



HAL
open science

Study of discounting methods applied to canonical decomposition of belief functions

Romain Guyard, Véronique Cherfaoui

► **To cite this version:**

Romain Guyard, Véronique Cherfaoui. Study of discounting methods applied to canonical decomposition of belief functions. 21st International Conference on Information Fusion (FUSION 2018), Jul 2018, Cambridge, United Kingdom. pp.2505-2512, 10.23919/ICIF.2018.8455350 . hal-01897865

HAL Id: hal-01897865

<https://hal.science/hal-01897865>

Submitted on 17 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Study of discounting methods applied to canonical decomposition of belief functions ¹

Romain Guyard¹, Veronique Cherfaoui¹

Abstract—In Dempster-Shafer theory the discounting operation can be used to weak the belief according to the reliability of the source of information. This is usually done by modifying basic belief assignment also called mass function. The canonical decomposition- well adapted for some combination rules- is an other way to represent the belief of an agent. We propose in this paper to focus on discounting methods that are directly applied to the canonical decomposition and we compare them to methods previously studied in the literature. Then we propose an equation that implements the discounting on canonical decomposition strictly equivalent to the classical discounting method. This approach reduces the computation cost significantly. Finally, we illustrate the validity of the method by demonstrating the convergence of the distributed data fusion algorithm using an operator based on the cautious rule and discounting, using an application on a distributed data fusion algorithm, we demonstrate the convergence when the cautious operator is used.

Keywords: Dempster-Shafer theory, canonical decomposition, discounting

I. INTRODUCTION

Dempster-Shafer theory introduced in [1] is often used to model both imprecision and uncertainties of an agent or a data source thanks to belief functions that can be expressed by different representations. The most common one at credal level is the mass function. Mass values are the quantity of knowledge spread among the subsets of the frame of discernment. There are other representations such as belief (bel), plausibility (pl) and commonality (q) functions. This paper focuses on the conjunctive decomposition [2], [3] of the belief functions. This decomposition produces values known as weights that represent the quantity of ignorance regarding one particular element.

Several algorithms use a discounting operation that reduces the importance of data coming from less trusted sources [4] or older sources [5]. Classically the discounting operation modifies masses and this has been well studied in [1], [6]. In some cases, to improve simplicity and efficiency, the discounting operation could be preferred to be applied on weights. For instance, the distributed data fusion algorithm presented in [7] is based on the cautious operator proposed in [3]. The cautious operator uses the conjunctive decomposition as input and output. To avoid heavy computations due to conversions between masses and weights, the discounting is applied on weights. This operation is necessary to guaranty the self-stabilization property. We propose to study more attentively the discounting on the weights.

This paper first introduces two discounting methods based on weight representation that try to follow the behavior of the classical discounting. Then a comparison with mass discounting shows flaws in the weight-based discounting methods. We introduce then an original method to compute the well studied discounting directly on the canonical decomposition with a reduced amount of computations and thus make it implementable in low computing capacity devices. Finally the convergence of the algorithm from [7] is proven when the cautious operator is used in order to perform distributed data fusion .

II. DISCOUNTING USING CANONICAL DECOMPOSITION OF BELIEF FUNCTIONS

A. Mass functions and weight representation

Belief functions proposed by Dempster and Shafer are similar to probabilities but instead of splitting amount of belief on singletons, quantities, called masses, are split among subsets. Let be a variable ω , taking values in a finite set Ω called frame of discernment, 2^Ω is the powerset of Ω , $A \subseteq \Omega$ is an element of 2^Ω . The quantity $m(A)$ is interpreted as the part of the belief allocated to the hypothesis “the answer ω is in the subset A of Ω ”. The equation 1 shows the mass function property.

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

The mass $m(\Omega)$ represents the ignorance. Every A of Ω such that $m(A) > 0$ is a “focal element” of m . A “simple mass function” has one or two focal element including Ω . If Ω is the only focal element then the mass function is called “vacuous” and represents total ignorance. If Ω is not a focal element then the mass function is called “dogmatic”. In this paper we assume that mass functions are always non-dogmatic which is always the case since sensors are always imperfect.

The commonality of A is noted $q(A)$ and is defined by the formula 2. The commonality of A represents the quantity of knowledge on A and on every set that includes A . Commonalities can be used as an equivalent representation of the mass function.

$$q(A) = \sum_{B \supseteq A} m(B) \quad (2)$$

Another representation of belief functions used in this paper is the canonical decomposition presented by [2]. The result of the decomposition is a weight value for each subset $A \subseteq$

¹ The authors are with Sorbonne Universités, Université de Technologie de Compiègne, CNRS Heudiasyc UMR 7253, France

$\Omega, A \neq \Omega$ noted $w(A)$. The equation 3 shows how to compute weight from masses. One can see that this function requires exponential computations and thus its usage should be avoid in systems with limited computational capacity.

$$w(A) = \prod_{B \supseteq A} q(B)^{(-1)^{|B|-|A|+1}} \quad (3)$$

To convert weights to masses $|2^\Omega|$ simple support belief functions are created as shown in equation 4. Then they are all combined using the conjunctive combination operator shown in 5.

$$\begin{cases} \mu(A) = 1 - w(A) \\ \mu(\Omega) = w(A) \end{cases} \quad (4)$$

$$(m_1 \odot m_2)(A) = \sum_{B \cap C = A} m_1(B)m_2(C) \quad (5)$$

B. Belief combination

In order to improve the knowledge and reduce uncertainties, mass functions can be combined. The well-known method of combination is the Dempster conjunctive rule shown in equation 6.

$$\begin{cases} (m_1 \oplus m_2)(\emptyset) = 0 \\ A \neq \emptyset & (m_1 \oplus m_2)(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B)m_2(C) \\ \text{where } K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C) \end{cases} \quad (6)$$

