# Fast computation of $L^p$ norm-based specialization distances between bodies of evidence

Mehena Loudahi, John Klein, Jean-Marc Vannobel and Olivier Colot

Lille1 University, LAGIS UMR CNRS 8219 - firstname.name@univ-lille1.fr

**Abstract.** In a recent paper [12], we introduced a new family of evidential distances in the framework of belief functions. Using specialization matrices as a representation of bodies of evidence, an evidential distance can be obtained by computing the norm of the difference of these matrices. Any matrix norm can be thus used to define a full metric. In particular, it has been shown that the $L^1$ norm-based specialization distance has nice properties. This distance takes into account the structure of focal elements and has a consistent behavior with respect to the conjunctive combination rule. However, if the frame of discernment on which the problem is defined has $n$ elements, then a specialization matrix size is $2^n \times 2^n$. The straightforward formula for computing a specialization distance involves a matrix product which can be consequently highly time consuming. In this article, several faster computation methods are provided for $L^p$ norm-based specialization distances. These methods are proposed for special kinds of mass functions as well as for the general case.

## 1 Introduction

The belief function theory, or evidence theory [6, 16], is a formal framework for reasoning under uncertainty. In the past decades, a growing interest has been shown toward determining meaningful dissimilarity measures between bodies of evidence. These measures are used in belief function approximation computation [1, 4, 5, 2], in belief functions clustering [15, 8], in evidential data classification [19], in evidential sensor reliability assessment [9, 18] or in estimation of some parameters feeding refined belief function combinations [7, 13] or update processes [11]. All (dis)similarity measures attempt to describe the degree of (non-)alikeness between belief functions in a meaningful way for the widest range of applications. Indeed, the choice of a particular measure is most of the time application-dependent.

A thorough survey about dissimilarity measures in the evidence theory and their properties was presented by Jousselme and Maupin [10]. The authors also provided generalizations of some distances thereby introducing families of new measures.

We introduced in [12] a new family of evidential distances based on specialization matrices as a representation of bodies of evidence. In particular, the $L^1$ norm-based specialization distance has unprecedented properties as compared to state-of-the-art approaches. Unfortunately, a straightforward implementation of specialization distances requires a rather large computation time. In this work, we provide several faster computation methods for $L^p$ norm-based specialization

distances. These methods allow a computation at least as fast as for usual evidential metrics.
The rest of this paper is organized as follows. Section 2 provides the preliminaries of evidence theory. In section 3, faster computation methods for the $L^p$ norm are proposed for special kinds of mass functions as well as for the general case. Finally, we conclude the paper in section 4.

## 2 Belief function framework: notations and definitions

In this section, mathematical notations for classical belief function concepts are given. The reader is expected to be familiar with belief function basics and consequently some definitions are not recalled. More material on belief functions basics is found in [12]. A greater stress is given to a reminder on matrix calculus as part of the belief function framework and on some specialization distances.

### 2.1 Belief function basics

For a given body of evidence $Ev_i$, the corresponding **mass function** representing this piece of evidence is denoted by $m_i$. These functions are set-functions with respect to a **frame of discernment** denoted by $\Omega$. The power set $2^\Omega$ is the set of all subsets of $\Omega$ and it is the domain of mass functions. For any $A \in 2^\Omega$, the **cardinality** of this set is denoted by $|A|$ and $|\Omega| = n$. The cardinality of $2^\Omega$ is denoted by $N = 2^n$. Mass functions have $[0,1]$ as codomain and they sum to one. A **focal element** of a mass function $m_i$ is a set $A \subseteq \Omega$ such that $m_i(A) > 0$. A function having only one focal element $A$ is called **categorical mass** function and denoted by $m_A$.
The most widely used way to combine pieces of evidence is to apply the **conjunctive combination rule** to their corresponding mass functions. This rule is denoted by $\textcircled{\cap}$. For two given mass functions $m_1$ and $m_2$, their conjunctive combination is denoted by $m_1 \textcircled{\cap} 2 = m_1 \textcircled{\cap} m_2$. The conjunctive rule is a generalization of evidential conditioning (or Dempster's conditioning) which is itself a generalization of Bayesian conditioning. Indeed when mass functions are replaced with probability distributions, then Bayes' theorem is retrieved. An updated mass function given $A$ is denoted by $m(.|A) = m \textcircled{\cap} m_A$. The conjunctive rule can be applied if all sources of information providing pieces of evidence are fully reliable in the sense that the pieces of evidence they provide are true.

