Possibilistic information fusion using maximal coherent subsets

Sebastien Destercke and Didier Dubois and Eric Chojnacki

***Abstract*—When multiple sources provide information about the same unknown quantity, their fusion into a synthetic interpretable message is often a tricky problem, especially when sources are conflicting. In this paper, we propose to use possibility theory and the notion of maximal coherent subsets (MCS), often used in logic-based representations, to build a fuzzy belief structure that will be instrumental both for extracting useful insight about various features of the information conveyed by the sources and for compressing this information into a unique possibility distribution. Extensions and properties of the basic fusion rule are also studied.**

***Index Terms*—Information fusion, maximal coherent subsets, possibility theory, fuzzy sets, fuzzy belief functions**

I. Introduction

When multiple sources deliver information tainted with uncertainty about some unknown quantity, aggregating this information can be a tedious task, especially when information is conflicting. This problem was first addressed in the framework of probability theory, and still constitutes an active area of research (see [1] for a review).

Some shortcomings of probabilistic methods are emphasized in [2], where it is shown that probabilistic methods tend to confuse randomness with imprecision. The shortcomings of the arithmetic mean (the most used and founded fusion operator for probabilities) are also discussed. Namely it tends to suggest, as being plausible, values none of the sources considered possible.

An alternative approach is to consider other theories of uncertainty, such as imprecise probabilities [3], evidence theory [4] or possibility theory [5]. These theories allow to faithfully model incomplete or imprecise data, a feature that probability theory arguably cannot account for. When it comes to aggregating data from multiple sources, these theories possess far more flexibility in the treatment of conflicting information, mainly due to the flexible use of set-operations (conjunction and disjunction).

In this paper, we focus on uncertainty modeled by possibility distributions, for they can be easily elicited and interpreted as collection of confidence intervals, and are attractive from a computational viewpoint. On the other hand, possibility distributions can sometimes be judged insufficiently expressive in regard with available information (other theories should then be used).

S. Destercke and D. Dubois are with the Institut de Recherche en Informatique de Toulouse (IRIT, CNRS & Université de Toulouse), France (email: desterck@irit.fr;dubois@irit.fr).

E. Chojnacki is with the Institut de Radioprotection et de Sûreté Nucléaire, Cadarache, France (email: eric.chojnacki@irsn.fr).

Many fusion rules have been proposed to aggregate conflicting possibility distributions, using combinations of conjunction and disjunction operations, possibly exploiting additional data (e.g. reliability of sources); see [6] for review. Most of these proposals result in a single final possibility distribution built from the original distributions provided by the sources, thus eliminating inconsistency between them during the fusion process. In this paper, we explore a fusion method based on maximal coherent subsets (a natural way of coping with inconsistent knowledge basis in logic [7]). The proposed fuzzy information fusion method does not preserve the consonance property of possibility distribution and produces a fuzzy belief structure.

The use of the notion of maximal coherent subsets in uncertainty theories is not new: in the theory of imprecise probabilities, the notion is thoroughly studied by Walley in [8]. It is also used in [9] as a step in a fusion process, and the result of the rule proposed in [10] can be seen as a weighted average of maximal coherent subsets of sources. In the context of evidence theory, the notion of maximal coherent subsets is used in [11] to detect subgroups of coherent sensors.

The paper is divided as follows: theoretical preliminaries are introduced in Section II and Section III provides a quick review of existing possibilistic fusion rules. Section IV then explains how maximal coherent subsets are applied to obtain the fuzzy belief structure. Some properties of the proposed method are laid bare in comparison with other fusion rules in Section V. Section VI presents and discusses some means of extracting useful information from this structure, especially a possibility distribution. Finally, Section VII proposes some possible ways of taking into account additional information concerning the sources.

II. Preliminaries

Zadeh introduced the link between fuzzy sets and possibility theory, and he was the first to propose an extension of Shafer belief structures [4] when focal sets are fuzzy sets [12]. Since then, many proposals appeared, for example by Yager [13], Dubois and Prade [14], Yen [15] and Denoeux [16]). This section presents the framework adopted in the paper to handle fuzzy belief functions.

A. Possibility theory

A possibility distribution π is a mapping from a space X to $[0, 1]$ such that $\pi(x) = 1$ for some element x of X, and is formally equivalent to the definition of a normalized fuzzy membership function. One can interpret a quasi-concave possibility distribution on the real line, that is a distribution equivalent to a fuzzy interval, as a set of nested closed intervals, with various confidence levels [17]. From a possibility distribution, possibility and necessity measures are respectively defined as:

$$\Pi(A) = \sup_{x \in A} \pi(x)$$
$$N(A) = 1 - \Pi(A^c)$$

where A^c stands for the complement of A. A possibility degree $\Pi(A)$ quantifies to what extent the event A is plausible, while the necessity degree quantifies the certainty of A, in the face of incomplete information modelled by π. These measures can also be interpreted as probability bounds [18].

An α-cut E^α of the distribution π is defined as the set

$$E^\alpha = \{x | \pi(x) \geq \alpha\}$$

The core $c(\pi)$ and the support $s(\pi)$ of π respectively correspond to E^1 and $lim_{\epsilon \to 0}\ E^\epsilon$

B. Fuzzy belief structure

A belief structure consists of a mapping m from the power set $\wp(X)$ of a space X to $[0, 1]$ such that $\sum_{E \subseteq X} m(E) = 1$, $m(E) \geq 0$ and $m(\emptyset) = 0$. Sets E that have positive mass are called focal sets. From this mapping, we can again define two set-functions, the plausibility and belief functions, which read [4]:

$$Bel(A) = \sum_{\substack{E \in \wp(X) \\ E \subseteq A}} m(E)$$

$$Pl(A) = \sum_{\substack{E \wp(X) \\ E \cap A \neq \emptyset}} m(E) = 1 - Bel(A^c)$$

where the belief function quantifies the amount of information that surely supports A, and the plausibility function reflects the amount of information that potentially supports A. When focal sets are nested, a belief structure is equivalent to a possibility distribution, and the belief (plausibility) function is also a necessity (possibility) measure. In this model the mass $m(E)$ should be interpreted as the probability of only knowing that the unknown quantity lies in E.

A natural way of putting fuzzy sets and belief functions together is to assume that focal sets are fuzzy. Suppose there are p fuzzy focal sets denoted F_i. The set of fuzzy focal sets along with masses $m(F_i)$ can be viewed as a fuzzy random variable. The degrees of belief and plausibility of a fuzzy event A are defined as follows:

$$Pl_m(A) = \sum_{i=1}^{p} m(F_i) \int_0^1 \sup_{w \in F_i^\alpha} \mu_A(w)\ d\alpha \qquad (1)$$

$$Bel_m(A) = \sum_{i=1}^{p} m(F_i) \int_0^1 \inf_{w \in F_i^\alpha} \mu_A(w)\ d\alpha \qquad (2)$$

where F_i^α is the α-cut of the fuzzy focal element F_i. This is Yen's [15] definition. The reason for choosing this generalization rather than another one is that the part involving fuzzy sets F_i in Equations (1) and (2) comes down to compute the Choquet integral [19] of the (possibly fuzzy) event A with respect to the possibility and necessity measures induced by the distribution $\pi_i = \mu_{F_i}$, with μ_{F_i} the membership function of F_i. This means that linear operators are used in every part of the equation, which sounds more coherent than using a mix of linear operators and of maximum/minimum definition of possibility and necessity measures of fuzzy events. Note that Yen's work is not based on these two considerations, but rather on optimization criteria. Also note that Yen's approach is in concordance with Smets definition of a fuzzy event [20] (Equations (1) and (2) reduce to Smets definitions when the focal sets are crisp).

In the finite case, let $\{\alpha_1 = 1 > \cdots > \alpha_q \geq 0\}$ be the ordered collection of distinct values of the membership functions of focal sets F_i, $i = 1, \ldots, p$ (i.e., $\{\alpha_1 > \cdots > \alpha_q\} = \cup_{i=1,\ldots,p; x \in X} \mu_{F_i}(x)$). The degrees of belief and plausibility of a fuzzy event A become:

$$Pl_m(A) = \sum_{i=1}^{p} m(F_i) \sum_{\alpha_j} (\alpha_j - \alpha_{j-1}) \max_{w \in F_i^{\alpha_j}} \mu_A(w) \quad (3)$$

$$Bel_m(A) = \sum_{i=1}^{n} m(F_i) \sum_{\alpha_j} (\alpha_j - \alpha_{j-1}) \min_{w \in F_i^{\alpha_j}} \mu_A(w) \quad (4)$$

with μ_A the membership function of the fuzzy event A (in the case of classical events, this function takes only values in $\{0, 1\}$).

Remark that this generalization of belief structures to structures with fuzzy focal sets has another theoretical justification. In fact, it comes down to reducing a random fuzzy set to a regular random set where each cut $F_i^{\alpha_j}$ has mass $m(F_i)(\alpha_j - \alpha_{j-1})$ [21]. In the continuous case [22], it is equivalent to consider the convex combination of possibility and necessity measures (viewed as continuous consonant plausibility and belief functions) induced by π_i.

C. Problem statement and illustration

In this paper, we will consider a set of N sources, each of them providing a possibility distribution π_i as their evaluation of an unknown quantity $x \in X$. We will use maximal coherent subsets to summarize the information and will then work on the resulting structure. We note $[\![N]\!] = \{1, \ldots, N\}$ the set of natural numbers from 1 to N.

To illustrate our purpose, consider the following illustrative example : four sources (experts, computer code, sensor, ...) providing information in term of a best-estimate and a conservative interval, and the possibility distributions are supposed to have trapezoidal shapes. The information, represented in Figure 1, is summarized in Table I.

