority of others it can be assumed to be malfunctioning, and would hopefully be easily identified since distance metrics involving the malfunctioning sensor would be consistently higher than those involving functional sensors. This “plurality” method is one simple technique to accomplish FDI in a system with redundant uncorrelated sensors.

#### 4.3.1. Results with an accurate assumed noise

Distance metric performance is evaluated first in the instance where sensor noise is precisely known such that  $\sigma_s = 0.025$  and  $\sigma_{bpa} = 0.025$ . The two lefthand plots of Fig. 7 show the three computed distance metrics between each pair of sensors for both the proposed distance  $d_H$  and Josselme’s distance  $d_J$ . Note that because there is some noise associated with each sensor, they do exhibit some disagreement, but in general the distances are low and uniform across all sensor pairs indicating that all sensor outputs correlate well.

Next consider the case in which sensor 3 fails. Failure is simulated by generating a uniformly random signal of sensor outputs (although this failure mode for most sensors is quite extreme, it demonstrates the concept of consistent disagreement). The two righthand plots of Fig. 7 show the three computed distance metrics for both  $d_H$  and  $d_J$  for the sensor failure example. Here, both distances are able to differentiate between functioning sensors and malfunctioning sensors. However, it is more difficult to set a tolerance for Josselme’s distance such that, if the distance is above this tolerance it could be assumed that a sensor is malfunctioning. This is because the inherent distance created by noise is already close to the maximum value of the distance, a situation analogous to that analyzed in Section 2. In contrast, the proposed metric differentiates easily between functioning pairs and those that involve malfunctioning sensors, so that a threshold can be set to determine when sensors display consistent extreme disagreement.

#### 4.3.2. Results with an inaccurate assumed noise

When using cardinality-based metrics, knowledge of sensor noise characteristics becomes imperative so that the Gaussian

BPA’s can be designed appropriately and precisely. If assumed sensor noise is too low, the Gaussian BPA’s may be too narrow, resulting in sensor BPA’s that do not overlap simply due to measurement noise being higher than anticipated. In such cases, it may become impossible to set a distance threshold using cardinality-based metrics that distinguishes between malfunctioning sensors and sensors whose noise or bias characteristics are simply larger than expected. To demonstrate this, assume that now sensor noise is increased to  $\sigma_s = 0.10$  while the BPA standard deviation is held constant at  $\sigma_{bpa} = 0.0393$ . The two lefthand plots of Fig. 8 show filtered distances calculated assuming all three sensors are functioning normally with this higher noise value. Note that, once again, all metrics show agreement between each pair of sensors since all are assumed operational. However, also note that as expected the average value of these filtered metrics is higher than in the lower noise example of Fig. 7.

Now, consider the case where sensor 3 has again failed, producing purely random outputs. Filtered distance metrics for this case are shown in the plots on the right in Fig. 8. Here, it is clear that because the actual noise is greater than anticipated, filtered metric values for  $d_J$  are nearly saturated even for the functioning sensor pair. Alternatively,  $d_H$  maintains good performance, clearly differentiating between the failed sensors and those that show relative agreement. In an FDI system, it would likely be necessary to determine a threshold above which a sensor would be considered unreliable. When using cardinality-based metrics such as  $d_J$ , it would clearly be difficult if not impossible to establish this threshold given imprecise *a priori* knowledge of noise characteristics. The proposed metric, on the other hand, still maintains a wide range for threshold placement and is therefore more robust to uncertainties in sensor noise. This robustness can lead directly to less effort spent in sensor characterization and noise identification.

#### 4.3.3. Results with an ideal mass distribution, but inaccurate assumed noise

In order to verify that the results of Figs. 7 and 8 were not influenced by the Gaussian noise approximation, the example case was run again using the rigorously-correct form of the measurement