### 2.2 Belief functions and matrix calculus

Mass functions can be viewed as vectors belonging to a vector space spanned by categorical mass functions. Since mass functions sum to one, the set of mass functions is the simplex of that vector space. In this paper, the following notations and conventions are used :

- Vectors are written in bold small letters and matrices in bold capital letters.
- Vectors are column vectors and their length is $N$. The $i^{\text{th}}$ element of a mass function vector $\mathbf{m}$ is such that $\mathbf{m}_i = m(A)$ with $i$ the index of set $A$ according to the binary order. The binary order [17] is a common way to index elements of $2^\Omega$ without supposing any order on $\Omega$.

- Matrices are square and their size is $N \times N$. A matrix can be represented by $\mathbf{X} = [X_{ij}]$, or alternatively by the notation $\mathbf{X} = [X(A,B)], \quad \forall A,B \in \Omega$. The row and column indexes $i$ and $j$ are those corresponding to the subsets $A$ and $B$ using the binary order.

Matrix calculus as part of the BFT is especially interesting when it comes to conjunctive combination computation. Indeed, from Smets [17], one has:

$$\mathbf{m}_{1\bigcirc\!\!\!\!\cap 2} = \mathbf{M}^{-1}\text{diag}\left(\mathbf{M}\mathbf{m}_1\right)\mathbf{M}\mathbf{m}_2, \tag{1}$$

with diag the operator turning a vector into a diagonal matrix and $\mathbf{M}$ a matrix such that $M(A,B) = 1$ if $A \subseteq B$ and $M(A,B) = 0$ otherwise. Note that this matrix can be computed using $n$ iterations of the following recurrence:

$$\mathbf{M}^{(i+1)} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \otimes \mathbf{M}^{(i)}, \tag{2}$$

with $\otimes$ the Kronecker matrix product and $\mathbf{M}^{(0)} = [1]$. Furthermore, the matrix $\mathbf{S}_1$ such that $\mathbf{S}_1\mathbf{m}_2 = \mathbf{m}_{1\bigcirc\!\!\!\!\cap 2}$ is called the Dempsterian specialization matrix of $m_1$. This matrix thus writes:

$$\mathbf{S}_1 = \mathbf{M}^{-1}\text{diag}\left(\mathbf{M}\mathbf{m}_1\right)\mathbf{M}. \tag{3}$$

Each element of $\mathbf{S}_1$ represents actually the mass assigned to a set $A$ after conditioning on $B$: $S_1(A,B) = m_1(A|B)$. In other words, $\mathbf{S}_1$ does not only represent the current state of belief depicted by $m_1$ but also all reachable states from $m_1$ through conditioning. From a geometric point of view [3], a specialization matrix contains the vertices the conditional subspace associated with the function $m$. Specialization matrices are consequently relevant candidates for assessing dissimilarities between bodies of evidence in consistence with evidential conditioning and the conjunctive combination rule.

### 2.3 Specialization distances

The most natural way to design distances between specialization matrices is to rely on a given matrix norm $\| \, . \, \|_x$. Indeed, suppose two bodies of evidence[1] represented respectively by $m_1$ and $m_2$, then the following function $d_{Sx}$ is a normalized full metric:

$$d_{Sx}(m_1,m_2) = \frac{1}{\rho} \| \mathbf{S_1} - \mathbf{S_2} \|_x \tag{4}$$

with $\rho = \max_{m_i,m_j} \| \mathbf{S_i} - \mathbf{S_j} \|_x$ a normalization coefficient. Such distances are called **specialization distances**.

In this article, we focus on specialization distances relying on the $L^p$ matrix norms. These distances are denoted by $d_p$. For these distances, the normalization coefficient is known in closed form: $\rho = \left(2\left(2^n - 1\right)\right)^{\frac{1}{p}}$. Choosing $L^p$ norm-based specialization distances is justified by the fact that, in particular, the distance $d_1$ has interesting properties [12]. It takes into account the interactions[2] between focal elements (structural property) and two mass functions are necessarily closer

[1] For the sake of clarity, the distinction between a body of evidence and its corresponding mass function is omitted in equations.