III. Existing fusion rules: a quick review

Fusion rules mainly follow three different kinds of behaviors [23]:

Conjunctive mode: comes down to retaining information common to all sources. A conjunctive fusion rule presupposes that all sources are reliable, which is often too

TABLE I
EXAMPLE INFORMATION FROM SOURCES

Source	Conservative interval	Best estimate
1	[1,5]	[2,4]
2	[1,13]	[3,6]
3	[3,11]	[7]
4	[5,13]	[10,12]

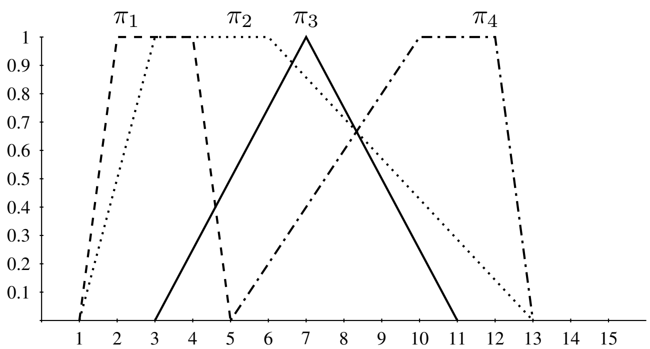

Fig. 1. Example distributions

optimistic. In case of conflicting information, such rules lead to poorly reliable results and cannot be applied if the conflict is total between some sources. In the context of possibility theory, the conjunction reads

$$\pi_\cap = \bigcap_{i=1,\dots,N} (\pi_i)$$

where $\cap$ is a t-norm operator (often the minimum or the product), which generalizes set-intersection.

Disjunctive mode: opposite to the conjunctive mode, it performs the union of all (fuzzy) sets that model the pieces of information provided by sources. It makes the pessimistic assumption that **at least** one source is reliable, without knowing which one. The pure disjunctive rule gives reliable results, but they are often too imprecise to be really useful. In possibility theory, it reads

$$\pi_\cup = \bigcup_{i=1,\dots,N} (\pi_i)$$

where $\cup$ is a t-conorm operator (often the maximum) generalizing set-union.

Trade-off mode: this kind of fusion rule lies between conjunctive and disjunctive mode, and is often used when sources are partially conflicting. Usually, it tries to maintain a good balance between reliability and informativeness. The resulting possibility distribution π^* of a trade-off rule is such that

$$\pi_{\min} < \pi^* < \pi_{\max}.$$

There are many possible trade-off rules, and we only recall here the most commonly used (see [24] for a review of existing trade-off rules in possibility theory, and [25] for compromise aggregation operators in general).

Weighted arithmetic mean: It is the most popular and commonly used trade-off combination. It reads

$$\pi_{WA} = \sum_{i=1}^{N} \lambda_i \pi_i$$

where λ_i can be considered as a measure of source i reliability. Weighted average can be interpreted as a statistical counting procedure, where a source i is considered as an independent sample of weight λ_i. Many other trade-off fusion rules are based on weighted average: Yager introduces the use of ordered weighted average (OWA) in [26] and proposes extensions in [27].

Adaptive rule: the aim of an adaptive rule is to progressively go from conjunctive to disjunctive behavior as conflict between sources increases. In case of total conflict (agreement) between sources, the conjunctive (disjunctive) mode is retrieved. The following adaptive rule, proposed by Dubois and Prade [28], is often used as a reference, even if partially ad hoc:

$$\pi_{AD}(x) = \max\left(\frac{\pi_{(c_*)}(x)}{h(c^*)}, \min(\pi_{(c^*)}(x), 1 - h(c^*))\right) \quad (5)$$

with

$$\begin{aligned} h(T) &= \sup_x \left(\min_{i \in T} \pi_i(x)\right) \\ c_* &= \sup_{T \subset [\![N]\!]} (|T|, h(T) = 1) \\ c^* &= \sup_{T \subset [\![N]\!]} (|T|, h(T) > 0) \\ h(c^*) &= \max(h(T), |T| = c^*) \\ \pi_{(k)}(x) &= \max_{|T|=k} \left(\min_{i \in T} \pi_i(x)\right) \end{aligned}$$

with $T \subseteq [\![N]\!]$ a subset of sources, and $|T|$ its cardinality. c_* is the greatest number of sources that completely agree together (the cores of distributions intersect), while c^* is the greatest number of sources that partially agree together (the supports of distributions intersect). Distribution $\pi_{(k)}$ is the disjunction of conjunctions of distributions stemming from subsets of k sources (it is equivalent to the t-norm $\min$ if $k = N$, and to the t-conorm $\max$ if $k = 1$). $h(T)$ can be interpreted as a measure of the agreement between the sources in subset T (it is the height of the conjunction between the distributions from sources in T). $h(c_*)$ is the maximal level of agreement between sources in subsets of size c_*. Equation (5) can thus be interpreted as a trade-off between an optimistic ($\pi_{(c^*)}$) and a pessimistic ($\pi_{(c_*)}$) distribution, with the last one discounted by a level $1-h(c^*)$.

In [29], alternatives to Equation (5) that consider the distance between possibility distributions are proposed. Their aim is to account for the metric structure of space X, and to ignore potential outliers. These alternatives mainly consist of reformulating $\pi_{(c_*)}$ into a distribution $\pi'_{(c_*)}$, its shape depending on a threshold distance d_0 and on the distance of a point from a consensus zone (e.g. the core of $\pi_{(c^*)}$).

A generalization of Equation (5) using the Hamacher t-norm family (instead of operators max and min) is proposed in [30]. In [31], another adaptive rule using reliability of sources is proposed.

In the sequel, the new proposal is also an adaptive rule, in the sense that it respectively reduces to a disjunction or a conjunction when sources respectively conflict or agree together. Nevertheless, an important difference with the schemes mentioned above is that instead of directly producing a final synthetic possibility distribution, we propose to build a fuzzy belief structure, more faithfully reflecting all the information delivered by the multiple sources. The resulting structure is theoretically meaningful but can be hard to handle in practice, and we will propose practical tools to exploit it in different ways (one of them being the construction of a final synthetic possibility distribution that can then be compared to the other proposals).

IV. A METHOD BASED ON MAXIMAL COHERENT SUBSETS (MCS)

When no information is available about the sources reliability, and when these sources are conflicting, a reasonable fusion method should take account of the information provided by all sources (i.e. without discarding any). At the same time, it should try to gain a maximum of informativeness. The notion of maximal coherent subsets (MCS) is a natural way to achieve these two goals. It consists of applying a conjunctive operator inside each non-conflicting subset of sources, and then to use a disjunctive operator between the partial results [7], [32]. With such a method, as much precision as possible is gained while not neglecting any source, an attractive feature in information fusion problem. We now explain in detail how this approach applies to possibility distributions on the real line.

A. Computing maximal coherent subsets of intervals

Assume the set $[\![N]\!]$ of sources supply N intervals $I_i = [a_i, b_i]$, $i = 1, \ldots, N$. Using the method of maximal coherent subsets on these intervals comes down to finding every maximal subset $K_j \subset [\![N]\!]$ of sources such that $\cap_{i \in K_j} I_i \neq \emptyset$ and then to performing the union of these partial results (i.e. $\cup_j \cap_{i \in K_j} I_i$). Algorithm 1, that finds maximal coherent subsets, was given by Dubois et al. in [33]. Contrary to what happens in logic (where the exhaustive search for maximal coherent subsets of formulas is of exponential complexity), once boundary values $\{a_i, b_i | i = 1, \ldots, N\}$ of all intervals have been sorted out, the Algorithm 1 is linear in the number of intervals, and thus computationally efficient.

The algorithm is based on increasingly sorting the interval end-points into a sequence $(c_i)_{i=1,\ldots,2N}$ that is scanned in this order. Each time (and only then) it meets an element c_i of type b, (i.e. the upper bound of an interval) followed by an element c_{i+1} of type a (i.e. the lower bound of another interval), a maximally coherent set of intervals is obtained. Figure 2 illustrates the situation for α-cuts of level 0.5 of our example. Using Algorithm 1, we find two maximal coherent

Algorithm 1: Maximal coherent subsets of intervals

Input: N intervals
Output: List of m maximal coherent subsets K_j

List = $\emptyset$;
j=1 ;
$K = \emptyset$;
Order in an increasing order $\{a_i | i = 1, \ldots, N\} \cup \{b_i | i = 1, \ldots, N\}$;
Rename them $\{c_i | i = 1, \ldots, 2N\}$ with $type(i) = a$ if $c_i = a_k$ and $type(i) = b$ if $c_i = b_k$;
for $i = 1, \ldots, 2N-1$ **do**
  if $type(i) = a$ **then**
    Add Source k to K s.t. $c_i = a_k$;
    if $type(i+1) = b$ **then**
      Add K to List ($K_j = K$) ;
      $j = j + 1$;
  else
    Remove Source k from K s.t. $c_i = b_k$;

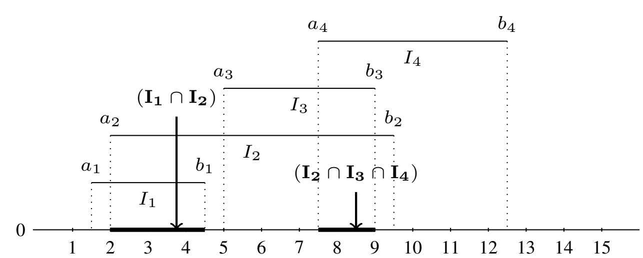

Fig. 2. Maximal coherent subsets on Intervals (0.5-cuts of example)

subsets : $K_1 = \{I_1, I_2\}$ and $K_2 = \{I_2, I_3, I_4\}$. After applying the maximal coherent subset method, the result is $(I_1 \cap I_2) \cup (I_2 \cap I_3 \cap I_4) = [2, 4.5] \cup [7.5, 9]$, as pictured in bold lines on the figure. They can be thought of as the most likely intervals where the unknown value may lie.