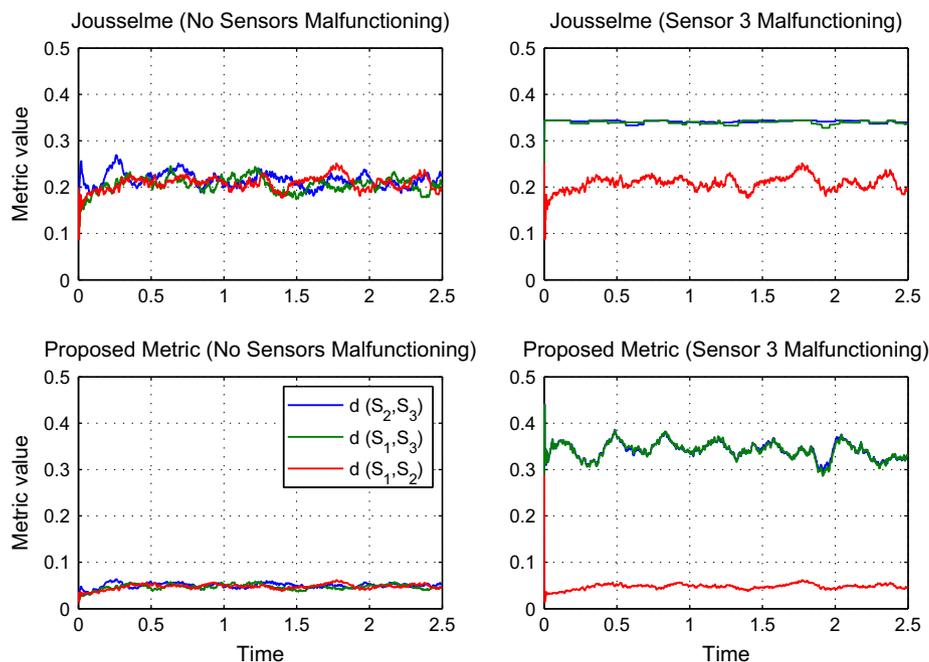


Fig. 7. Filtered distances calculated with accurate knowledge of sensor noise.

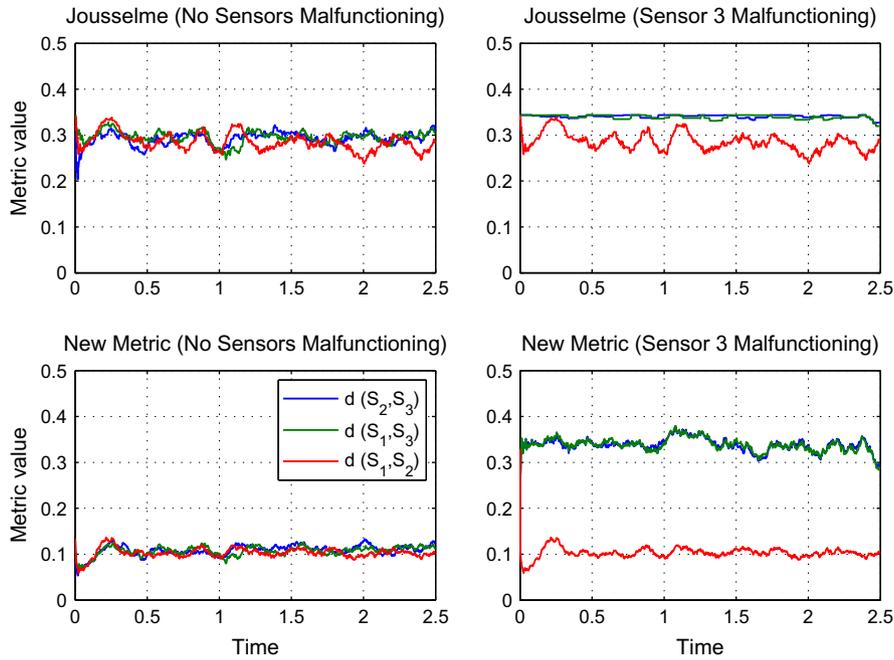


Fig. 8. Filtered distances calculated by the two metrics with inaccurate assumed noise.

noise distribution. In this case, given measurements  $m_j$  BPAs are created according to the distribution in Eq. (34). Thus, the width of the BPAs change with respect to roll angle.