[2] For instance, focal elements may have a non-empty intersection.

after conjunctive combination with addionnal evidence (conjunctive consistency property).
The straightforward computation of distances $d_p$ is given by :

$$d_p(m_1,m_2) = \frac{1}{\rho} \parallel \mathbf{M}^{-1}\text{diag}\left(\mathbf{M}(\mathbf{m}_1 - \mathbf{m}_2)\right)\mathbf{M} \parallel_p . \tag{5}$$

Unfortunately, equation (5) involves a matrix product. Its complexity is thus $O(N^3)$. Such a complexity can be prohibitive for many application contexts and appears to be greater than the complexity of other evidential distances. Consequently, faster ways to compute distances $L^p$ are investigated in the next section.

# 3 Faster computation of $L^p$ norm-based specialization distances

In this section, new computation methods are introduced for $L^p$ norm-based specialization distances. First, some results are given for special cases of mass functions and in the last subsection, an algorithm is provided for the general case.

## 3.1 Distances between categorical mass functions

A fast way to compute $L^p$ norm-based specialization distances between categorical mass functions is already given in [12]. Indeed, it is proved in this article that there exists an bijective link between the Hamming set distance and distances $d_p$ restricted to categorical mass functions. More precisely, one has:

$$d_p\left(m_A,m_B\right) = \left(\frac{N - 2^{n-|A\Delta B|}}{N-1}\right)^{\frac{1}{p}}, \tag{6}$$

with $\Delta$ the set symmetric difference. The cardinality of the set symmetric difference is the Hamming set distance.
The interest of equation (6) is twofold: first the computation for such distances is now just $O\left(1\right)$, and second, it also sheds light on the fact that there is an order isomorphism between the Hamming set distance and the specialization distance. This latter property is extremely important for evidential distances as it proves that the distance abides by the structural principle stated in [10].

## 3.2 Distances between a categorical mass function and any mass function

In this subsection, a broader case is investigated: computation of distances $d_p$ between a categorical mass function and any mass function. We provide a result only for the $L^1$ norm-based specialization distance $d_1$:

**Proposition 1.** *Suppose m is a mass function on a frame $\Omega$. Suppose $A \subseteq \Omega$ and $m_A$ is its corresponding categorical mass function. The specialization matrix of m is denoted by $\mathbf{S}$ and that of $m_A$ by $\mathbf{S}_A$. One has:*

$$d_1\left(m,m_A\right) = \frac{N - \parallel \mathbf{S} \circ \mathbf{S_A} \parallel_1}{N-1}, \tag{7}$$

$$= \frac{N - tr(\mathbf{S}^t\mathbf{S}_A)}{N-1}, \tag{8}$$

*with $\circ$ the Hadamard matrix product*[3]*, $tr$ the matrix trace operator and ${}^t\mathbf{S}_A$ the transpose matrix of $\mathbf{S}_A$.*

*Proof.* By definition of the $L^1$ norm, one has :

$$\| \mathbf{S} - \mathbf{S_A} \|_1 = \sum_{X,Y \subseteq \Omega} |S(X,Y) - S_A(X,Y)|.$$

It is known that $S_A(X,Y) = 1$ if $A \cap Y = X$ and $S_A(X,Y) = 0$ otherwise, which gives:

$$\begin{aligned}
\| \mathbf{S} - \mathbf{S_A} \|_1 &= \sum_{\substack{X,Y \subseteq \Omega \\ X = A \cap Y}} (1 - S(X,Y)) + \sum_{\substack{X,Y \subseteq \Omega \\ X \neq A \cap Y}} S(X,Y), \\
&= \sum_{\substack{X,Y \subseteq \Omega \\ X = A \cap Y}} 1 + \sum_{\substack{X,Y \subseteq \Omega \\ X \neq A \cap Y}} S(X,Y) - \sum_{\substack{X,Y \subseteq \Omega \\ X = A \cap Y}} S(X,Y), \\
&= \| \mathbf{S_A} \|_1 + \sum_{X,Y \subseteq \Omega} S(X,Y) - 2 \sum_{\substack{X,Y \subseteq \Omega \\ X = A \cap Y}} S(X,Y), \\
&= \| \mathbf{S_A} \|_1 + \| \mathbf{S} \|_1 - 2 \| \mathbf{S} \circ \mathbf{S_A} \|_1 .
\end{aligned}$$

Since the $L^1$ norm of any specialization matrix is $N$, and remembering that $\rho = 2N - 2$ when $p = 1$, equation (7) is retrieved:

$$\begin{aligned}
\| \mathbf{S} - \mathbf{S_A} \|_1 &= 2N + 2 \| \mathbf{S} \circ \mathbf{S_A} \|_1, \\
\Leftrightarrow d_1(m, m_A) &= \frac{N - \| \mathbf{S} \circ \mathbf{S_A} \|_1}{N - 1}.
\end{aligned}$$