B. Building the fuzzy belief structure

Now the information provided by the N sources are supposed to be possibility distributions π_i formally equivalent to fuzzy intervals. At each level α, their α-cuts form a set of N intervals E_i^α. It is then possible to apply Algorithm 1 to them : Let K_j^α be the maximal subsets of intervals such that $\bigcap_{i \in K_j^\alpha} E_i^\alpha \neq \emptyset$. Define E^α as the union of the partial results associated to K_j^α as suggested in [24] :

$$E^\alpha = \bigcup_{j=1,\ldots,f(\alpha)} \bigcap_{i \in K_j^\alpha} E_i^\alpha \tag{6}$$

where $f(\alpha)$ is the number of subsets K_j^α of maximal consistent intervals at level α. In general, E^α is a union of disjoint intervals, and it does not hold that $E^\alpha \supset E^\beta \quad \forall \beta > \alpha$. So, the result is not a possibility distribution, since the sets E^α are not nested. In practice, for a finite collection of fuzzy intervals, there will be a finite set of $p + 1$ values

$0 = \beta_1 \leq \ldots \leq \beta_p \leq \beta_{p+1} = 1$ such that the sets E^α will be nested for $\alpha \in (\beta_k, \beta_{k+1}]$, $k = 1, \ldots, p$. Algorithm 2 offers a simple method to compute threshold values β_k. It simply computes the height of $\min(\pi_i, \pi_j)$ for every pair of possibility distributions π_i, π_j. Clearly, such a value is the threshold above which π_i and π_j do not belong to the same coherent subset anymore.

Algorithm 2: Values β_k of fuzzy belief structure

Input: N possibility distributions π_i
Output: List of values β_k

List = $\emptyset$;
i=1 ;
for $k = 1, \ldots, n$ **do**
  for $l = k+1, \ldots, n$ **do**
    $\beta_i = \max(\min(\pi_k, \pi_l))$;
    i=i+1 ;
  Add β_i to List ;
Order List by increasing order ;

If we apply the MCS method in (6) for all $\alpha \in (\beta_k, \beta_{k+1}]$, we can build a non-normalized fuzzy set F_k with membership range $(\beta_k, \beta_{k+1}]$ (since sets E^α are nested in that range). We can then normalize it (so as to expand the range to $[0, 1]$) by changing $\mu_{F_k}(x)$ into

$$\frac{\max(\mu_{F_k}(x) - \beta_k, 0)}{\beta_{k+1} - \beta_k}$$

while assigning weight $m_k = \beta_{k+1} - \beta_k$ to this fuzzy focal set. By abuse of notation, we still denote F_k these normalized fuzzy focal sets in the sequel. Overall, we built a fuzzy belief structure $(\mathcal{F}, m)$ with weights m_k bearing on normal focal sets F_k. The weight m_k can be interpreted as the confidence given to adopting F_k as the information provided by all the sources. Figure 3 illustrate the result on the example. The 0.5-cut, is exactly the result of Figure 2. The result of Equation (6) for each level $\alpha \in (0, 1]$ is in bold. The obtained fuzzy belief structure is thus a meaningful "fuzzification" of the MCS method used on classical intervals. Remark that if all sources agree at least on one common value, the result is a single fuzzy focal set equivalent to $\pi(x) = \min_{i=1,\ldots,N} \pi_i(x)$ (usual conjunction). On the contrary, if every pair of sources is in a situation of total conflict (i.e. $\sup_{x \in X} \min(\pi_i, \pi_j) = 0 \quad \forall i \neq j$), then the result is a unique fuzzy focal set $\pi(x) = \max_{i=1,\ldots,N} \pi_i(x)$ (usual disjunction). Thus, as mentioned before, the maximal coherent subset method has the behavior of an adaptive rule.

Belief and plausibility measures can be derived for events or fuzzy events from Equations (1)-(2) viewing the fuzzy random set as a convex combination of standard continuous consonant belief structures associated to the fuzzy focal sets. For crisp events A, these equations come down to

$$Pl_m(A) = \sum_{i=1}^{p} m(F_i) \sup_{x \in A} \pi_i(x); \tag{7}$$

$$Bel_m(A) = \sum_{i=1}^{p} m(F_i) \inf_{x \notin A} 1 - \pi_i(x). \tag{8}$$

The results of the MCS method can also be encoded in the form of a continuous belief structure [22] defined by the Lebesgue measure on the unit interval ($\alpha \in [0, 1]$) together with the mapping $\alpha \to E^\alpha$. The associated basic belief density will be denoted $m^c(E^\alpha) = 1 \ \forall \alpha \in [0, 1]$. One can then work on this continuous structure instead of working on $(\mathcal{F}, m)$. The corresponding plausibility and belief measures are then defined as

$$Pl_c(A) = \int_0^1 \sup_{w \in E^\alpha} \mu_A(w) \, d\alpha \tag{9}$$

$$Bel_c(A) = \int_0^1 \inf_{w \in E^\alpha} \mu_A(w) \, d\alpha \tag{10}$$

where μ_A is the membership function of the (fuzzy) event A. It can be proved that the two belief structures are equivalent. First consider a random fuzzy interval of the real line $\{(F_i, m_i), i = 1, \ldots, p\}$. Define $\alpha_{i+1} = \sum_{j=1}^{i} m_j, \forall i = 1, p$ with $\alpha_1 = 0$. The corresponding continuous belief function is defined by the Lebesgue measure on the unit interval together with the mapping

$$\alpha \to E^\alpha = (F_i)_{\phi_i(\alpha)}, \forall \alpha \in (\alpha_i, \alpha_{i+1}],$$

where $\phi(\alpha) = \frac{\alpha - \alpha_{i+1}}{m_i}$ maps $(\alpha_i, \alpha_{i+1}]$ to $(0, 1]$. Then we can prove the following:

Theorem 1: $\forall A \subseteq X, Pl_m(A) (= \sum_1^p m_i \Pi_i(A)) = Pl_c(A)$

Proof: Denote by $\mathbf{1}_A$ the function with value 1 except if $A = \emptyset$ where its value is 0. Let $\beta_i = \phi_i(\alpha)$ and notice that $d\alpha = m_i d\beta_i$. Then

$$Pl_c(A) = \int_0^1 \mathbf{1}_{A \cap E^\alpha} d\alpha = \sum_1^p \int_{\alpha_i}^{\alpha_{i+1}} \mathbf{1}_{A \cap (F_i)_{\phi_i(\alpha)}} d\alpha$$

$$= \sum_1^p \int_0^1 \mathbf{1}_{A \cap (F_i)_{\beta_i}} m_i d\beta_i = \sum_1^p m_i \Pi_i(A).$$

■

This formal result shows that the fuzzy belief structure $(\mathcal{F}, m)$ resulting from the fusion process can be reduced to an equivalent convex combination of possibility measures or to a continuous random set, and that the three structures have the same informative content. These structures are theoretically attractive and well represent the information delivered by the sources. Nevertheless, such structures would be difficult to interpret and have little utility for the analyst, due to their complexity. We thus see the structure described in this section and pictured in Figure 3 as a theoretically sound model summarizing the (potentially conflicting) information provided by the sources, from which can then be extracted, by proper tools, useful and interpretable information. Such tools are proposed in the next sections.

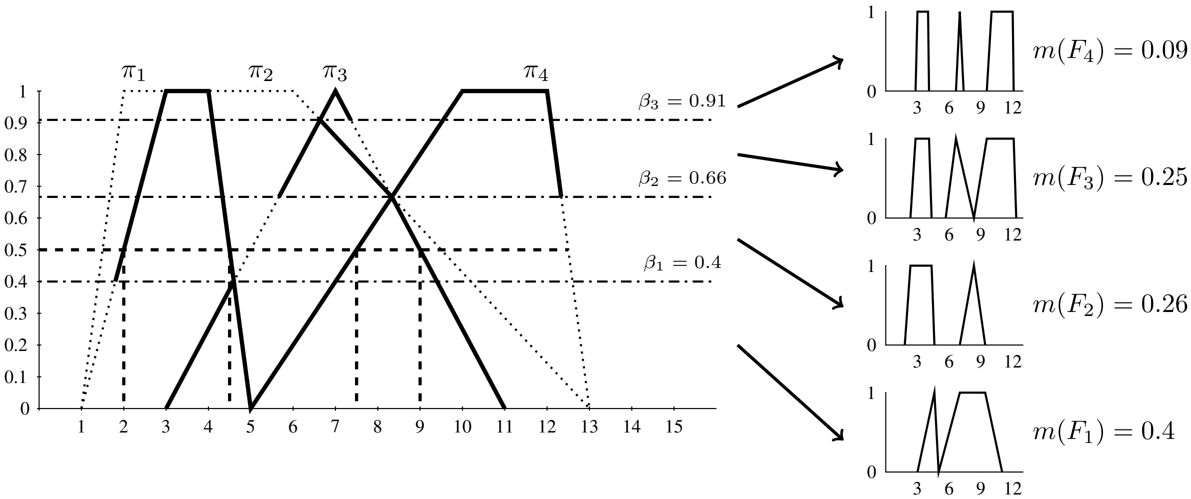

Fig. 3. Result of maximal coherent subset method on example (—) and 0.5-cut (- - -)