As in Section 4.3.2, it is assumed that the level of sensor noise is not known precisely so that the actual  $\sigma_s = 0.1$  while the standard deviation assumed when using Eq. (34) is  $\sigma_{s,assumed} = 0.025$ . Fig. 9 shows the distance metric outputs in this case for both  $d_j$  and  $d_H$ . Even with the correct non-Gaussian belief mass distribution, the righthand plots clearly show that  $d_H$  provides clear separation between functioning and nonfunctioning sensors, allowing an FDI threshold to be established. It is also worth noting that, because the shape of the mass distribution changes depending on the mea-

surement, in this case neither metric has a convenient easily-calculated fixed maximum value as described in Section 3.3. In a more complicated scenario, such as those described in Ref. [7], this lack of a fixed maximum value would make the design of an FDI algorithm more challenging and thus the advantages of the proposed metric in reducing the likelihood of saturation become even more important.

#### 4.4. Sensor rejection with accelerometer data

Data were gathered using accelerometers on two inertial measurement units (IMUs) attached to a small unmanned helicopter

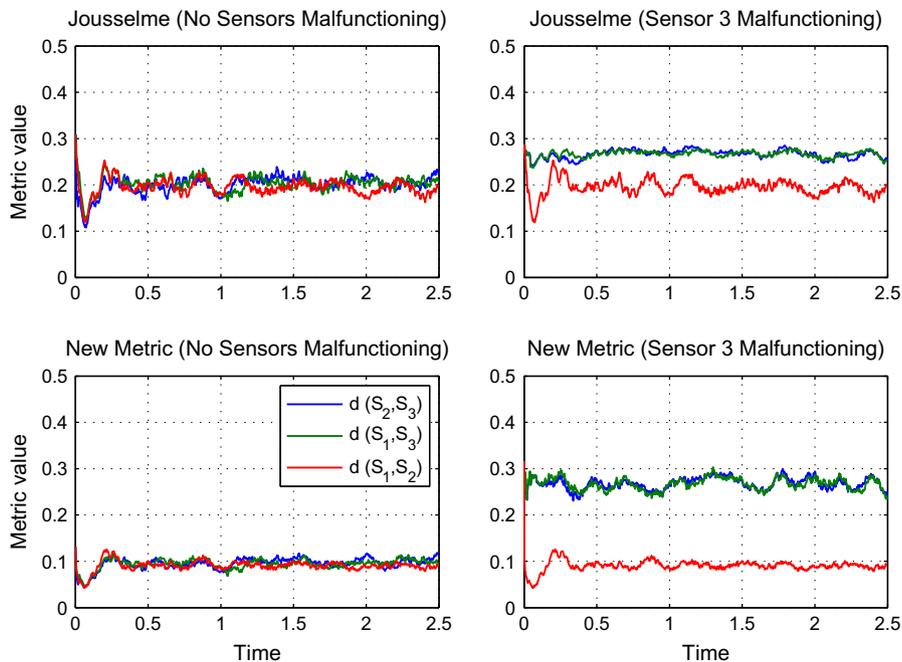


Fig. 9. Filtered distances calculated by the two metrics with the ideal mass distribution from Eq. 34, but inaccurate assumed noise.

for the purpose of testing metric performance in an experimental environment. The results further illustrate the superiority of the proposed metric over cardinality-based ones when using data collected from actual sensors.

#### 4.4.1. Low vibration test

For the first test, in order to limit vibration, an unpowered helicopter was moved by hand inside the lab. Data were collected from the accelerometers aligned with the forward-backward direction of the vehicle. This test shows how the metrics perform when used with a comparatively clean dataset. Note that the noise characteristics of these sensors are not known precisely in the experiment, and can vary among individual sensors and environmental conditions. The data from the accelerometers are sampled at a rate of 10 Hz. For each measurement, a BPA of singleton elements with a Gaussian distribution, similar to one of those shown in Fig. 6, is created. The standard deviation for these BPAs is  $15 \mu\text{g}$ , based loosely on the noise observed when the IMUs are not moving. Data from a hypothetical malfunctioning sensor were simulated by generating evenly distributed random measurements between  $-1 \text{ g}$  and  $1 \text{ g}$ . The gain parameter was adjusted to a value of  $K = 0.005$  and the distances were filtered using a moving average of 30 previous values. Results are shown in Fig. 10. Again, with realistic sensor data, the new metric more clearly differentiates the functioning sensors from spurious data. Using the proposed metric a threshold for detecting a malfunctioning sensor could easily be set at a value anywhere between 0.4 and 0.6 and correctly detect the malfunctioning sensor over the entire data set. There is no threshold value that would accomplish the same using Jousselme's metric.

