

A consistency-specificity trade-off to select source behavior in information fusion

Frédéric Pichon, Sébastien Destercke, Thomas Burger

Abstract—Combining pieces of information provided by several sources without or with little prior knowledge about the behavior of the sources is an old yet still important and rather open problem in belief function theory. In this paper, we propose an approach to select the behavior of sources based on a very general and expressive fusion scheme, that has the important advantage of making clear the assumptions made about the sources. The selection process itself relies on two cornerstones that are the notions of specificity and consistency of a knowledge representation, and that we adapt to the considered fusion scheme. We illustrate our proposal on different examples and show that the proposed approach actually encompasses some important existing fusion strategies.

Index Terms—Dempster-Shafer theory, Information fusion, Consistency, Specificity, Conflict.

I. INTRODUCTION

Determining, from information provided by multiple sources, the actual value taken by an ill-known variable x defined on a space \mathcal{X} is a central problem in many information systems, commonly known as information fusion. As previously argued in [41], [32], [10], the task of information fusion necessarily involves making some (possibly uncertain) assumptions about the relation between the sources of information and about their behavior (*e.g.*, their relevance and truthfulness [32]). A main concern in information fusion is thus to find, or to select, assumptions to make about the source behaviors for the fusion result to be sensible.

Two situations can be distinguished with respect to this problem: either one has some strong knowledge of the source behaviors, inherited from past experiences, or one has only vague or no knowledge about such behaviors. In the former case, depending on the form of the prior experience, *i.e.*, data or expert knowledge, one may resort to some learning procedures [27], [14], [9] or multicriteria aggregations [7], [33] to estimate the behavior of the sources. When there is no or little prior experience with the sources, which is the case treated in this paper, then the selection of an appropriate assumption about source behaviors needs to be based on other considerations.

This paper is an extended and revised version of [31] with full proofs.

F. Pichon is with Université Lille Nord de France, Lille F-59000, France, and also with the Université d'Artois, Béthune F-62400, France. E-mail: Frederic.Pichon@univ-artois.fr

S. Destercke is with Centre National de la Recherche Scientifique, UMR 7253 Heudiasyc, Université de Technologie de Compiègne, Compiègne F-60200, France. E-mail: Sebastien.Destercke@hds.utc.fr

T. Burger is with Université Grenoble-Alpes, CEA (iRSTV/BGE), INSERM (U1038), CNRS (FR3425), Grenoble F-38000, France. E-mail: Thomas.Burger@cea.fr

To tackle this problem, we propose a practical scheme that relies on the theoretical framework introduced by Pichon et al. [32] in the belief function theory [35], [42]. The problem of information fusion has indeed received a lot of attention within this theory [41] that allows for a flexible modeling of uncertainty. Pichon et al. [32] framework proposes to model source possible behaviors by generic sets of hypotheses, and provides a very expressive way of performing information fusion. It also has the advantage of making transparent the assumptions made about the sources, thus resulting in very readable and interpretable rules.

Yet, the framework remains theoretical, and does not propose any practical means to apply it nor ways to select a particular assumption. To make such a selection, we propose to base the selection on two criteria formalizing the two primary features one may seek regarding one's knowledge about x : to make it, first, as specific as possible and, second, as consistent as possible.

Indeed, a fusion result is all the better if it satisfies both criteria. On the one hand, a specific but poorly consistent knowledge is not desirable, as it cannot be trusted: in belief function theory, it is usual to question the fusion results of the (unnormalized) Dempster's rule [4], [35], which assumes all sources to be truthful and relevant [32], when the inconsistency (conflict) [8] resulting from it is too high. How to deal with this inconsistency is itself a hard problem, as shows the literature on conflict management (*e.g.*, [41], [23]). On the other hand, a consistent but poorly specific result is also not desirable, as it is indecisive: this explains why some rules that makes weaker assumptions about the sources, such as the disjunctive rule [11], [39] that corresponds to assuming that at least one source is relevant, despite ensuring better consistency, are seldom used. Note that the goals of consistency and specificity are somehow antagonists, as the more informative the sources are, the more likely they will be conflicting about x true value. Let us remark also that the need to balance between consistency and informativeness of knowledge is present in other fields facing similar issues: for instance, to deal with inconsistency in a logical knowledge base [15], [16], Grant and Hunter propose a stepwise procedure to improve consistency while minimizing information loss.