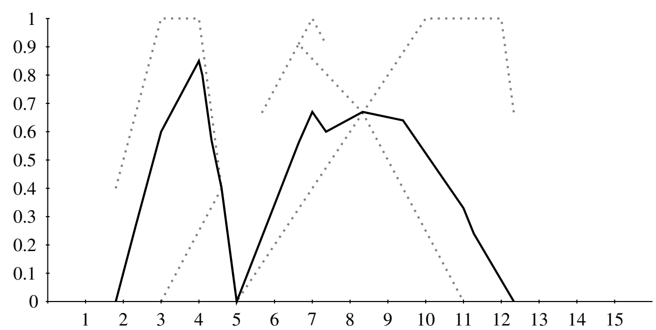

Fig. 4. Contour function π_c (—) with fuzzy focal sets (- - -)

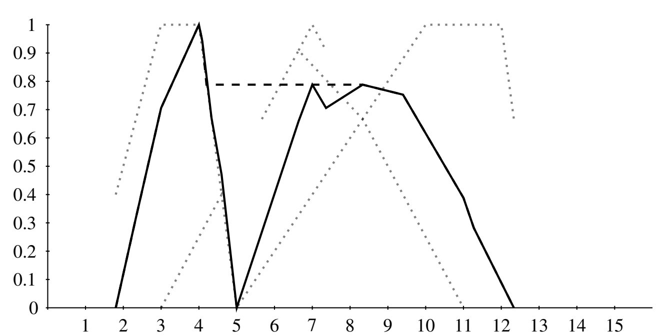

Fig. 5. Contour function, normalized (—) and convexified (- - -), with fuzzy focal sets (gray lines)

C. Building a final possibility distribution

As we said, it is hard to directly use the fuzzy belief structure representation in practical problems (such as uncertainty propagation through a mathematical model). In this case, a method that derives a unique possibility distribution from a fuzzy belief structure $(\mathcal{F}, m)$ is needed.

A natural candidate is to build the contour function of the obtained continuous belief structure:

$$\forall x \in X, \quad \pi_c(x) = Pl(x) = \sum_1^p m_i \pi_i(x), \tag{11}$$

i.e. boils down to computing the weighted arithmetic mean of the membership functions of (normalized) fuzzy focal sets F_i, the weight of F_i being equal to m_i. One can then normalize the resulting distribution π_c[1] and/or take its convex hull if needed.

Figure 4 shows the contour function π_c and Figure 5 shows the same function, once normalized and convexified, together with the fuzzy focal elements in the background. No assumptions are made on source reliabilities or the metric structure of the space. We can see on these figures that the contour function π_c and its (convexified) normalization can directly be used by an analyst, and provide a summary of the fuzzy belief structure $(\mathcal{F}, m)$.

On our example, the final result is a bimodal distribution, with one mode centered around value 8 and the other with a value of 4, this last value being the most plausible. This is so because these areas are the only ones supported by three sources whose information are highly (even if not perfectly) coherent. We can expect that the true value lies in one of these two areas, but it is hard to tell which one. Indeed, in this case, one should either take the normalized convex hull of π_c as the final representation of the parameter X, or find out the reason for the conflict (if feasible).

[1]by computing $\pi'_c(x) = \pi_c(x)/h(\pi_c)$ where $h(\pi_c)$ is the height of π_c

V. Properties of the MCS method

This section studies some properties of the MCS fusion rule in the light of requirements proposed by Oussalah [34]. Similar properties were studied by Walley [8] in the more general setting of imprecise probability. We use the same terminology as in [34] (we put between parentheses the name used in [8] for the same property when possible and relevant). For simplicity of notation (and to make comparison with other fusion rules easier), we will refer in the property definitions to the original distributions π_i and their relation with the resulting distribution π_c given by Equation (11), but we could have equally referred to the continuous belief structure m^c or the random fuzzy set $(\mathcal{F}, m)$, except for properties 10 and 12 which concern π_c only. In the sequel, φ denotes a general aggregation operator.

Prop. 1) Associativity (Aggregation of aggregates [8]): φ is associative if $\varphi(a, \varphi(b, c)) = \varphi(\varphi(a, b), c)$. The MCS method is not associative in general, and neither is its level-wise application to possibility distributions. Associativity is not verified in general by trade-off rules, and by our method in particular. It is also difficult to preserve under sophisticated conflict management, that require all sources to be considered at once. This property is quite useful for local or step-by-step computations, and usually allows to increase computational efficiency, but requiring it limits the potential fusion rules that can be used. Since the methodology proposed here do not demand high computational effort, we do not regard associativity as essential.

Prop. 2) Commutativity (Symmetry [8]): φ is commutative if $\varphi(a, b) = \varphi(b, a)$. Equation (6) does not depend on a particular order of the distributions π_i, thus the MCS method is commutative. Commutativity is necessary when sources cannot be ordered in a sensible way (that is the case here).

Prop. 3) Idempotence: φ is idempotent if $\varphi(a, a) = a$. After Equation (6), if the N sources supply the same fuzzy interval, we retrieve it using the MCS method, which is thus idempotent. When aggregating possibility distributions, idempotence can be seen as a cautious assumption in case of possible source dependencies. In particular there is no reinforcement effect when several sources supply the same information. If independence between sources must be acknowledged, one may combine the possibility distributions π_i, viewed as consonant belief structures, using Dempster rule of combination. It comes down to intersecting the cuts $E_i^{\alpha_i}$ for distinct values of α_i, combining the local mass functions multiplicatively. As this may result in conflict, one can apply the MCS method to such n-tuples of cuts, instead of doing it using a the same threshold α for all sources. In the case of two sources, note that it yields focal sets of the form $E_1^\alpha \cap E_1^\beta$ if not empty and $E_1^\alpha \cup E_1^\beta$ otherwise. This rule was already proposed by Dubois and Prade in 1988 [23]. Note that such a combination would have an exponential complexity, even on the real line.

Prop. 4) Weak zero preservation (Unanimity [8]): π_c satisfies weak zero preservation if $\pi_c \subseteq \max_{i=1,\dots,N}(\pi_i)$. This property states that if an element is considered as impossible by all the sources, then it is also impossible for the fusion result. This property corresponds to the informal requirement made in [31] that the support of the resulting distribution should be included in the union of the support of the source distributions. This property is verified by all adaptive rules (since they are equal to the disjunction only in case of pair-wise total conflict between all sources), and thus by the MCS method. Note that this property is called *strong* zero preservation in [34], but we choose to call it weak, since it puts less constraints on the result than its (here) strong counterpart. That a fusion rule should satisfy this property seems sensible, since not satisfying it would mean that information no sources provided at first can be present in the fusion rule result.

Prop. 5) Strong zero preservation (Conjunction [8]): π_c satisfies weak zero preservation if $\pi_c \subseteq \min_{i=1,\dots,N}(\pi_i)$. This property is verified when an element is considered as impossible if it is considered as impossible by at least one of the sources. This property is not generally verified by the SMC method. However, this property makes sense only if sources agree together, thus we do not regard it meaningful for an adaptive rule. In fact, requiring this property comes down to enforce a conjunctive behavior for the rule.

Prop. 6) Weak maximal plausibility (Indeterminacy [8]): π_c satisfies this property if $\pi_c \supseteq \min_{i=1,\dots,N}(\pi_i)$. A fusion rule verifies weak maximal plausibility if an element considered as possible by all sources is also considered possible by the fusion result. We can check that MCS method verifies this property (by an argument similar to the one used for weak zero preservation). Similarly to weak zero preservation, it seems sensible to require this property from a fusion rule, since satisfying it means that we want to be at least coherent with the information on which all sources agree.

Prop. 7) Strong maximal plausibility (Total reconciliation [8]): π_c satisfies this property if $\pi_c \supseteq \max_{i=1,\dots,N}(\pi_i)$. Strong interpretation of maximal plausibility is satisfied when an element is considered as possible in the fusion result if it is considered as possible by at least one of the sources. Although this ensures that every source will fully agree with the fusion result, this property leads most of the time to results that are too imprecise to be useful, since it enforces a disjunctive behavior.

Prop. 8) Information relevance (Reconciliation and strong reconciliation [8]): This property is informally stated in [29] as the requirement that all distributions π_i should be taken into account (unless explicitly stated otherwise by additional assumptions). Similar properties are more formally stated in [8], where they are called reconciliation and strong reconciliation. Let I be any maximal consistent subset of sources s.t. $\min_{i\in I}(\pi_i) \neq \emptyset$, then properties of reconciliation and strong reconciliation are respectively satisfied (in our context) if $\pi_c \cap \pi_i \neq \emptyset$ $i = 1, \dots, n$ and if $\pi_c \cap \left(\bigcap_{i\in I} \pi_i\right) \neq \emptyset$ for any MCS I. By its definition, the MCS method naturally satisfies (strong) reconciliation property. These properties are clearly desirable if we have no reason to discard some sources.

Prop. 9) Insensitivity to complete and relative ignorance [8]: Satisfying insensitivity to complete ignorance means that a source $N+1$ that provides no information at all should not influence the fusion result (i.e. $\pi_{N+1}(x) = 1$ if $x \in [l, u]$, 0 otherwise where $[l, u]$ is the whole domain). Insensitivity to relative ignorance is a stronger version in which a source $N+1$ that provide information implied by all the other sources taken together (i.e. $\pi_{N+1} \supset \max_{i=,\dots,N+1} \pi_i$) should not influence the fusion result. Again, MCS method naturally verifies these two properties (since a source $N+1$ as described above would be in every MCS). We regard them as desirable, since not satisfying them means that the result can get very imprecise from the information of just one source.

Prop. 10) Convexity: This property is satisfied if the fusion result is (fuzzy) convex (provided initial distributions are). This property is not generally satisfied by the MCS method, but it is always possible to take the convex hull of the result (which implies losing some information). As for associativity, convexity often simplifies mathematical and computational treatments, but there is no obvious theoretical reasons to require it.