This rule requires the independence of sources. If it is not the case (for example, when a source gives information which is generated using data from other sources) the fusion operator has to be idempotent. To fulfill this requirement, the cautious operator has been proposed in [3]. This operator is applied on weight using the operator minimum as shown in equation 7.

$$(w_1 \otimes w_2)(A) = \min(w_1, w_2) \quad (7)$$

C. Discounting operation

The discounting proposed by Shafer in [1] and called in this paper “*mass-linear-discounting*” is given by Equation 8. $m^\alpha(A)$ is the mass linear discounted with a factor of $\alpha \in [0, 1]$. The value of α depends on the reliability of the source.

$$\begin{cases} A \neq \Omega & m^\alpha(A) = \alpha m(A) \\ & m^\alpha(\Omega) = (1 - \alpha) + \alpha m(\Omega) \end{cases} \quad (8)$$

Some variants of mass discounting was proposed by Mercier in [8] to take the context into account and in [9] to consider the temporal evolution of information.

In [10], the authors proposed a discounting directly applied on weights. It consists in adding a value $\beta \in \mathbb{R}^+$ to the weights as described in Equation 9. We will called it “*weight-summed-discounting*”.

Let $w_{summed}^\beta(A)$ be the *weight-summed-discounting* of $w(A)$ with a factor of $\beta \in \mathbb{R}^+$

$$w_{summed}^\beta(A) = \min(1, w(A) + \beta) \quad (9)$$

This discounting on weights has been chosen because it works with an algorithm using cautious operator based on weight representation. Also, discounting the weights avoids the transformation from weight to mass representation and then go back to weights. These conversions require exponential computation and thus it becomes hard to implement it in embedded devices. Even if this discounting operation has been chosen to fulfill the convergence of the algorithm [11], there is no study to our knowledge that discuss its behavior.

It is easy to notice that the classic *mass-linear-discounting* is a proportional discounting and the *weight-summed-discounting* is additive. Thus we propose two other proportional discounting operations that are applied on weights. To decrease the quantity of information of a belief function it is necessary to increase the ignorance represented directly by the weights.

Let $w_{WLinear}^\gamma(A)$ be the *weight-linear-discount* of $w(A)$ with a factor of $\gamma > 1$ computed by the equation 10.

$$w_{WLinear}^\gamma(A) = \min(1, \gamma w(A)) \quad (10)$$

In order to follow the behavior of the *mass-linear-discounting*, it may be better to reduce the weight of a proportion of $1 - w(A)$ instead of $w(A)$. This discounting will be called “*one-minus-weight-linear-discounting*”.

Let $w_{1-WLinear}^\delta(A)$ be the *one-minus-weight-linear-discount* of $w(A)$ with a factor of $\delta > 1$ computed by the equation 11.

$$w_{1-WLinear}^\delta(A) = \min(1, w(A) + \delta(1 - w(A))) \quad (11)$$

The discounting operators must keep the masses positive and the sum of all masses equals to one. The conjunctive operator conserves those properties. Thus if the equations 4 are mass functions then the underlying weight function is also a belief function. In the case the belief function is not separable, weights are greater than 1 thus the masses of support functions computed in equation 4 are negative. Hence the minimum operator of the three last equations has been added. The impact of this minimum should be analyzed in further study.

In [12], the authors proposed a contextual discounting that will be compared in further studies.

III. ANALYSIS AND COMPARISON OF DISCOUNTING METHODS

This section discusses the discounting methods presented and proposed in the previous section. A “good” discounting operation reduces the quantity of knowledge evenly on focal elements. Since it is easier to interpret a belief function with mass values and for homogeneity, the following always show masses in figures even if the computation is done on weights.

A. General behavior differences using examples

In order to analyze the behavior of discounting operations, we propose to mitigate information several times using the same discounting operation. Graphs 1 to 4 show all discounting methods applied multiple times with different initial states. This method considers that there is no new information added to the system to focus on the behavior of the discounting and not on any potential fusion operators.

We will consider four examples to highlight differences between discounting methods. This is not exhaustive analysis but it will be enough to conclude on which function should be used.

For each example, the four discounting are given in the following order:

- top left: *mass-linear-discounting* with $\alpha = 0.3$.
- top right: *weight-summed-discounting* with $\beta = 0.1$.
- bottom left: *weight-linear-discounting* with $\gamma = 1.3$.
- bottom right: *one-minus-weight-linear-discounting* with $\delta = 0.3$.

In all examples we will consider $\Omega = \{e_1, e_2, e_3\}$. Only mass that are relevant for the analysis are plotted.

1) *Example 1*: The figure 1 shows an example with the initial state: $m(\{e_1\}) = 0.5$ $m(\{e_2\}) = 0.4$ and $m(\Omega) = 0.1$.

In this example we can see that the *mass-linear-discounting* (top left) is smoothly decreasing knowledge with a proportional amount of its value. Masses decrease indefinitely whereas the *weight-summed-discounting* (top right) reduces masses linearly until reaching 0 with rough angles. Only the *one-minus-weight-linear-discounting* (bottom right) looks as smooth as the first graph. Smoothness is important for transitivity reason. For instance, as discussed in [9] for temporal discounting, two discounting of one second should be equivalent to one discounting twice as strong. This property is only verified for the *mass-linear-discounting* and the *one-minus-weight-discounting*.

2) *Example 2*: The figure 2 shows an example with the initial state: $m(\{e_1\}) = 0.2$ $m(\{e_2\}) = 0.1$ $m(\{e_1, e_2\}) = 0.6$ and $m(\Omega) = 0.1$.

In this example we can see in both weigh summed discounting (top right) and *weight-linear-discounting* (bottom left) that the mass associated with the subset $\{e_1, e_2\}$ starts by increasing before decreasing. This is caused by the decrease of $m(\{e_1\})$ and $m(\{e_2\})$ also increases $m(\{e_1, e_2\})$. $m(\{e_1, e_2\})$ is also decreased but not enough to compensate this effect. A discounting operation should decrease knowledge and never increase it by definition. Thus, this behavior should be avoided. In this example, the *mass-linear-discounting* and *one-minus-weight-linear-discounting* have the same behavior.