Equation (8) is obtained from equation (7) using a classical algebra result. □

In terms of computation time, equation (7) should be preferred. Specialization matrices have $3^n$ non-null elements. The Hadamard product can be restricted to the entrywise product of these non-null elements. The complexity of equation (7) is thus:

$$\begin{aligned}
O(3^n) &= O\left(N^{\frac{\log(3)}{log(2)}}\right), \\
&\approx O\left(N^{1.58}\right).
\end{aligned}$$

### 3.3 Distances between any mass functions

We now address the $d_p$ distance computation problem in the general case. An algorithm for optimizing this computation will be introduced. This algorithm relies on the following result:

**Proposition 2.** *Suppose $m$ is a mass function on a frame $\Omega$. Suppose $X \subseteq Y \subseteq \Omega$ and let $z \notin Y$. The following result holds:*

$$m(X|Y) = m(X|Y \cup \{z\}) + m(X \cup \{z\} | Y \cup \{z\}). \tag{9}$$

[3] The Hadamard matrix product is the entrywise product or Schur product. Let $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ be three matrices such that $\mathbf{X} \circ \mathbf{Y} = \mathbf{Z}$, then we have $Z_{ij} = X_{ij} Y_{ij}$, $\forall i$ and $j$.

*Proof.* By definition of evidential conditioning, one has:

$$m(X|Y) = \sum_{\substack{A\subseteq\Omega \\ A\cap Y=X}} m(A) = \sum_{\substack{A\subseteq\Omega \\ A\cap Y=X, z\in A}} m(A) + \sum_{\substack{A\subseteq\Omega \\ A\cap Y=X, z\notin A}} m(A). \tag{10}$$

Let us deal with the first term in equation (10). We need to prove that $A\cap Y = X$ and $z\in A$ if and only if $A\cap(Y\cup\{z\}) = X\cup\{z\}$. Let us prove first that $A\cap Y = X$ and $z\in A$ implies $A\cap(Y\cup\{z\}) = X\cup\{z\}$. If $z\in A$ then

$$A\cap(Y\cup\{z\}) = (A\cap Y)\cup(A\cap\{z\}) = A\cap Y\cup\{z\}.$$

In addition if $A\cap Y = X$, we obtain:

$$A\cap(Y\cup\{z\}) = X\cup\{z\}.$$

Reciprocally, let us now prove that $A\cap(Y\cup\{z\}) = X\cup\{z\}$ implies $A\cap Y = X$ and $z\in A$. Suppose that $X\cup\{z\} = A\cap(Y\cup\{z\}) = (A\cap Y)\cup(A\cap\{z\})$. Since $z\notin Y$, $z\notin A\cap Y$, hence $z\in A\cap\{z\}$, which implies $z\in A$.
In addition, we also have $X\cup\{z\} = (A\cap Y)\cup\{z\}$. Since $z\notin Y$, then $(A\cap Y)\cup\{z\}\setminus\{z\} = A\cap Y$. Again, since $z\notin Y$, then $z\notin X$ and therefore $X\cup\{z\}\setminus\{z\} = X = A\cap Y$. From the above reasoning, we deduce:

$$\sum_{\substack{A\subseteq\Omega \\ A\cap Y=X, z\in A}} m(A) = \sum_{\substack{A\subseteq\Omega \\ A\cap(Y\cup\{z\})=X\cup\{z\}}} m(A) = m(X\cup\{z\}|Y\cup\{z\}).$$

Let us now deal with the remaining term in equation (10). We need to prove that $A\cap Y = X$ and $z\notin A$ if and only if $A\cap(Y\cup\{z\}) = X$. Let us prove first that $A\cap Y = X$ and $z\in A$ implies $A\cap(Y\cup\{z\}) = X$. Suppose $z\notin A$ and $A\cap Y = X$. One can write:

$$\begin{aligned} A\cap(Y\cup\{z\}) &= (A\cap Y)\cup(A\cap\{z\}), \\ &= A\cap Y\cup\emptyset, \\ &= X. \end{aligned} \tag{11}$$