Prop. 11) Robustness to small changes: This property means that small changes made to the original distributions (e.g. horizontally shifting a distribution $\pi(x)$ to $\pi'(x) = \pi(x + \epsilon)$, coarsening or reducing the support or the core of a distribution by a small value $\epsilon, \dots$) should only cause a small change on the final result. Since information is often approximately modeled, this property is often considered as desirable [29] (viz. the lack of robustness for the rule given by Equation (5), studied in [29]). Concerning the method proposed here, although small changes can have an important impact for a particular E^α by making a coherent maximal subset no longer coherent, small changes will have small impact on the overall structure $(\mathcal{F}, m)$ and on the distribution π_c (most of the time, small changes will only cause small shifts in values of β_k). Thus the MCS method is robust to small changes in the shape of the distributions π_i. Although robustness is important in automatic fusion procedures (where the result should be ensured to be non-empty), it is not essential in other cases.

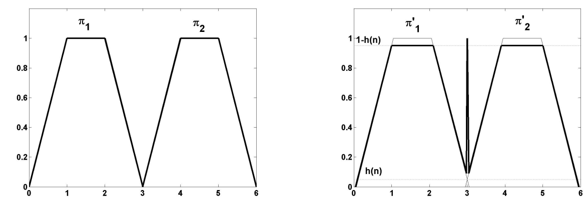

Fig. 6. Left : Equation (5) applied to two totally conflicting sources π_1, π_2
Right : Equation (5) applied to $\pi'_1(x) = \pi_1(x - 0.05)$ and $\pi'_2(x) = \pi_2(x + 0.05)$

Prop. 12) Core insensitivity under high conflict: The fact that, for some fusion rules, the core of the resulting distribution can be sensitive to small changes when data are highly conflicting has been emphasized in [31]. As an example, Figure 6 illustrates the sensitivity of the resulting core for Equation (5): When the two distributions are conflicting, then the core of the resulting distribution $c(\pi_{AD}) = c(\pi_1 \cup \pi_2)$, but as soon as $\min(\pi_1, \pi_2) \neq \emptyset$, $c(\pi_{AD}) = c(\min(\pi_1, \pi_2))$. With this kind of behavior, a value that both sources judge very unlikely can suddenly become the (only) most plausible value. This is indeed quite adventurous, and means that the core of the resulting distribution is not a continuous function of the conflict level. Nevertheless, even if one value can suddenly shift from impossible to the most plausible, the other values can remain highly plausible, and similarly to robustness, we do not regard this property as forcefully necessary. In comparison, the core of π_c resulting from the MCS method does not exhibit such a discontinuous behavior. Indeed, the core changes from $c(\pi_1 \cup \pi_2)$ to $c(\min(\pi_1, \pi_2))$ as $h(\pi_1, \pi_2)$[2] come close to 1 (complete agreement). Figure 7 illustrates the behavior of the MCS rule as the agreement level h between distributions π_1, π_2 of Figure 6 increase (i.e. as π_1 and π_2 are respectively shifted to the right and left). The figure shows that the disjunctive part of the MCS is dominant in the result until $h = 0.5$, after which the conjunctive part becomes dominant in the resulting distribution. For $h = 0.5$, disjunctive and conjunctive parts balance each other.

To summarize, when information is (partially) conflicting and when no specific assumptions can be made about the

[2]Note that $h(T) = \sup_x \left(\min_{i\in T} \pi_i(x) \right)$ is a measure of concordance inside subset T

	Properties											
	1	2	3	4	5	6	7	8	9	10	11	12
Fusion Rule												
min	√	√	√	√	√	√			√	√		
max	√	√	√	√		√	√	√			√	√
AD		√	√	√		√			√			
MCS		√	√	√		√		√	√		√	√

TABLE II
COMPARISON OF THE PROPERTIES OF THE SMC METHOD WITH OTHER RULES (AD : ADAPTIVE RULE, min: MINIMUM T-NORM, max: MAXIMUM T-CONORM).

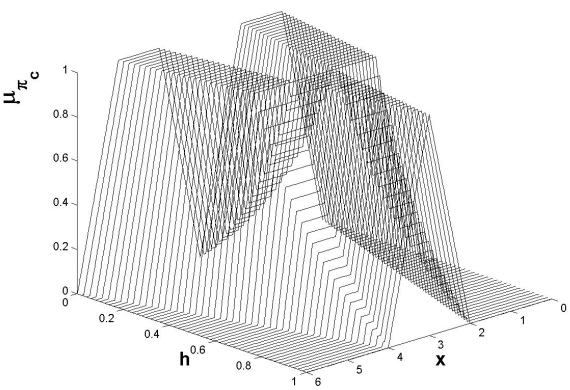

Fig. 7. Normalized final distribution π_c for two sources, in function of agreement level h

sources, we regard Properties 2, 3, 8, 9, 6, 4 as strongly desirable. Properties 1 and 10 can be required if computational efficiency is an issue, but we do not regard them as necessary, since they can greatly limit the scope of rules that can be used. Properties 5 and 7 are very strong (since requiring them enforces the rule to follow specific behaviors) and incompatible, thus we do not regard them as desirable. Properties 11 and 12 are important in automatic fusion procedures, but their necessity in other situations is arguable. Outside automatic fusion procedures, we do not regard them as forcefully desirable.

Table II summarizes the properties satisfied by the MCS method in contrast with some other known fusion rules. It satisfies all properties, except associativity, convexity, strong versions of zero preservation and maximal plausibility, the two latter being only satisfied in specific cases (respectively when all sources totally agree or are totally conflicting, that is when MCS method reduces to classical disjunction or conjunction). The MCS method satisfies all the properties of fusion rules that we regard as desirable, as well as some others. Associativity is incompatible with adaptiveness and convexity can hide the presence of conflict between sources. Overall, the MCS method meets all requirements advocated in [29], [31]. This fact indicates that the method is likely to be useful in practical applications, and compete with other rules.

VI. EXTRACTING USEFUL INFORMATION

The fuzzy belief structure $(\mathcal{F}, m)$ resulting from the MCS method is a good representation of the information provided by the overall group of sources. But it can be hard to draw conclusions or useful information directly from it (see Figure 3 to be convinced) if not simplified using, for instance the contour possibility distribution. However it has rich content. In this section, we present various evaluations that provide additional insights into the resulting information, and can be of practical usefulness.

A. Finding groups of coherent sources

For each threshold in $(\beta_k, \beta_{k+1}]$, merging the cuts applying Algorithm 1 exploits the same maximal coherent subsets $K_j^{(\beta_k,\beta_{k+1}]}$ of sources. Changing the value of this threshold yields a finite collection of coverings of the set of sources. Increasing the threshold from 0 to 1, we go from the largest sets of agreeing sources (i.e. those for which the supports of distributions π_i intersect), to the smallest sets of agreeing sources (i.e. those for which cores intersect). Subsets $K_j^{(\beta_k,\beta_{k+1}]}$ can be interpreted as clusters of sources that agree up to a confidence level β_{k+1}.

Analyzing these clusters can give some information as to which groups of sources are consistent, i.e. agree together with a high confidence level (possibly using some common evidence to supply information) and which ones are strongly conflicting with each other (and which items of information are plausibly based on different pieces of evidence). The groups in our example are summarized in the following table

Subsets	Clusters	Max. Conf. level
$K^{(0,0.4]}$	$[1,2,3][2,3,4]$	0.4
$K^{(0.4,0.66]}$	$[1,2][2,3,4]$	0.66
$K^{(0.66,0.91]}$	$[1,2][2,3][4]$	0.91
$K^{(0.91,1]}$	$[1,2][3][4]$	1.0

In our example, only few conclusions can be drawn from the clusters, showing that, if this kind of summary can be useful, it is not sufficient. Results show that some sources are totally conflicting (since there is more than one subset in $K^{(0,\beta_1]}$), and that source 4 looks more isolated than the three others (at a confidence level higher than 0.66, it is strongly conflicting with all other sources). This type of analysis can trigger further investigations on reasons for conflict.

B. Measuring the gain in precision

It is interesting to measure how much precision is gained by applying the MCS method to a set of N possibility distributions. Let $\pi_\cup$ be the disjunction such that $\pi_\cup = \max_{i=1,\dots,N} \pi_i$. We consider that the overall imprecision of the information provided by all the sources is equal to

$$IP = |\pi_\cup| = \int_X \pi_\cup(x)dx$$

where $|\pi_\cup|$ is the fuzzy cardinality of $\pi_\cup$, an extension of usual interval cardinality (the cardinality being a natural candidate to measure imprecision). After fusion by the MCS method, the imprecision of the resulting fuzzy belief function can be measured as

$$IP' = \sum m_k |F_k|$$

The difference $GP = IP - IP'$ quantifies the precision gained due to the fusion process. This index is 0 in case of total conflict and when the sources supply the same possibility distribution. Indeed, the MCS method increases the precision when sources are consistent with one another but supply distinct pieces of information.

In our example, we have $IP = 11.195$, $IP' = 5.412$ and the normalized index is 0.52, which indicates a reasonable gain of precision after fusion.