#### 4.4.2. High vibration test

For the second test using accelerometer data, flight data was collected on-board the small unmanned helicopter. Vibration of the helicopter structure induces a very large amount of noise in the accelerometer data, making it difficult to extract any type of usable information from the data including the information required for FDI. This low signal-to-noise ratio represents a very difficult environment in which to attempt to differentiate functional from non-functional sensors, but is not uncommon in many realis-

tic sensor fusion scenarios. Furthermore, FDI and AI systems designed for low-noise environments may often need to be made robust against unexpected low-frequency noise.

The raw data collected from this flight is shown in Fig. 11. The data from the accelerometers are now aligned with the vertical direction of the vehicle, and thus the mean acceleration value returned is approximately  $-1 \text{ g}$ . The data for the hypothetical malfunctioning sensor consist of uniformly distributed randomly generated values between  $-2 \text{ g}$  and  $2 \text{ g}$ . Belief is assigned in the manner described in Section 4.4.1 with the same standard deviation of  $15 \mu\text{g}$ . The distances, filtered with a moving average filter considering of 50 past values, are shown in Fig. 12.

Jousselme's metric is unable to differentiate between the malfunctioning and correctly functioning sensors because the value is saturated on most measurements. With the proposed metric, it may be significantly more difficult to establish a threshold between sensors in agreement and disagreement than in previous examples. However, the new metric still does offer a usable degree of differentiation in this case. Furthermore, it is easy for a human to differentiate the data from the two functioning sensors and the malfunctioning sensor in Fig. 11, so for AI and FDI applications one would require a conflict measure that reflects this intuition. The proposed metric clearly offers more intuitive outputs than cardinality-based metrics in this case, and is further proven to be more robust to unexpected failure or noise disturbances.

#### 4.5. Summary of example results

In all of the cases examined above, the new metric distinguishes a malfunctioning sensor from correctly functioning sensors more clearly than Jousselme's metric. It should be noted that Jousselme's metric could possibly be used successfully in any of the cases. However this would require much more filtering of the data and fine tuning of the parameters of the FDI system or more advanced and accurate allocation of the sensor's belief based on a highly refined sensor model. The advantage of the new metric lies in its robustness in several senses. It is tolerant of errors and engineering approximations in sensor modeling and can therefore relieve source characterization requirements during the system design

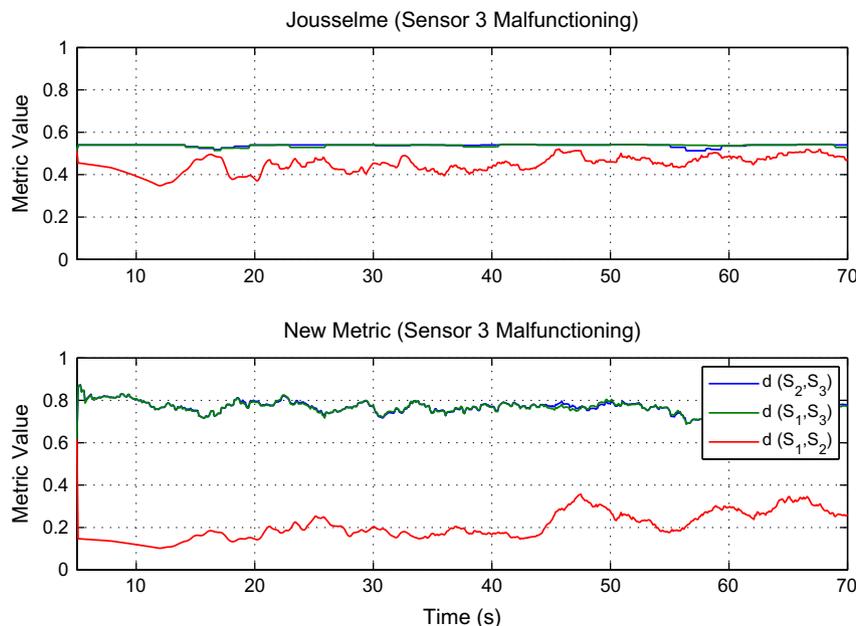


Fig. 10. Filtered metric values based on accelerometer data gathered while moving a small unmanned aircraft smoothly by hand in a lab. The simulated output of the malfunctioning sensor is a string of randomly generated values.