Reciprocally, let us prove that $A\cap(Y\cup\{z\}) = X$ implies $A\cap Y = X$ and $z\notin A$. Suppose that $X = A\cap(Y\cup\{z\}) = (A\cap Y)\cup(A\cap\{z\})$. Since $z\notin Y$, then $z\notin X$. We thus have $z\notin A\cap\{z\}$, which implies $z\notin A$. In addition, this leads to $X = A\cap Y\cup\emptyset = A\cap Y$. From this reasoning, we deduce:

$$\sum_{\substack{A\subseteq\Omega \\ A\cap Y=X, z\notin A}} m(A) = \sum_{\substack{A\subseteq\Omega \\ A\cap(Y\cup\{z\})=X}} m(A) = m(X|Y\cup\{z\}).\square$$

Proposition 2 is especially interesting when it comes to specialization matrix computation as it shows that any element of the matrix can be obtained by adding two other elements belonging to a right-hand column and lower lines. Since the last column vector is equal to $\mathbf{m}$, this matrix can be built incrementally starting from the column with index $N-1$ down to the first column. In each column, we start with the lowest element up to the top one. This procedure is given by algorithm 1.
This fast specialization matrix computation algorithm can be directly used with $m_1 - m_2$ as entry in order to obtain the matrix difference $\mathbf{S_1} - \mathbf{S_2}$. This algorithm

**Algorithm 1** Fast computation of a specialization matrix $\mathbf{S}$

entries : $m$, $N$, $\mathbf{M}$.
$\mathbf{S} \leftarrow \mathbf{0}$, the null matrix.
**for** $X \subseteq \Omega$ **do**
  $S(X,\Omega) \leftarrow m(X)$
**end for**
**for** $Y \subsetneq \Omega$ (following the decreasing binary order) **do**
  **for** $X \subseteq Y$ (following the decreasing binary order) **do**
    **if** $M(X,Y) > 0$ **then**
      choose $z \in \bar{Y}$.
      $S(X,Y) \leftarrow S(X,Y \cup \{z\}) + S(X \cup \{z\}, Y \cup \{z\})$.
    **end if**
  **end for**
**end for**
return $\mathbf{S}$.
**End**

can also compute recursively distance $d_p$ by updating its value each time a new element $S_1(X,Y) - S_2(X,Y)$ is obtained. Given the definition of matrix $\mathbf{M}$, there are $3^n$ loops in algorithm 1. Similarly to the previous subsection, the distance $d_p$ computation complexity for any mass functions is thus $O\left(N^{1.58}\right)$. Figure 1 illustrates the computation time ease induced by algorithm 1 as compared to the computation time when using equation (5). These results were obtained using a laptop with an Intel® centrino2 2.53 GHz CPU and GNU Octave© programming environment. It can be seen that the log-ratio of computation times is linear with respect to $n$ which is compliant with the claim that the complexity dropped from $O\left(N^3\right)$ to $O\left(N^{1.58}\right)$.

Concerning distance computation, it should also be noted that algorithm 1 does not only improve the time-complexity but also the memory occupation. Indeed, it is unnecessary to store the whole matrix $\mathbf{S}_1 - \mathbf{S}_2$ when computing a specialization distance because some colums will never be used again and one can anticipate that.

Most of state-of-the-art evidential metrics[4], such as Jousselme distance, resort to a product between a matrix and a mass function vector. Their complexity is thus $O\left(N^2\right)$ (using naive programming). Consequently, we have succeeded in making the $L^p$ norm-based specialization distances at least as attractive as other evidential metrics in terms of computation time.

## 4 Conclusion

In this article, several methods for a faster computation of $L^p$ norm-based specialization distances are introduced. Initially, such distances are computed in $O\left(N^3\right)$ with $N$ the size of the power set of the frame of discernment. We provide an algorithmic way to reduce this complexity to $O\left(N^{1.58}\right)$ in the general case. In case of categorical mass functions, the complexity is just $O(1)$.

Using these approaches, $L^p$ norm-based specialization distances become usable tools for several potential applications. In particular, we plan to tackle mass function approximation problems using specialization distance minimization. The

[4] Perry and Stephanou [14] introduced a full metric with $O(N)$ complexity but it fails to grasp structural mass function aspects (see [10]).