3) *Example 3*: The figure 3 shows an example with the initial state: $m(\{e_1\}) = 0.2$ $m(\{e_2\}) = 0.2$ $m(\{e_1, e_3\}) = 0.2$ $m(\{e_2, e_3\}) = 0.2$ and $m(\Omega) = 0.2$.

As expected in this example where all masses are equal at the initial state, the *mass-linear-discounting* (top left) shows similar decreasing for each mass. That is not the case with the three others. This example shows that an element with an initial mass equal to 0 ($\{e_3\}$ in this example) can increase

as well. With those discounting functions an element that has not been observed can be non null. This behavior should be avoided.

4) *Example 4*: The figure 4 shows an example 4 with initial state: $m(e_1) = 0.4$ $m(e_2) = 0.4$ $m(\{e_1, e_2\}) = 0.1$ and $m(\Omega) = 0.1$.

In this last example, we can see that in the three discounting based on weights, the conflict can raise significantly. In order to reduce this conflict, the mass function can be normalized but this operation can only be done using masses. This remark nullifies the interest of using discounting on weights. We can also observe a slight increase of $m(\{e_1, e_2\})$ on the graph of the *one-minus-weight-linear-discounting*.

5) *Comparison conclusion*: Those examples show that the *mass-linear-discounting* has no misbehavior and are easy to interpret. The three other discounting can cause issues in most scenarios. They can increase masses even if there were null. Moreover masses that are equal before can be different after discounting. The power of the discounting also depends of the number of focal elements which lead to a linear by parts discounting. The mass disappears after few iterations whereas the *mass-linear-discounting* always carries a small part of a data over the multiple iterations. Finally the three weight discounting operations generate mass on the empty set which may cause issues in some situations.

In most cases the *mass-linear-discounting* should be preferred. But it has two problems that may reduce its interest for an implementation in real scenarios. Firstly it is applied on masses and in some application using weights this will force conversion and then a lot of computations. The second issue is when the discounting is used in combination with the cautious operator as proposed in [11]. The convergence of a distributed algorithm has been proven in [10] only if the discounting increase weights and *mass-linear-discounting* is not proved to be in this case. Solutions to both this problems will be demonstrated in the next part of this document.

IV. MASS-LINEAR-DISCOUNTING APPLIED ON WEIGHTS

This part will demonstrate an original equation to use *mass-linear-discounting* on weights with a reduced amount of computation than doing representation conversions.

A. General mass functions properties

Usually weights are defined in \mathbb{R}^+ and could be null. In the following of this paper, we need to be able to divide by a weight or a product of weights. The lemma 1 shows that weights will not be null in our case.

Lemma 1.

Weights of non dogmatic mass function can not be null.

Proof: By definition, commonalities are sum of masses. This sum always includes $m(\Omega)$ because Ω includes all its subsets by definition. Non-dogmatic mass function has always Ω as a focal element. Every masses are strictly greater than 0 then commonalities are also greater than 0 strictly. Weights