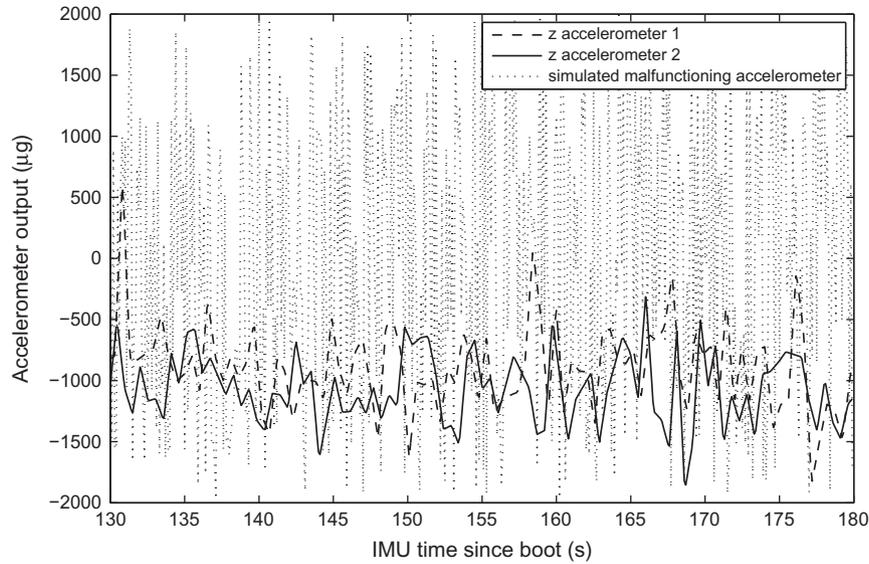


Fig. 11. Raw accelerometer data collected from test flight, along with random noise used to simulate a malfunctioning sensor.

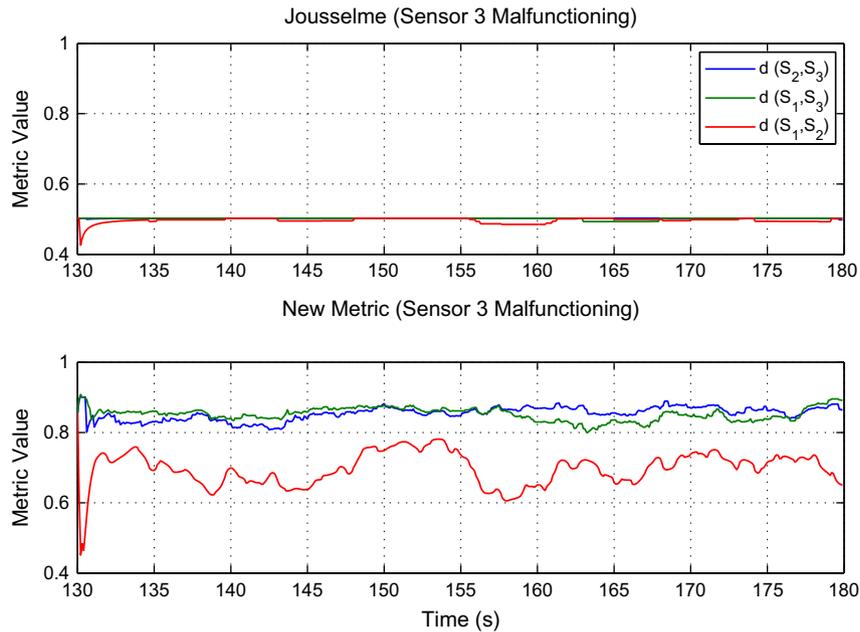


Fig. 12. Filtered distance metric values based on flight test accelerometer data.

phase. It is also robust to unforeseen disturbances that could cause a system or sensors to behave in a way that undermines more fragile cardinality-based metrics.