Since fusion rule presented here is based on a level-wise application of the maximal coherent subset methodology, it is natural to investigate the behaviour of the level-wise gain of precision. That is, we can compute, for each threshold α

$$IP(\alpha) = |\pi_\cup^\alpha| \qquad IP'(\alpha) = |E^\alpha|$$

where $|\pi_\cup^\alpha|$ is the cardinality of the α-cut of $\pi_\cup$. Since these evaluations depend on α, they can be viewed as gradual numbers [35], [36]. A gradual number is formally a mapping from $(0, 1]$ to the real line $\mathbb{R}$, such as $IP(\alpha)$ and $IP'(\alpha)$. Clearly, $IP(\alpha)$ measures the imprecision of the continuous belief structure $m^{E_\cup}$ which assigns to each $\alpha \in [0, 1]$ the set $E_\cup^\alpha = \bigcup_{i=1,\dots,N} E_i^\alpha$ (E_i^α is the α-cut of π_i). $IP(\alpha)$ is a gradual evaluation the imprecision of the belief structure resulting from the level-wise disjunction of α-cuts. It is a monotonic gradual number. The gradual number $IP'(\alpha)$ measures the imprecision of m^c likewise. However it is generally neither continuous nor monotone. The gradual number $GP(\alpha) = IP(\alpha) - IP'(\alpha)$ is thus a level-wise measure of the precision gained by applying the maximal coherent subset method. The following equality formalizes the link between these gradual numbers and their scalar counterparts IP, IP' and GP:

$$IP = \int_0^1 IP(\alpha)d\alpha,$$

and likewise for IP' and GP. Since $m_k|F_k| = \int_{\beta_{k-1}}^{\beta_k} |E^\alpha|d\alpha$, we effectively have $IP' = \int_0^1 IP'(\alpha)d\alpha$. The validity of the other equality $IP = \int_0^1 IP(\alpha)d\alpha$ follows from the definition of fuzzy cardinality.

C. Group confidence in an event, in a source

Since we consider the fuzzy belief structure $(\mathcal{F}, m)$ resulting from the MCS method as a good representative of the group of sources, plausibility and belief functions of an event A can be interpreted respectively as an upper and a lower confidence level given to A by the sources. In particular, if $A = \pi_i$, plausibility and belief can be used to evaluate the resulting upper and lower "trust" in the information given by source i in view of all the sources.

In our example, values $[Bel_m(\pi_i), Pl_m(\pi_i)]$ for sources 2 and 4 are, respectively, $[0.38, 1]$ and $[0, 0.93]$ (using Equations (2)-(1) or (10)-(9)). We see that information provided by source 2 is judged totally plausible by the group, and also strongly supported (source 2 is undoubtedly the less conflicting of the four). Because one source completely disagrees with source 4, its belief value drops to zero, but the information delivered by it is nevertheless judged fairly plausible (since source 4 is not very conflicting with sources 2 and 3).

Belief and plausibility functions are natural candidates to measure the overall confidence in a source, but their informativeness can sometimes be judged too poor. Indeed, if a distribution π_i given by a source i is in total conflict with the others, the resulting fuzzy belief structure $(\mathcal{F}, m)$ will give the following measures for $\pi_i : [Bel_m(\pi_i), Pl_m(\pi_i)] = [0, 1]$ (total ignorance). It means that in the presence of strong conflict, the MCS method grants no confidence in individual sources, even though no source can be individually discarded. On the contrary, if the pieces of information are fully consistent, $Bel_m(\pi_i) \geq 0.5$ and $Pl_m(\pi_i) = 1$. Note that it suffices that one source contradicts other globally consistent sources for $Bel_m(\pi_i)$ to vanish because the MCS method deteriorates precision (even if to a limited extent) in the case of inconsistency.

An alternative to reduce this potential imprecision is to take a fuzzy equivalent of the so-called pignistic probability, namely

$$BetP(A) = \sum_{k=1}^{p} m(F_k) \frac{|F_k \cap A|}{|F_k|} \tag{12}$$

where $|F_k \cap A|/|F_k|$ is taken as the degree of subsethood , also called relative cardinality, of the fuzzy set F_k in A, with A a (fuzzy) event. This pignistic probability is zero if A is strongly conflicting with every focal set F_k and one if every F_k is included in A (here, F_k is included in A iff $\mu_{F_k}(x) < \mu_A(x) \forall x$). In the example, Equation (12) applied to sources 2 and 4 ($A = \pi_2$ and $A = \pi_4$) respectively gives confidence 0.80 and 0.49, confirming that source 2 is more trusted by the group than source 4.

Let us note that other formulas instead of $|F_k \cap A|/|F_k|$ could have been chosen to measure the subsethood of F_k in A. Such other measures are considered in [37], [38]. One could also choose to consider the continuous random set E_α and to use the continuous extension of the pignistic probability proposed in [22], which would give yet another result. Further research are needed to know the properties of each of these measures and the relations existing between them, and it is presently not clear when to choose one measure rather than the others. From our standpoint, the important criteria satisfied by these measures is that they are consistent ways to measure the coherence of A with respect to the fuzzy belief structure coming out from the MCS method (e.g., in our example, source 2 would be judged more reliable than source 4, irrespectively of the chosen formula for the pignistic probability, and only the scalar evaluations would change).

VII. Taking additional information into account

The fact that one needs no further information than the distributions π_i to apply the MCS method described above is an advantage: this means that the method is applicable to any situation where information is modeled by possibility distributions and to any space X (not only the real line)[3].

However, it is often desirable to account for reasonable assumptions or some additional information (either about the sources or the particularities of the underlying space) when using the fusion rule. It can be assumptions related to the credible number of reliable sources, the existence of a metric on the space X, information about individual source reliabilities,...

We thus propose such extensions of the MCS method, that accommodate such assumptions or information in simple ways.

A. Number of reliable sources

Suppose there is information on the number r of sources that can be expected to be reliable, or at least that some assumptions can be made about this number. Given the lower bound r on the number of reliable sources, we propose to adapt Equation (6) as follows

$$E_r^\alpha = \bigcup_{\substack{j=1,\dots,f(\alpha),\\ |K_j^\alpha|\geq r}} \bigcap_{i\in K_j^\alpha} E_i^\alpha \tag{13}$$

where $|K_j^\alpha|$ is the number of sources in the maximal coherent subset K_j^α. Namely, for each level α, only coherent subsets which contain at least r sources are taken into account.

Using this threshold r, the contribution of isolated or small groups of consistent sources is lessened. The proposed values of r can of course be decreased or increased according to the situation and the available information. Figure 8 illustrates the fuzzy belief structure resulting from our example when $r = 2$. This choice leads to discard all the information given by source 4 after $\alpha = 0.66$, as well as a small part of source 3 information (modifying only focal sets F_3 and F_4 when compared to Figure 3). Our final structure is thus more informative, as the (assumed) poorly reliable information supporting values above 11 has been discarded.

[3]Note that Algorithm 1 is only applicable to completely (pre-)ordered spaces X, and if X is a finite space, the continuous belief structure m^c just become a usual discrete belief function

B. Accounting for the reliability of sources

Suppose that some numerical evaluation of the reliability of each source is available. Denote λ_i the reliability of source i, and suppose, without loss of generality, that $\lambda_i \in [0, 1]$, value 1 meaning that the source is fully reliable, 0 representing a useless source. There are at least two ways of taking this reliability indices into account, the first one increasing the result imprecision by modifying (i.e. discounting) the possibility distributions, the second one decreasing the imprecision by discarding poorly reliable subgroups of sources:

- *Discounting*: discounting consists of transforming π_i into a distribution π'_i whose imprecision increases all the more as λ_i is low. In other words, the lower λ_i is, the more irrelevant π_i becomes. A common discounting operation is:

  $$\pi'_i(x) = \max(1 - \lambda_i, \pi_i(x))$$

  Once discounted, sources are assumed to be reliable. The effect of the discounting operation on MCS method possesses a nice interpretation. Indeed, applying the MCS method to discounted sources means that the information modeled by π_i will only be considered for levels higher than $1-\lambda_i$, since below that level, source i is present in every subsets K_j, as no information coming from it will be considered. A draw-back of this method is that if values λ_i are too low, the result will be highly imprecise.
- *Discarding unreliable sources*: assume the overall reliability of a subgroup K is of the form

  $$\lambda_K = \perp_{i\in K}(\lambda_i)$$

  where $\perp$ is a t-conorm. Choosing a particular t-conorm to aggregate reliability scores then depends on the dependence between the sources. For example, the t-conorm $\perp(x, y) = \max(x, y)$ correspond to the cautious assumption that agreeing sources are dependant (i.e. use the same information), thus the highest reliability score is not reinforced by the fact that other sources agree. On the contrary, the t-conorm $\perp(x, y) = x + y - xy$ (the dual of the product t-norm) can be associated to the hypothesis that sources are independent (reliability score increases as more sources agree together). A limit value $\underline{\lambda}$ can then be fixed, such that only subsets of sources having a reliability score over this limit are kept. Equation (6) then becomes

  $$E_{\underline{\lambda}}^\alpha = \bigcup_{\substack{j=1,\dots,f(\alpha),\\ \lambda_{K_j^\alpha}\geq\underline{\lambda}}} \bigcap_{i\in K_j^\alpha} E_i^\alpha \tag{14}$$

  Remark that this method does not modify the pieces of information π_i.