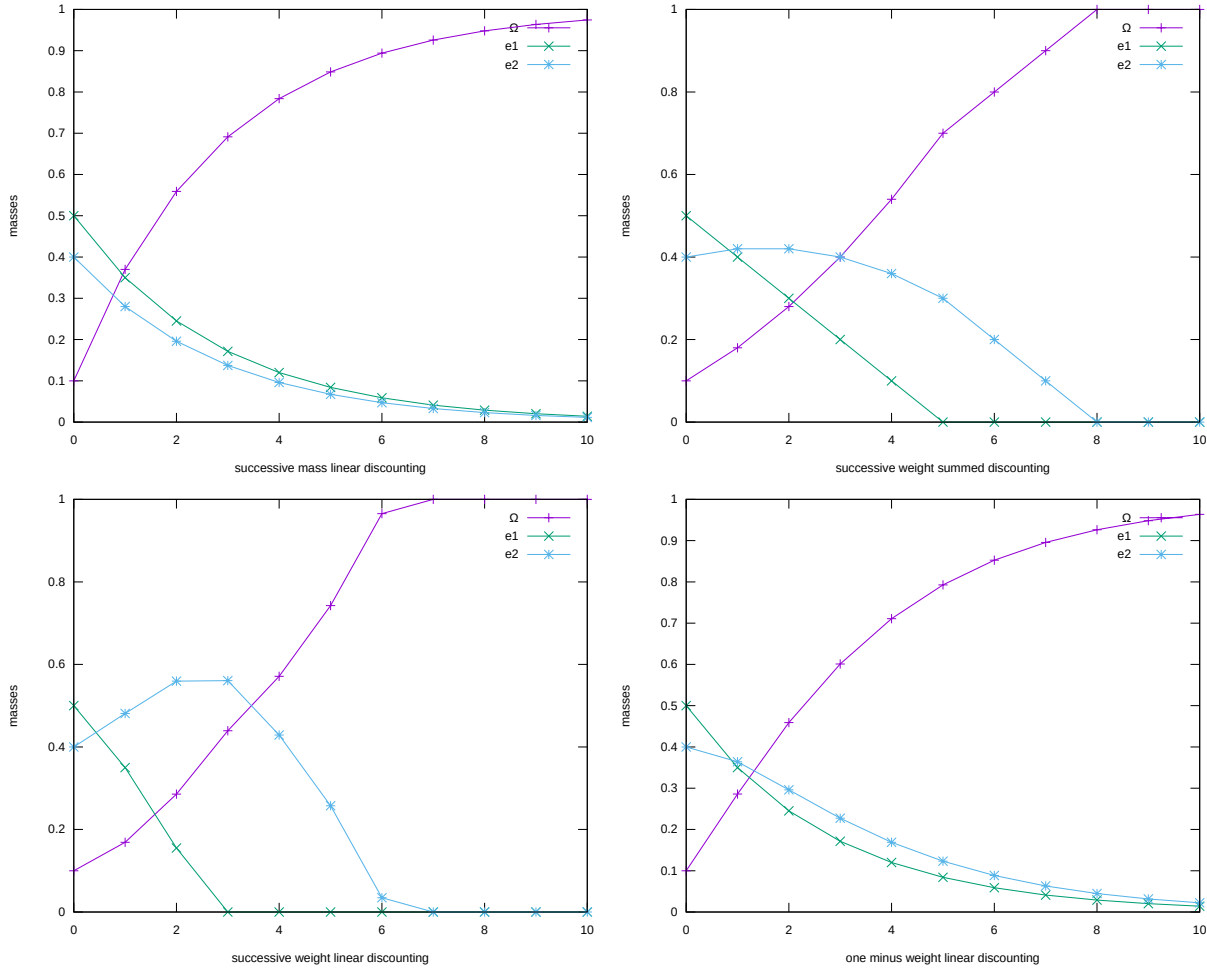


Figure 1: Example 1: Discounting shapes

are defined by a product of commonalities and since no commonality are null then the product is not null. Thus weights of non dogmatic belief function can not be null. ■

Lemma 2.

$$\forall A \subseteq \Omega, B \subseteq \Omega \sum_{B \supseteq A} (-1)^{|B|-|A|+1} = 0$$

The lemma 2 is proven using Pascal equivalence. This equation shows that there is as many even cardinalities than odd cardinalities in the power set of a set.

Proof: Let $A \subseteq \Omega, B \supseteq A$ and 2^Ω the power set of Ω .

Let P_{Even}^Ω the set of elements of 2^Ω having even cardinal.

Let P_{Odd}^Ω the set of elements of 2^Ω having odd cardinal.

Let $x = |P_{Even}^\Omega| - |P_{Odd}^\Omega|$.

Let $C \subseteq P(\Omega)$ with $\forall D \in C, B \supseteq D$ and $\Omega' = \Omega \setminus A$.

Let $E \subset \Omega', E_{Even}$ the set of E even and E_{Odd} the set of E odd.

$\forall E \subset \Omega', A \cup E \in C$

Thus, there is a bijective function f :

$$f: \begin{array}{l} 2^{\Omega'} \mapsto C \\ E \mapsto A \cup E \end{array}$$

$$|E| = |f(E)| - |A|$$

$$\text{Thus } \begin{cases} E \in E_{Even} & |f(E)| - |A| + 1 \equiv |E| + 1 \equiv 1 \pmod{2} \\ E \in E_{Odd} & |f(E)| - |A| + 1 \equiv |E| + 1 \equiv 0 \pmod{2} \end{cases}$$

So $|E_{Even}| = |P_{Odd}^\Omega|$ and $|E_{Odd}| = |P_{Even}^\Omega|$

Let $a \in [0, |\Omega'|[$ and E_a the set of every $E \subset E_a, |E| = a$

$$|E_a| = \binom{|\Omega'|}{a} \text{ with the binomial coefficient: } \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Let Σ_{Even} the sum of $|E_a|$ for every even a and Σ_{Odd} the sum of $|E_a|$ for every odd a .

$$\begin{cases} \Sigma_{Even} = \sum_{i=0}^{i=\lfloor \frac{|\Omega'|}{2} \rfloor} \binom{|\Omega'|}{2i} \\ \Sigma_{Odd} = \sum_{i=0}^{i=\lfloor \frac{|\Omega'|}{2} \rfloor - 1} \binom{|\Omega'|}{2i+1} \end{cases}$$

$$\Sigma_{Even} = \binom{|\Omega'|}{0} + \binom{|\Omega'|}{2} + \binom{|\Omega'|}{4} + \dots + \binom{|\Omega'|}{|\Omega'| - 2}$$

Using Pascal equivalence we have:

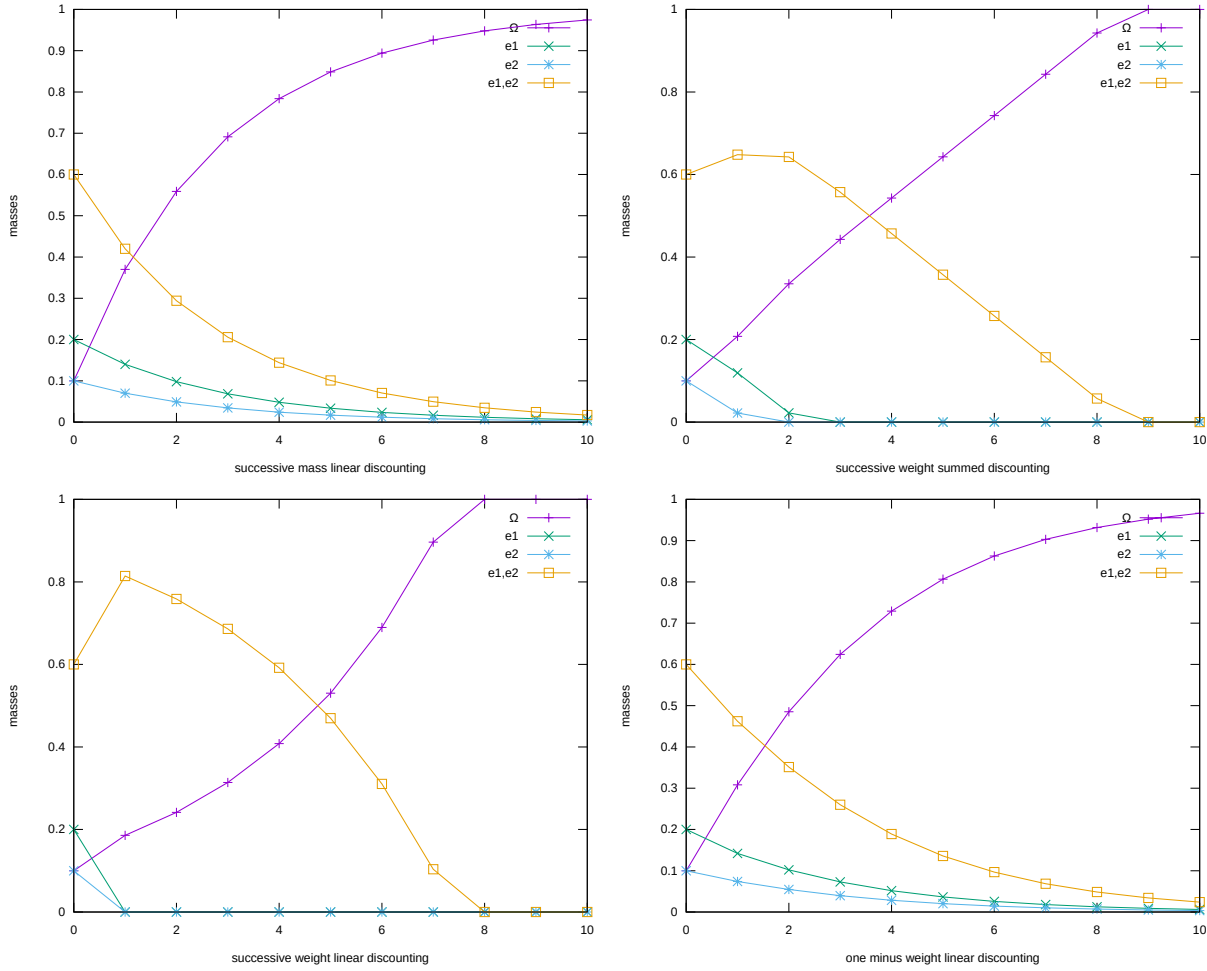


Figure 2: Example 2: Monotonousness of discounting

$$\begin{aligned}
\Sigma_{Even} &= \left[\binom{|\Omega'| - 1}{0} \right] \\
&+ \left[\binom{|\Omega'| - 1}{1} + \binom{|\Omega'| - 1}{2} \right] \\
&+ \left[\binom{|\Omega'| - 1}{3} + \binom{|\Omega'| - 1}{4} \right] \\
&\dots + \left[\binom{|\Omega'| - 1}{|\Omega'| - 3} + \binom{|\Omega'| - 1}{|\Omega'| - 2} \right] \\
&+ \left[\binom{|\Omega'| - 1}{|\Omega'| - 1} \right] \\
\Sigma_{Even} &= \sum_{i=0}^{|\Omega'| - 1} \binom{|\Omega'| - 1}{i}
\end{aligned}$$

Symmetrically:

$$\begin{aligned}
\Sigma_{Odd} &= \left[\binom{|\Omega'| - 1}{0} + \binom{|\Omega'| - 1}{1} \right] \\
&+ \left[\binom{|\Omega'| - 1}{2} + \binom{|\Omega'| - 1}{3} \right] \\
&\dots + \left[\binom{|\Omega'| - 1}{|\Omega'| - 2} + \binom{|\Omega'| - 1}{|\Omega'| - 1} \right] \\
\Sigma_{Odd} &= \sum_{i=0}^{|\Omega'| - 1} \binom{|\Omega'| - 1}{i}
\end{aligned}$$

Thus $\Sigma_{Even} = \Sigma_{Odd}$ So $|E_{Even}| = |E_{Odd}|$ So $|P_{Odd}^\Omega| = |P_{Even}^\Omega|$ So $x = 0$ So $\sum_{B \supseteq A} (-1)^{|B| - |A| + 1} = 0$

We can now also prove that: ■

Corollary 3.

$$\forall A \subseteq \Omega, \forall x \in \mathbb{R}^* \\
\prod_{B \supseteq A} (x \cdot q(B))^{(-1)^{|B| - |A| + 1}} = \prod_{B \supseteq A} q(B)^{(-1)^{|B| - |A| + 1}}$$

Thus it is possible to simplify every non null real number in the computation of weights

Theorem 4.

$$m(\Omega) = q(\Omega) = \prod_{A \subset \Omega} w(A)$$

Proof: By definition, to compute masses with weights, we have to create simple mass functions for each subset of Ω defined by the equation 4. We can see that the each weight became a mass associated with Ω . The second step of the conversion is to combine the simple mass functions with the conjunctive operator show in equation 5. There is no other set that includes Ω except itself. Thus the result of the conjunctive combination for Ω is the product of every weight and then it is the mass of Ω . ■