**5. Conclusion**

The ability to measure disagreement between belief probability assignments is a cornerstone of belief function theory applied to real-time sensing and artificial intelligence applications. In this paper, a new distance metric was proposed that is designed specifically for orderable frames of discernment. The primary advantage of the proposed metric is that it avoids saturation even when BPAs do not overlap, since the physical distance (in the form of the Hausdorff distance) between focal elements plays a role in the metric. The proposed metric leverages the quadratic form originally proposed by Joussleme, and metric properties are explored in detail. Tuning of the metric’s gain parameter, which allows sca-

lability to problems of different discretizations, was also investigated. Several example cases using orderable measurement spaces demonstrated the superiority of the proposed metric over cardinality-based formulations, especially in common instances where sensor noise is not accurately known *a priori*. When characterizing sensor agreement on any orderable space, the proposed metric invariably proves to be more robust, flexible, and accurate when compared to any previous distance metrics that rely on cardinality.

**Acknowledgements**

This material is based upon work supported by, or in part by, the US Army Research Laboratory and the US Army Research Office under Grant No. W911NF-12-1-0274 from the Complex Dynamics and Systems program.

This material is based upon work supported by the National Science Foundation under Grant No. 1252521.

### Appendix A. Proof of positive definite property of a $3 \times 3$ similarity matrix

**Given:** a real, symmetric  $3 \times 3$  matrix  $A$  such that

$$A = \begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}, \quad (\text{A.1})$$

with the following constraints:

$$0 < a < 1, (\text{A.2})$$

$$0 < b < 1, (\text{A.3})$$

$$0 < c < 1, (\text{A.4})$$

$$\frac{|a-b|}{ab} \leq \frac{1-c}{c} \leq \frac{a+b-2ab}{ab}. \quad (\text{A.5})$$

**Prove:**  $A$  is symmetric positive definite.

**Proof.** If all of the principle minors of a matrix are positive, then that matrix is positive definite. The two principle minors of  $A$  are given by

$$A_1 = [1] \quad A_2 = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}. \quad (\text{A.6})$$

Now

$$\det(A_1) = 1 > 0 \quad \det(A_2) = 1 - a^2 > 0, \quad (\text{A.7})$$

and thus all principal minors are positive. Therefore, in order to prove that  $A$  is positive definite, we now need only show that  $\det(A) > 0$ . The determinant of  $A$  is given by

$$\det(A) = 1 + 2abc - a^2 - b^2 - c^2. \quad (\text{A.8})$$

While it may be possible to manipulate constraints A.2, A.3, A.4 and A.5 directly to show that the expression in (A.8) is always positive, this manipulations highly nontrivial, and instead a calculus-based approach will be used. Consider

$$f(a, b, c) = 1 + 2abc - a^2 - b^2 - c^2. \quad (\text{A.9})$$

The gradient of  $f$  is given by

$$\nabla f = \begin{bmatrix} -2a - 2bc \\ -2b - 2ac \\ -2c - 2bc \end{bmatrix}. \quad (\text{A.10})$$

From Eq. (A.10), it is clear that  $f$  has only two critical points located at  $(a, b, c) = (0, 0, 0)$  or  $(1, 1, 1)$ . As evidenced by the constraints in Eqs. A.2, A.3, A.4 these points occur on the boundary of the open set under consideration. The critical points will be evaluated using the Hessian of  $f$ , given by

$$H = \begin{bmatrix} -2 & 2c & 2b \\ 2c & -2 & 2a \\ 2b & 2a & -2 \end{bmatrix}. \quad (\text{A.11})$$

Evaluating  $H$  at the two critical points, we find that the eigenvalues of  $H(0, 0, 0)$  are all negative, and thus  $f$  has a local maximum of 1 at  $(a, b, c) = (0, 0, 0)$ . Alternatively, we find that the eigenvalues of  $H(1, 1, 1)$  have mixed sign and thus  $(a, b, c) = (1, 1, 1)$  is a saddle point. Thus,  $f$  has no local minima on the open set or the open set boundary.

Since  $f$  is continuous everywhere, the Extreme Value Theorem dictates that the minimum of  $\det(A)$  must lie either at a crit-

ical point within the open set defined by constraints A.2, A.3, A.4 and A.5 or along the open set boundary. Thus  $\det(A)$  has no local minima anywhere within the open set.

We will now evaluate  $f$  along each of these boundaries individually to find minima along each boundary. The smallest minimum of  $\det(A)$  evaluated over all boundaries will globally minimize  $\det(A)$  within our bounded region.