We now consider our example with $\lambda_1 = 0.2$, $\lambda_2 = 0.6$, $\lambda_3 = 0.8$,$\lambda_4 = 0.2$. Figure 9 shows the result of the MCS method after discounting (the bounds of the variation domain are assumed to be $[1, 14]$). The result using discounting is very different from the result obtained with the original

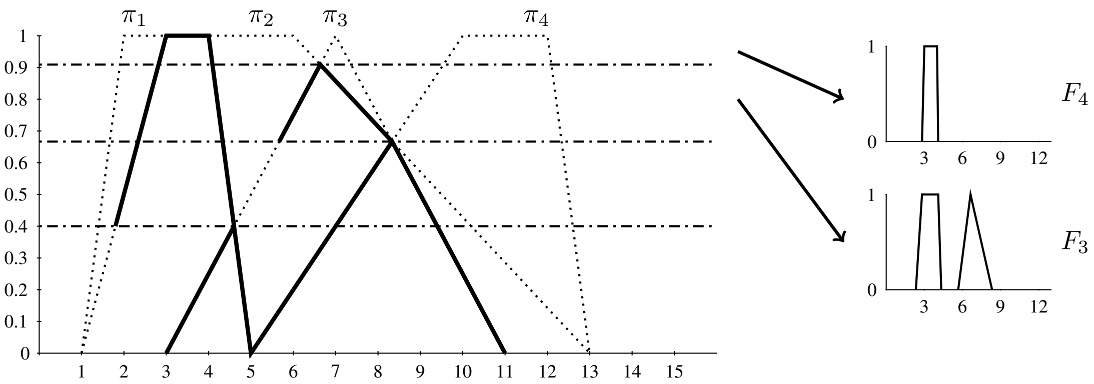

Fig. 8. Result of MCS method with number of reliable sources $r = 2$

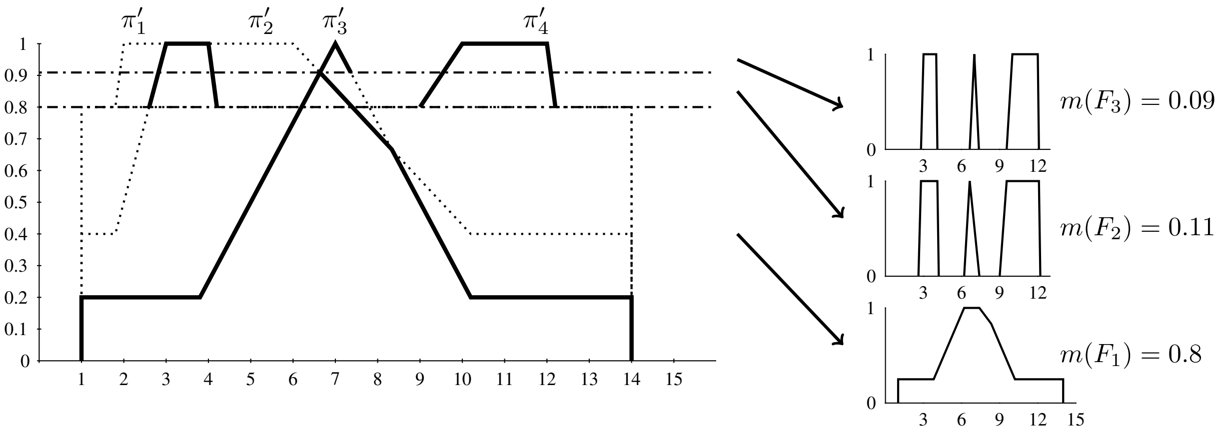

Fig. 9. Result of MCS method with reliability scores $\lambda = (0.2, 0.6, 0.8, 0.2)$ and discounting

method, and it is clear that distribution π_3' (i.e. the most reliable source) dominate the others. Figure 10 shows the result of discarding poorly reliable sources, where independence is assumed and $\underline{\lambda} = 0.5$. As we can see, the result is this time very close to the result of Figure 8, except that now all the information delivered by source 3 is taken into account, due to its high reliability. Comparing Figure 8 to Figure 3, we still have that only fuzzy focal sets F_3 and F_4 are modified. From Figure 10, we can see that the fact of discounting sources can have a significant impact on the result.

C. Accounting for the metric

In the original MCS method, if an isolated source is totally conflicting with the others, then it will constitute a maximal coherent subset of its own. If the notion of distance makes some sense in X (X is a metric space), this will be true whatever the distance of the isolated source distribution from the others is. As stressed in [29], it is sometimes desirable to take the distance between distributions into account, with the aim of neglecting the information lying outside a certain zone. Let $k_\alpha = \max_{j=1,\dots,f(\alpha)} |K_j^\alpha|$ be the maximal number of consistent sources at level α. Denote $E_{K_j^\alpha} = \bigcap_{i \in K_j^\alpha} E_i^\alpha$. At each level α a so-called consensus zone can be defined as the interval:

$$E_{K^\alpha} = \left[\cup_{j,|K_j^\alpha|=k_\alpha} \left(E_{K_j^\alpha}\right)\right] = [\underline{k}_\alpha, \overline{k}_\alpha]$$

where [ ] denote the convex hull of a collection of (possibly) disjoint sets. Now, let $A = [\underline{a}, \overline{a}]$, $B = [\underline{b}, \overline{b}]$ be two intervals. We define the closeness $C(A, B)$ between A and B as

$$C(A, B) = \inf_{a \in A, b \in B} (d(a, b))$$

where $d(a, b)$ is the distance between two points a and b of the space X. Let us note that $C(A, B)$ is not a distance (it does not satisfy triangle inequality), but is a measure of consistency between sets A and B accounting for the metric. Indeed, it will be 0 as soon as $A \cap B \neq \emptyset$. Since the proposed method emphasizes the concept of consistency, this choice looks sensible [4]. Moreover, between two thresholds β_k, β_{k+1}, the closeness $C(E_{K_j^\alpha}, E_{K_i^\alpha})$ between any two sets $E_{K_j^\alpha}, E_{K_i^\alpha}$ $i \neq j$ is an increasing function of α, due to the nestedness of these sets [5].

Like [29], the metric of the space can affect the MCS method by fixing a distance threshold d_0 to the consensus

[4]Genuine distances between sets like the Hausdorff distance are less meaningful in our context.

[5]this would not be true for the Hausdorff distance.

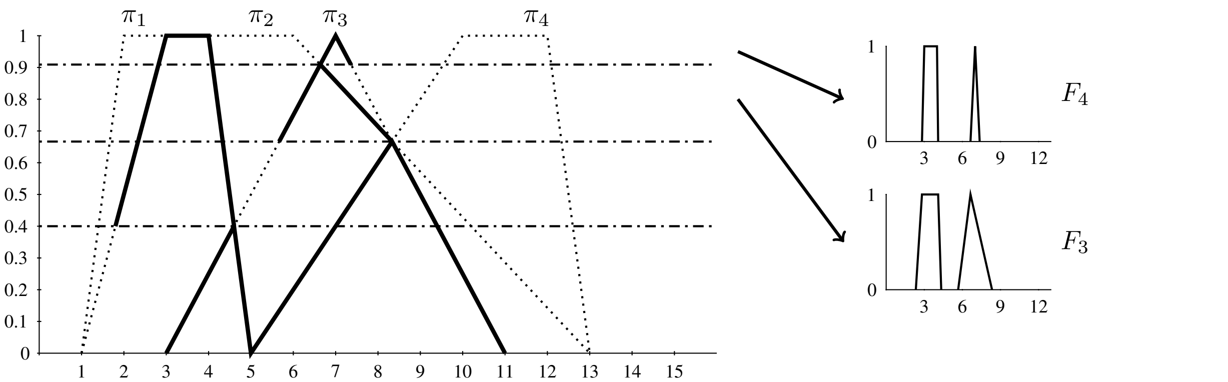

Fig. 10. Result of MCS method with reliability scores $\lambda = (0.2, 0.6, 0.8, 0.2)$ and discarding of poor reliable subgroups $(\perp(x, y) = x + y - xy)$

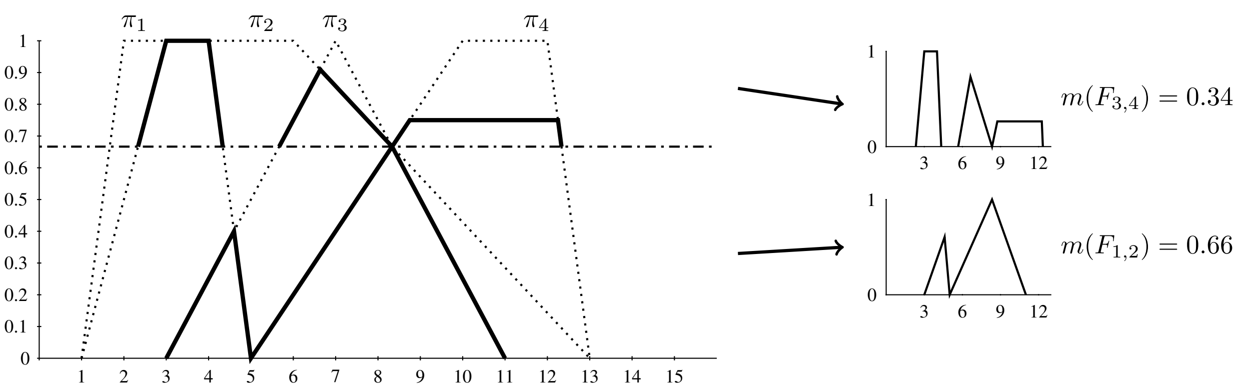

Fig. 11. Result of MCS method taking metric into account with $d_0 = 1$

zone, adapting Equation (6) as

$$E_C^\alpha = \bigcup_{\substack{j=1,\dots,f(\alpha),\\ C(E_{K_j^\alpha}, E_{K^\alpha}) \leq d_0}} \bigcap_{i \in K_j^\alpha} E_i^\alpha \tag{15}$$

Pieces of information away from the consensus zone are regarded as outliers and deleted. Figure 11 illustrates the method when $d_0 = 1$. Except F_1, all fuzzy focal elements are affected by the considered method. In the focal element F_3, distribution π_4 is taken into account until $\alpha = 0.75$ (After this level, $D(E_{K_j^\beta}, K^\beta) < d_0$). In F_2 and F_4, the previous contributions of respectively π_1, π_2 and of π_3 are discarded. Moreover, the structure $(\mathcal{F}, m)$ is simplified and composed of only two fuzzy sets (F_1, F_2 and F_3, F_4).