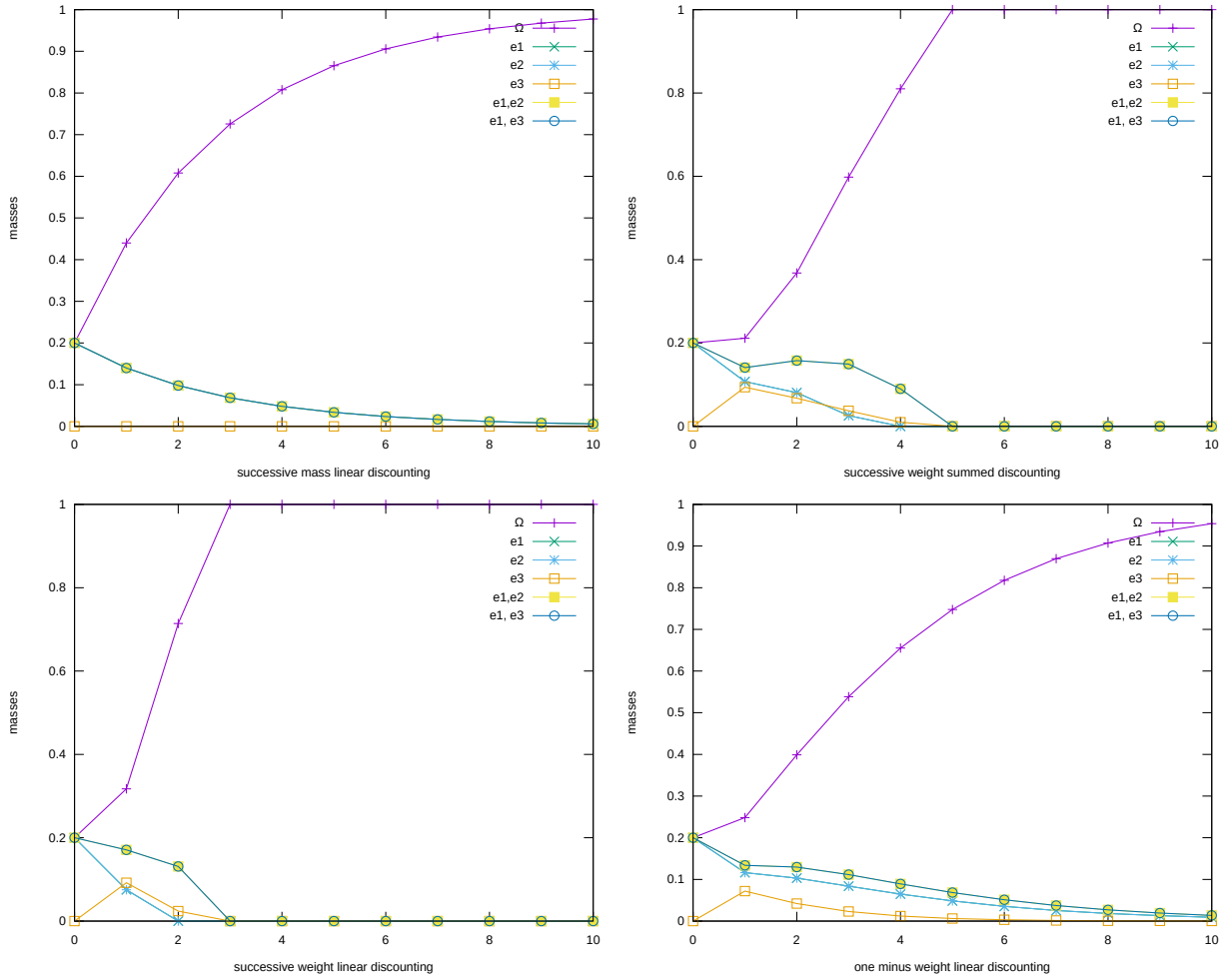


Figure 3: Example 3: Change of focal element number

Conjecture 5.

$$q(A) = \prod_{B \not\subseteq A} w(B) = \frac{\prod_{B \subset \Omega} w(B)}{\prod_{B \supseteq A} w(B)}$$

The theorem 4 validates the conjecture 5 in the particular case of $A = \Omega$. The second part of the equality is founded by multiplying the numerator and the denominator by $m(\Omega)$ as shown in theorem 4.

The first equality requires less computations if $|A| < \frac{|\Omega|}{2}$, otherwise the second equality requires less. The numerator of the second formula $\prod_{B \subset \Omega} w(B)$ is in fact $m(\Omega)$. As it doesn't depend of A it can be computed once and used in every other computation.

B. Simplified version of mass-linear-discounting applied to weights

In this part we will build an mathematical equivalent of the *mass-linear-discounting* usually done on mass functions using weights. In order to reach this propriety, we convert weights to masses, we apply the *mass-linear-discounting* and then we go back to weights. Since going from weights to masses and from masses to weights are very heavy on computation, we

will use theorems from the previous section to simplify the equations. Let respectively be $m^\alpha(A)$, $q^\alpha(A)$ and $w^\alpha(A)$ the mass, commonality and weight of A discounted with a factor of $\alpha \in [0, 1]$.

Theorem 6. *If $\alpha = 0$ then*

$$w^\alpha(A) = 1$$

Proof: Using the formula 8 we are looking for a mass function where the only focal element is Ω . $m(\Omega)$ is defined by the lemma 1 : $m(\Omega) = \prod_{A \subset \Omega} w(A) = \prod_{A \subset \Omega} 1 = 1$. If $m(\Omega)$ is equal to 1 then all other masses are null. P. Smets has proven in [2] that there is only one representation of a belief function using weights. Thus since we have found one then it is the only one. ■