#### A.1. Boundary 1

Boundary 1 is defined as  $a = 0$ . In this region, we have  $0 \leq b \leq 1$ . From constraint (A.5),  $c \rightarrow 0$  in the limit that  $a \rightarrow 0$ , so  $c = 0$  lies at the open set boundary. Therefore, the determinant function evaluated along the boundary  $f_1$  is given by

$$f_1(a, b, c) = 1 - b^2. \quad (\text{A.12})$$

This function has a minimum  $f_1 = 0$  when  $b = 1$ .

#### A.2. Boundary 2

Boundary 2 is defined as  $a = 1$ . Constraint (A.5) evaluated here gives  $0 \leq b \leq 1$  and  $c = b$ , so the determinant function evaluated along this boundary  $f_2$  is given by

$$f_2(a, b, c) = -(b-c)^2 = 0. \quad (\text{A.13})$$

Thus, this function is zero over the entire boundary.

#### A.3. Boundary 3

Boundary 3 is defined as  $b = 0$ . In this region,  $0 \leq a \leq 1$ . Again, from constraint (A.5),  $c \rightarrow 0$  in the limit that  $b \rightarrow 0$ , so  $c = 0$  lies at the open set boundary. Therefore, the determinant function evaluated along this boundary  $f_3$  is given by

$$f_3(a, b, c) = 1 - a^2. \quad (\text{A.14})$$

This function has a minimum  $f_3 = 0$  when  $a = 1$ .

#### A.4. Boundary 4

Boundary 4 is defined as  $b = 1$ . Constraint (A.5) evaluated here gives  $0 \leq a \leq 1$  and  $c = a$ , so the determinant function evaluated along this boundary  $f_4$  is given by

$$f_4(a, b, c) = -(a-c)^2 = 0. \quad (\text{A.15})$$

Thus, this function is zero over the whole boundary.

#### A.5. Boundary 5

Boundary 5 is defined at the surface in which  $c$  reaches its minimum value for a given  $a$  and  $b$ , defined by

$$c = \frac{ab}{a+b-ab}. \quad (\text{A.16})$$

Along this boundary,  $0 \leq a \leq 1$  and  $0 \leq b \leq 1$ . Therefore, the determinant function evaluated along this boundary,  $f_5$ , is given by

$$f_5(a, b, c) = -a^2 - b^2 - \frac{a^2 b^2}{(a+b-ab)^2} + \frac{2a^2 b^2}{a+b-ab} + 1. \quad (\text{A.17})$$

This function achieves a minimum value of  $f_5 = 0$  along the line  $a = 0$  or  $b = 0$ .

#### A.6. Boundary 6

Boundary 6 is defined at the surface in which  $c$  reaches its maximum value for a given  $a$  and  $b$  defined by

$$c = \frac{ab}{|a-b|+ab} \quad (\text{A.18})$$

Along this boundary, we have  $0 \leq a \leq 1$  and  $0 \leq b \leq 1$ . Therefore, the determinant function evaluated along this boundary  $f_6$  is given by

$$f_6(a, b, c) = -a^2 - b^2 - \frac{a^2b^2}{(|a-b|+ab)^2} + \frac{2a^2b^2}{|a-b|+ab} + 1 \quad (\text{A.19})$$

This function achieves a minimum value of  $f_6 = 0$  along the line  $a = 0$ ,  $b = 0$ , or  $a = b$  (where  $c = 0$ ).

The minimum values of  $f$  along the 6 boundaries of the open region defined by constraints A.2, A.3, A.4 and A.5, all occurring at  $f = 0$ , and the absence of any local minima of  $f$ , verify that the minimum value of the determinant occurs on the boundary, and that this global minimum value is 0. Thus the determinant has a value greater than 0 everywhere in the open set and the matrix  $A$  is positive definite under the conditions defined in A.2, A.3, A.4 and A.5.