Except for the discounting technique (which affects the shape of the distributions), all other adaptations result in minor modifications of Equation (6). Thus, the methodology is adapted to available information without modifying its foundamental idea (i.e. using a level-wise notion of MCS). This implies that computational costs of these adaptations are not much higher than the costs of the original method. Except for the variant involving discounting operations (which can lead either to a gain or to a loss of information), all adaptations lead to more informative results, corresponding to the fact that more information is taken into account.

VIII. Conclusions

An adaptive method for merging possibility distributions, based on the notion of maximal coherent subsets is proposed. This method is simple (it can be applied without any additional information, and its computational complexity remains affordable) and the way it summarizes information is conceptually attractive (maximal coherent subsets are the best we can do in the presence of conflict). While most existing fusion rules only aim at directly building a final synthetic distribution from the initial ones, the result of our method is a fuzzy belief structure from which useful information can be extracted. The MCS method satisfies natural requirements expected from an adaptive fusion rule, while avoiding drawbacks of other fusion rules. Additional information concerning source reliability can be accounted for, and outlier information can be discarded from metric considerations if needed.

Close links between fuzzy belief structures and continuous belief structures have been exhibited, thus giving theoretical grounds to the fusion rule. Moreover, these links show how fuzzy random variables can be reinterpreted in term of continuous random sets[6]. This allows to apply results concerning random sets to fuzzy random variable.

[6]Provided fuzzy random variables are interpreted as first-level imprecise probabilistic models

We have proposed various ways to extract useful information from the result of the fusion, making it usable for non-expert analysts. More specifically, we concentrated on how to characterize the situation in term of sources (which sources agree/disagree and to which level, how to measure information gain or overall confidence in each source). This kind of information is useful to figure out where future efforts should be spent (finding the causes of a conflict, or suspecting redundancy of sources, . . .).

We have also proposed a means to get a final distribution coherent with the available information, using the fuzzy belief structure resulting from our method. This allows the decision maker to build a synthetic distribution, easy to understand and to manipulate, which is a good representative of the information delivered by the sources. To summarize, the proposed fusion rule is:

- Simple, generic and conceptually attractive
- Theoretically sound (i.e. not based on ad hoc considerations)
- Flexible
- Useful both for synthesis and analysis of multiple information sources

Axiomatic and theoretical aspects of the MCS method are discussed in this paper. It still remains to validate its use in practical applications in contrast with other fusion rules. We plan to use this method to analyze information issued from the benchmark BEMUSE [39], concerning uncertainty analysis of thermal-hydraulic codes in nuclear safety.

Acknowledgment

This paper has been supported by a grant from the Institut de Radioprotection et de Sûreté Nucléaire (IRSN).

References

[1] R. Clemen and R. Winkler, “Combining probability distributions from experts in risk analysis,” *Risk Analysis*, vol. 19, no. 2, pp. 187–203, 1999.

[2] S. Sandri, D. Dubois, and H. Kalfsbeek, “Elicitation, assessment and pooling of expert judgments using possibility theory,” *IEEE Trans. on Fuzzy Systems*, vol. 3, no. 3, pp. 313–335, August 1995.

[3] P. Walley, *Statistical reasoning with imprecise Probabilities*. New York: Chapman and Hall, 1991.

[4] G. Shafer, *A mathematical Theory of Evidence*. New Jersey: Princeton University Press, 1976.

[5] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press, 1988.

[6] D. Dubois, H. Prade, and R. Yager, *Fuzzy Sets in Approximate Reasoning and Information Systems*, ser. The Handbooks of Fuzzy Sets Series. Boston: Kluwer, 1999, ch. Merging fuzzy information, pp. 335–401.

[7] N. Rescher and R. Manor, “On inference from inconsistent premises,” *Theory and Decision*, vol. 1, pp. 179–219, 1970.

[8] P. Walley, “The elicitation and aggregation of beliefs,” University of Warwick, Tech. Rep., 1982.

[9] S. Moral and J. Sagrado, “Aggregation of imprecise probabilities,” in *Aggregation and Fusion of Imperfect Information*, B. BouchonMeunier, Ed. Heidelberg: Physica-Verlag, 1997, pp. 162–188.

[10] M. Troffaes, “Generalising the conjunction rule for aggregating conflicting expert opinions,” *I. J. of Intelligent Systems*, vol. 21, no. 3, pp. 361–380, March 2006.

[11] A. Ayoun and P. Smets, “Data association in multi-target detection using tramnsferable belief model,” *I. J. of intelligent systems*, vol. 16, no. 10, October 2001.

[12] L. Zadeh, “Fuzzy sets and information granularity,” in *Advances in Fuzzy Sets Theory and Applications*, R. Ragade, M. Gupta, and R. Yager, Eds. Amsterdam: North Holland, 1979, pp. 3–18.

[13] R. Yager, “Generalized probabilities of fuzzy events from fuzzy belief structures,” *Information sciences*, vol. 28, pp. 45–62, 1982.

[14] D. Dubois and H. Prade, “Evidence measures based on fuzzy information,” *Automatica*, vol. 21, no. 5, pp. 547–562, 1985.

[15] J. Yen, “Generalizing the dempster-shafer theory to fuzzy sets,” *IEEE Trans. on Systems, Man and Cybernet.*, vol. 20, no. 3, pp. 559–569, 1990.

[16] T. Denoeux, “Modeling vague beliefs using fuzzy-valued belief structures,” *Fuzzy Sets and Systems*, vol. 116, pp. 167–199, 2000.

[17] D. Dubois, L. Foulloy, G. Mauris, and H. Prade, “Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities,” *Reliable Computing*, vol. 10, pp. 273–297, 2004.

[18] D. Dubois and H. Prade, “When upper probabilities are possibility measures,” *Fuzzy Sets and Systems*, vol. 49, pp. 65–74, 1992.

[19] G. Choquet, “Theory of capacities,” *Annales de l’institut Fourier*, vol. 5, pp. 131–295, 1954.

[20] P. Smets, “The degree of belief in a fuzzy event,” *Information Science*, vol. 25, pp. 1–19, 1981.

[21] C. Baudrit, D. Guyonnet, and D. Dubois, “Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment,” *IEEE Trans. Fuzzy Systems*, vol. 14, pp. 593–608, 2006.

[22] P. Smets, “Belief functions on real numbers,” *I. J. of Approximate Reasoning*, vol. 40, pp. 181–223, 2005.

[23] D. Dubois and H. Prade, “Representation and combination of uncertainty with belief functions and possibility measures,” *Computational Intelligence*, vol. 4, pp. 244–264, 1988.

[24] ——, “Possibility theory in information fusion,” in *Data fusion and Perception*, G. D. Riccia, H. Lenz, and R. Kruse, Eds. Berlin: Springer Verlag, 2001, vol. CISM Courses and Lectures N 431, pp. 53–76.

[25] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*. Berlin: Springer, 2008.

[26] R. Yager, “On ordered wieghted averaging aggregation operators in multicriteria decision making,” *IEEE Trans. on Syst., Man, and Cybern.*, vol. 18, pp. 183–190, 1988.

[27] R. Yager and D. Filev, “Induced ordered weighted averaging operators,” *IEEE Trans. on Systems, Man and Cybernet.*, vol. 29, pp. 141–150, 1999.

[28] D. Dubois and H. Prade, “Combination of fuzzy information in the framework of possibility theory,” in *Data Fusion in Robotics and Machine Intelligence*, M. Abidi and R. Gonzalez, Eds. New York: Academic press, 1992, pp. 481–505.

[29] M. Oussalah, H. Maaref, and C. Barret, “From adaptative to progressive combination of possibility distributions,” *Fuzzy sets and systems*, vol. 139, pp. 559–582, 2003.

[30] M. Oussalah, “On the use of hamacher’s t-norms family for information aggregation,” *Information Science*, vol. 153, pp. 107–154, 2003.

[31] F. Delmotte, “Detection of defective sources in the setting of possibility theory,” *Fuzzy Sets and Systems*, vol. 158, pp. 555–571, 2007.

[32] S. Benferhat, D. Dubois, and H. Prade, “Reasoning in inconsistent stratified knowledge bases,” in *Proc. of the 26th Inter. Symp. on Multi-Valued Logic (ISMVL)*, 1996, pp. 184–189.

[33] D. Dubois, H. Fargier, and H. Prade, “Multi-source information fusion: a way to cope with incoherences,” in *Proc. of French Days on Fuzzy Logic and Applications (LFA)*, Cepadues, Ed. La rochelle: Cepadues, 2000, pp. 123–130.

[34] M. Oussalah, “Study of some algebraic properties of adaptative combination rules,” *Fuzzy sets and systems*, vol. 114, pp. 391–409, 2000.

[35] D. Dubois and H. Prade, “Fuzzy elements in a fuzzy set,” in *Proc. 10th Int. Fuzzy Syst. Assoc. (IFSA) Congress*, 2005, pp. 55–60.

[36] J. Fortin, D. Dubois, and H. Fargier, “Gradual numbers and their application to fuzzy interval analysis,” *IEEE Transactions on Fuzzy Systems*, vol. Accepted for publication, 2006.

[37] M. Delgado, D. Sanchez, M. Martin-Bautista, and M. Vila, “A probabilistic definition of a nonconvex fuzzy cardinality,” *Fuzzy Sets and Systems*, vol. 126, pp. 177–190, 2002.

[38] D. Dubois and H. Prade, “Gradual elements in a fuzzy set,” *Soft Computing*, vol. 12, pp. 165–175, 2008.

[39] OCDE, “Bemuse phase 1,2,3 report,” Nuclear Energy Agency (NEA), Tech. Rep. NEA/CSNI/R(2006), 2006.