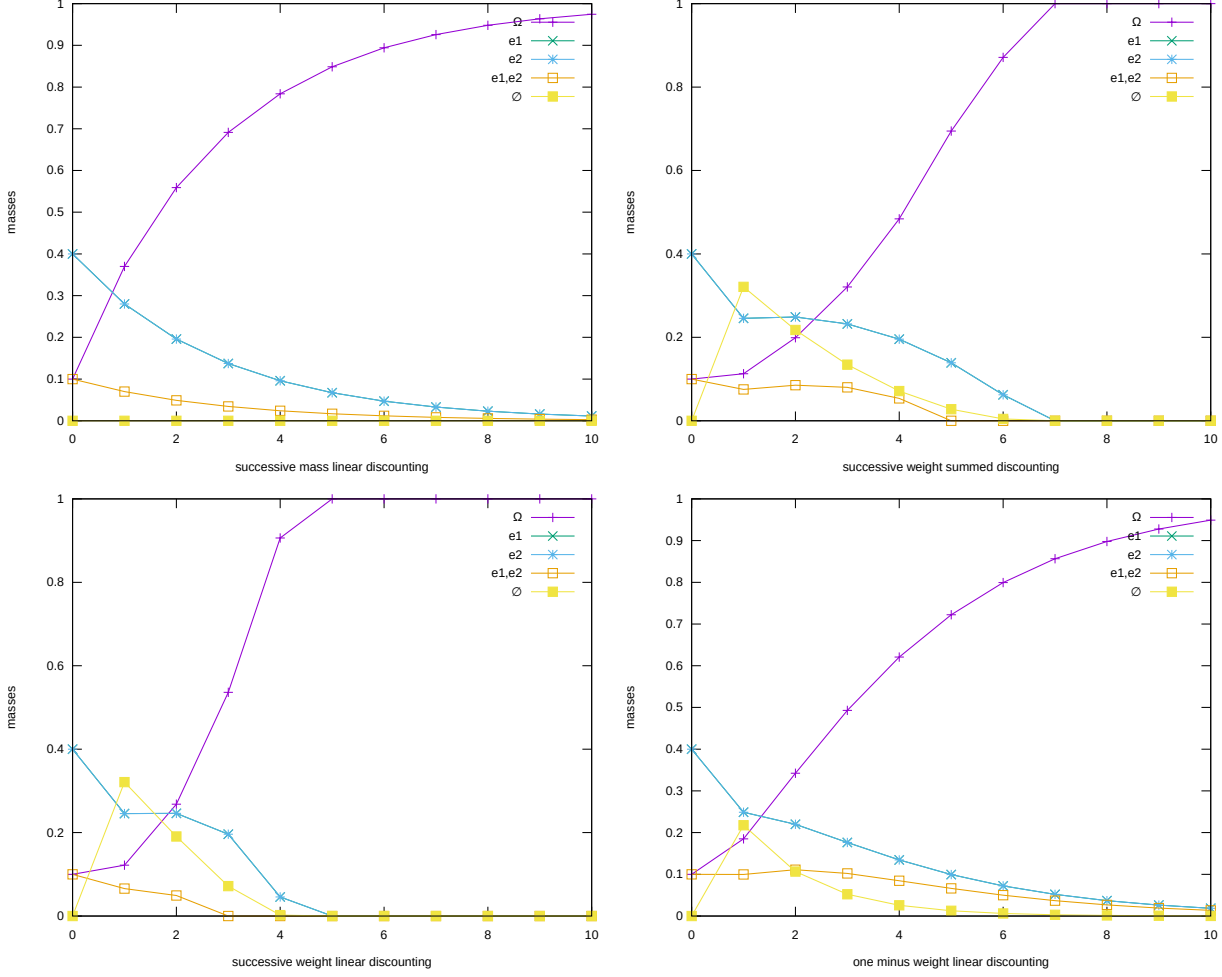


Figure 4: Example 4: Discounting method creating conflict

Theorem 7. if $\alpha \neq 0$ and if the conjecture 5 is true then

$$w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \not\supseteq B} w(C) + \frac{1-\alpha}{\alpha} \right]^{(-1)^{|B|-|A|+1}}$$

$$w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \supseteq B} \frac{1}{w(C)} + \frac{1-\alpha}{\alpha \prod_{D \subset C} w(D)} \right]^{(-1)^{|B|-|A|+1}}$$

Proof: If we use the commonality definition equation 2 and we apply the discounting of the formula 8 we have:

$$q^\alpha(A) = \sum_{B \supseteq A, B \neq \Omega} \alpha m(B) + (1-\alpha) + \alpha m(\Omega)$$

By getting out α of the sum we found the definition of the commonality of A .

$$q^\alpha(A) = \alpha [q(A) - m(\Omega)] + (1-\alpha) + \alpha m(\Omega) = \alpha q(A) + (1-\alpha)$$

Using the conjecture 5:

$$q^\alpha(A) = \alpha \left[\prod_{B \not\supseteq A} w(B) \right] + (1-\alpha)$$

By replacing $q(A)$ by $q^\alpha(A)$ in the equation 3 we have:

$$w^\alpha(A) = \prod_{B \supseteq A} \left[\left(\alpha \prod_{C \not\supseteq B} w(C) \right) + (1-\alpha) \right]^{(-1)^{|B|-|A|+1}}$$

The corollary of the lemma 1 is used to simplify α if $\alpha \neq 0$:

$$w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \not\supseteq B} w(C) + \frac{1-\alpha}{\alpha} \right]^{(-1)^{|B|-|A|+1}}$$

Using the same way with the second part of the equation of the conjecture 5 we have:

$$w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \supseteq B} \frac{m(\Omega)}{w(C)} + \frac{1-\alpha}{\alpha} \right]^{(-1)^{|B|-|A|+1}}$$

$m(\Omega) \neq 0$ because the belief function is non dogmatic. We can simplify $m(\Omega)$ to get the form of 7 as expected:

$$\begin{cases} w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \not\supseteq B} w(C) + \frac{1-\alpha}{\alpha} \right]^{(-1)^{|B|-|A|+1}} \\ w^\alpha(A) = \prod_{B \supseteq A} \left[\prod_{C \supseteq B} \frac{1}{w(C)} + \frac{1-\alpha}{\alpha m(\Omega)} \right]^{(-1)^{|B|-|A|+1}} \end{cases}$$

Like the formula 5, there is two ways to have the same result but with different amount of computation depending of the cardinal of A . The first equation requires less computations if $|A| < \frac{|\Omega|}{2}$ and otherwise if $|A| > \frac{|\Omega|}{2}$ then the second equation is faster.

This new formula to compute masses discounting on weights reduce the number of operations needed. The tables Ia and Ib illustrate this difference for an Ω of cardinality 3 and 4. The table Ia shows the number of operation needed at each step of the computation of the discounting by doing the full conversion. Table Ib shows that using the theorem 7 we need fewer computations to have the same result. With a cardinal of $\Omega = 4$, we have a reduction close to a factor of 18.

$ \Omega $	$w()$ to $m()$	discounting	$m()$ to $w()$	Total
3	294	8	34	336
4	1440	16	62	1518

(a) By full conversion

$ \Omega $	Number total of operations
3	26
4	90

(b) Using the theorem 7

Table I: Number of operations needed to discount

More $|\Omega|$ increase more this difference is important. This new formula enables the possibility to implement this discounting in a real embedded device with low computation capabilities.