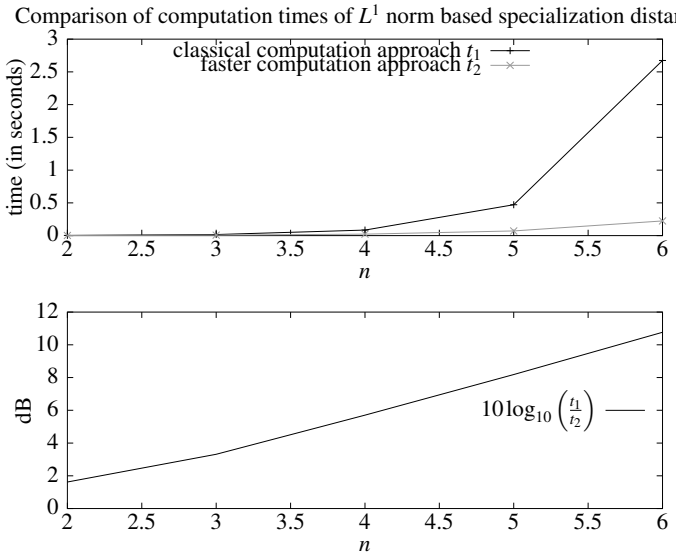


**Fig. 1.** Comparison of computation times using the classical approach (equation (5)) and using our faster approach (algorithm 1).

hope is that the one-of-a-kind properties of such distances will help resolving approximation problems more efficiently.

## References


1. Bauer, M.: Approximation algorithms and decision making in the Dempster-Shafer theory of evidence an empirical study. International Journal of Approximate Reasoning 17(23), 217 – 237 (1997)
2. Cuzzolin, F.: $L_p$ consonant approximations of belief functions. Fuzzy Systems, IEEE Transactions on 22(2), 420–436 (April 2014)
3. Cuzzolin, F.: Geometry of Dempster's rule of combination. IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics 34(2), 961 – 977 (2004)
4. Cuzzolin, F.: Geometric conditioning of belief functions. In: Proceedings of BELIEF 2010, International workshop on the theory of belief functions. pp. 1–6. Brest, France (2010)
5. Cuzzolin, F.: On consistent approximations of belief functions in the mass space. In: Liu, W. (ed.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Lecture Notes in Computer Science, vol. 6717, pp. 287–298. Springer Berlin Heidelberg (2011)
6. Dempster, A.: Upper and lower probabilities induced by a multiple valued mapping. Annals of Mathematical Satistics 38, 325 – 339 (1967)
7. Denœux, T.: Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. Artificial Intelligence 172, 234 – 264 (2008)

8. Denœux, T., Masson, M.H.: EVCLUS: Evidential clustering of proximity data. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics 34(1), 95–109 (2004)
9. Elouedi, Z., Mellouli, K., Smets, P.: Assessing sensor reliability for multisensor data fusion within the transferable belief model. IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics 34(1), 782 – 787 (2004)
10. Jousselme, A.L., Maupin, P.: Distances in evidence theory : Comprehensive survey and generalizations. International Journal of Approximate Reasoning 53, 118 – 145 (2012)
11. Klein, J., Colot, O.: Automatic discounting rate computation using a dissent criterion. In: Proceedings of BELIEF 2010, International workshop on the theory of belief functions. pp. 1 – 6. Brest. France (2010)
12. Loudahi, M., Klein, J., Vannobel, J.M., Colot, O.: New distances between bodies of evidence based on dempsterian specialization matrices and their consistency with the conjunctive combination rule. International Journal of Approximate Reasoning 55, 1093 – 1112 (2014)
13. Mercier, D., Quost, B., Denœux, T.: Refined modeling of sensor reliability in the belief function framework using contextual discounting. Information Fusion 9, 246 – 258 (2008)
14. Perry, W., Stephanou, H.: Belief function divergence as a classifier. In: Intelligent Control, 1991., Proceedings of the 1991 IEEE International Symposium on. pp. 280–285 (Aug 1991)
15. Schubert, J.: Clustering decomposed belief functions using generalized weights of conflict. International Journal of Approximate Reasoning 48, 466 – 480 (2008)
16. Shafer, G.: A mathematical theory of evidence. Princeton University Press (1976)
17. Smets, P.: The application of the matrix calculus to belief functions. International Journal of Approximate Reasoning 31, 1–30 (2002)
18. Sunberg, Z., Rogers, J.: A belief function distance metric for orderable sets. Information Fusion 14, 361 – 373 (2013)
19. Zouhal, L.M., Denœux, T.: An evidence-theoretic k-nn rule with parameter optimisation. IEEE Transactions on Systems, Man and Cybernetics. Part C: Application and reviews 28(2), 263 – 271 (1998)