## References

- [1] A.P. Dempster, A generalization of Bayesian inference, *Journal of the Royal Statistical Society* 30 (2) (1968) 205–247, [http://dx.doi.org/10.1007/978-3-540-44792-4\\_4](http://dx.doi.org/10.1007/978-3-540-44792-4_4).
- [2] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Application*, Mathematics in Science and Engineering, vol. 14, Academic Press, 1980.
- [3] B.R. Cobb, P.P. Shenoy, On the plausibility transformation method for translating belief function models to probability models, *International Journal of Approximate Reasoning* 41 (3) (2006) 314–330, <http://dx.doi.org/10.1016/j.ijar.2005.06.008>.
- [4] D. Biezad, *Integrated Navigation and Guidance Systems: Vol. I*, AIAA Education Series, American Institute of Aeronautics and Astronautics, Reston, Virginia, 1999.
- [5] J. Guan, D. Bell, *Evidence Theory and its Applications*, Elsevier Science Publishers, Amsterdam, 1991.
- [6] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
- [7] J. Rogers, M. Costello, Smart projectile state estimation using evidence theory, *Journal of Guidance, Control, and Dynamics* 35 (3) (2012) 824–833.
- [8] W. Liu, Analyzing the degree of conflict among belief functions, *Artificial Intelligence* 170 (11) (2006) 909–924, <http://dx.doi.org/10.1016/j.artint.2006.05.002>.
- [9] A.-L. Jousselme, D. Grenier, E. Bossé, A new distance between two bodies of evidence, *Information Fusion* 2 (2) (2001) 91–101 ([http://dx.doi.org/10.1016/S1566-2535\(01\)00026-4](http://dx.doi.org/10.1016/S1566-2535(01)00026-4)).
- [10] A. Martin, A.-L. Jousselme, C. Osswald, Conflict measure for the discounting operation on belief functions, in: *Proceedings of the 11th International Conference on Information Fusion (FUSION 2008)*, Cologne, Germany, 2008, pp. 1–8.
- [11] D. Fixsen, R. Mahler, The modified Dempster–Shafer approach to classification, *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* 27 (1) (1997) 96–104, <http://dx.doi.org/10.1109/3468.553228>.
- [12] J. Diaz, M. Rifqi, B. Bouchon-Meunier, A Similarity Measure between Basic Belief Assignments, in: *Proceedings of the 9th International Conference on Information Fusion (FUSION 2006)*, Florence, Italy, 2006, pp. 1–6. (<http://dx.doi.org/10.1109/ICIF.2006.301730>).
- [13] Z. Elouedi, K. Mellouli, P. Smets, Assessing sensor reliability for multisensor data fusion within the transferable belief model, *IEEE Transactions on Systems, Man, and Cybernetics – Part B* 34 (1) (2004) 782–787.
- [14] B. Tessem, Approximations for efficient computation in the theory of evidence, *Artificial Intelligence* 61 (2) (1993) 315–329.
- [15] L. Zouhal, T. Denœux, An evidence-theoretic k-NN rule with parameter optimization, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 28 (2) (1998) 263–271, <http://dx.doi.org/10.1109/5326.669565>.
- [16] A.-L. Jousselme, P. Maupin, Distances in evidence theory: comprehensive survey and generalizations, *International Journal of Approximate Reasoning* 53 (2012) 118–145.
- [17] M. Florea, A. Jousselme, E. Bossé, D. Grenier, Robust combination rules for evidence theory, *Information Fusion* 10 (2) (2009) 183–197, <http://dx.doi.org/10.1016/j.inffus.2008.08.007>.
- [18] R. Zohar, R. Yehuda, The maximum weight hierarchy matching problem, *Information Fusion* 10 (2) (2009) 198–206, <http://dx.doi.org/10.1016/j.inffus.2008.08.005>.
- [19] H. Wu, M. Siegel, S. Ablay, Sensor fusion using Dempster–Shafer theory II: static weighting and Kalman filter-like dynamic weighting, in: *Proceedings of the 20th IEEE Instrumentation Technology Conference*, vol. 2, Vail, CO, 2003, pp. 907–912. (<http://dx.doi.org/10.1109/IMTC.2003.1207885>).
- [20] F. Hausdorff, *Set Theory*, American Mathematical Society (Translation from original German work Mengenlehre), 1957.
- [21] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions, *Bulletin of the Calcutta Mathematical Society* 35 (4) (1943) 99–109.
- [22] M. Bouchard, A.-L. Jousselme, P. Doré, A proof for the positive definiteness of the Jaccard index matrix, *International Journal of Approximate Reasoning* (2013), <http://dx.doi.org/10.1016/j.ijar.2013.01.006>.