V. CONVERGENCE USING MASS-LINEAR-DISCOUNTING IN DISTRIBUTED FUSION CONTEXT

In a distributed data fusion algorithm, the result must converge to a common value. This is achieved by fusing step by step with data coming from other nodes. If the fusion function were used alone, a potential transient error could be kept in the final result. Moreover, if the situation evolves and the observation is different from previous data, the algorithm will always keep a bias. To solve this problem, a discounting operation must be done. The algorithm in [11] uses a discounting function that mitigates the importance of old and distant data in a smart cars data fusion network.

The algorithm from [11] has been proved to be convergent in [10]. This proof is valid for any discounting function that follows two rules. First rule : the discounting must be an endomorphism of (\mathbb{K}, \otimes) , ie the result of a discounting of a belief function stays a belief function and the discounting operator must be distributive with \otimes . Second rule : the discounting operator must be expensive on weights. Both rule are easily checked with weight discounting but the second rule has to be checked with the *mass-linear-discounting*.

Theorem 8. *The mass-linear-discounting of a non-dogmatic belief function always increases weights if the discounting coefficient α is positive.*

Proof: The proof of theorem 7 shows that:

$$q^\alpha(A) = \alpha [q(A) - m(\Omega)] + (1 - \alpha) + \alpha m(\Omega) = \alpha q(A) + (1 - \alpha)$$

$q(A) - q^\alpha(A) = (1 - q(A))(1 - \alpha)$ and $q(A) > 0$ if the function is non-dogmatic.

Then $q(A) - q^\alpha(A)$ is always positive if $\alpha > 0$. A weight is a product of commonalities and since $q(A) - q^\alpha(A) > 0$ then $w^\alpha(A)$ increases if $\alpha > 0$. ■

The theorem 8 proves that the mass discounting increase weights and thus using all discounting methods presented in this paper guarantees the convergence of the distributed algorithm.

VI. CONCLUSION

The discounting operations are an important part of algorithms using belief functions. For instance they are responsible of the importance of old and distant information in data fusion

algorithms. Tuning the discounting has to be done to map the time and geographic diffuseness characteristics of the observed event. In this paper, we have proposed two new ways of computing discounting using the canonical decomposition. Then we have analyzed the difference between discounting methods coming both from the literature and proposed in this paper. This study shows some flaws in weight discounting functions using specific scenarios. This analysis shows that all discounting operators defined on weights have behaviors that can be problematic. Finally we propose an original method for computing mass discounting on weights with less computations that are needed to do the conversion between this two representations of belief function. This new method enables the possibility to implement in embedded systems the well studied and flawless discounting while using for example the cautious operator on weights. Future works should extend this method to an application of distributed data fusion in vehicular network to robustly detect road events.

VII. ACKNOWLEDGMENTS

This work was carried out in the framework of the challenge DAPAD (Distributed and Augmented vehicle Perception to support Autonomous Driving) funded by the Labex MS2T, supported by the French Government, through the program ‘‘Investments for the future’’ managed by the National Agency for Research (Reference ANR-11-IDEX-0004-02)

REFERENCES

- [1] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [2] P. Smets, *The canonical decomposition of a weighted belief*. Morgan Kaufmann Publishers Inc., Aug 1995.
- [3] T. Denœux, ‘‘Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence,’’ *Artificial Intelligence*, vol. 172, pp. 234–264, Feb 2008.
- [4] E. Lefevre, O. Colot, and P. Vannoorenbergh, ‘‘Belief function combination and conflict management,’’ *Information Fusion*, vol. 3, no. 2, pp. 149 – 162, 2002.
- [5] M. Kurdej, J. Moras, V. Cherfaoui, and P. Bonnifait, ‘‘Controlling remanence in evidential grids using geodata for dynamic scene perception,’’ *International Journal of Approximate Reasoning*, vol. 55, pp. 355–375, January 2014.
- [6] P. Smets, ‘‘Belief Functions: The Disjunctive Rule of Combination and the Generalized Bayesian Theorem,’’ *SpringerLink*, pp. 633–664, 2008.
- [7] J. Radak, B. Ducourthial, V. Cherfaoui, and S. Bonnet, ‘‘Detecting Road Events Using Distributed Data Fusion: Experimental Evaluation for the Icy Roads Case,’’ *IEEE Trans. Intell. Transp. Syst.*, vol. 17, pp. 184–194, Jan 2016.
- [8] D. Mercier, B. Quost, and T. Denœux, ‘‘Refined modeling of sensor reliability in the belief function framework using contextual discounting,’’ *Inf. Fusion*, vol. 9, pp. 246–258, Apr 2008.
- [9] M. Kurdej and V. Cherfaoui, ‘‘Conservative, Proportional and Optimistic Contextual Discounting in the Belief Functions Theory,’’ *16th International Conference on Information Fusion (FUSION 2013)*, pp. 2012–2018, Jul 2013.
- [10] B. Ducourthial, V. Cherfaoui, and T. Denœux, ‘‘Self-stabilizing Distributed Data Fusion,’’ *Richa A.W., Scheideler C. (eds) Stabilization, Safety, and Security of Distributed Systems. SSS 2012*, vol. 7596, pp. 148–162, Oct 2012.
- [11] B. Ducourthial and V. Cherfaoui, ‘‘Experiments with Self-stabilizing Distributed Data Fusion,’’ *2016 IEEE 35th Symposium on Reliable Distributed Systems (SRDS)*, pp. 289–296, Sep 2016.
- [12] D. Mercier, É. Lefèvre, and F. Delmotte, ‘‘Belief functions contextual discounting and canonical decompositions,’’ *International Journal of Approximate Reasoning*, vol. 53, no. 2, pp. 146–158, 2012.