Our contribution may be summarized as follows: we define the notion of specificity and consistency by relying on the notion of specialization between belief functions and by extending recent works on conflict [8] to Pichon et al. [32] framework. We do this for the cases of single (Section III)

and multiple (Section IV) sources. We then propose (Section V) a practical yet generic fusion scheme to find source assumptions that lead to a good consistency-specificity trade-off of the final result. To our knowledge, this scheme is the first to propose a practical way to apply Pichon et al. [32] theoretical results, and one of the few that explicitly exploits the consistency-specificity trade-off. The scheme is then illustrated in two ways: we propose (Section VI) various instances of sets of assumptions from which the selection can be done, connecting them to existing fusion strategies and shedding some new light on those strategies; we apply (Section VII) our fusion scheme to a nuclear safety problem, in which rule interpretability is as important as the fusion result. Background material is recalled in Section II.

II. PRELIMINARIES

A. Basics of belief functions

Belief functions have been originally introduced by Dempster [4], mainly to model imprecise observations in statistical inferences. In this view, belief functions usually describe some imprecisely known probability distribution of a random variable. Shafer [35] then extended the theory so that belief functions could also model uncertainty related to ill-known but possibly fixed (deterministic) quantities. This interpretation, unrelated to probabilities and statistics, was taken over by Smets and Kennes [42] in the so-called Transferable Belief Model. In this paper, we consider this latter interpretation, called *singular* in [8], where beliefs concern a fixed quantity. This is most often the case in information fusion problems, where one searches a unique true value of some variable of interest.

Accordingly, we assume the beliefs held by an agent about the actual value taken by a variable \mathbf{x} defined on a finite domain \mathcal{X} , to be modeled using belief functions and to be represented using associated mass functions. A mass function $m^{\mathcal{X}}$ on \mathcal{X} is defined as a mapping from the power set $2^{\mathcal{X}}$ to $[0, 1]$ satisfying $\sum_{A \subseteq \mathcal{X}} m^{\mathcal{X}}(A) = 1$. The mass $m^{\mathcal{X}}(A)$ may be understood as the amount of belief given to the assumption that the agent knows that the value of the variable of interest lies somewhere in set A , and nothing more specific [13].

From the mass function are usually defined two uncertainty measures, the belief and plausibility measures, which respectively read for an event $A \subseteq \mathcal{X}$:

$$Bel(A) = \sum_{\emptyset \neq B \subseteq A} m^{\mathcal{X}}(B) \text{ and } Pl(A) = \sum_{B \cap A \neq \emptyset} m^{\mathcal{X}}(B).$$

That is, Bel is the sum of masses of sets that implies A , Pl the sum of masses of sets that are consistent with A . The *contour function* [35] $pl^{\mathcal{X}} : \mathcal{X} \rightarrow [0, 1]$ associated to a mass function $m^{\mathcal{X}}$ is defined by $pl^{\mathcal{X}}(x) = Pl^{\mathcal{X}}(\{x\})$. The *focal sets* of a mass $m^{\mathcal{X}}$ are the subsets A of \mathcal{X} such that $m^{\mathcal{X}}(A) > 0$. We will denote by \mathcal{F} the set of focal sets of $m^{\mathcal{X}}$. The classical notion of set E is modeled by the *categorical mass* $m(E) = 1$. Besides, a mass function $m^{\mathcal{X}}$ is called Bayesian if its focal sets are singletons.

B. Comparing informative contents

The most natural way to compare the informative content of two sets E_1, E_2 is to say that E_1 is more informative than E_2 if $E_1 \subseteq E_2$. This can be extended in several ways to compare the informative content of mass functions in terms of specificity [13]. In a singular interpretation, the most sensible extension is arguably the notion of specialization, that we use here.

Definition 1 (Specialization). *A mass function $m_1^{\mathcal{X}}$ defined on \mathcal{X} with $\mathcal{F}_1 = \{E_1, \dots, E_q\}$ is said to be a specialization of another mass function $m_2^{\mathcal{X}}$ defined on \mathcal{X} with $\mathcal{F}_2 = \{F_1, \dots, F_p\}$ if and only if there exists a non-negative matrix $W = [w_{ij}]$ of size $q \times p$ such that*

$$\begin{aligned} \text{for } i = 1, \dots, q, \quad & \sum_{j=1}^p w_{ij} = m_1^{\mathcal{X}}(E_i), \\ \text{for } j = 1, \dots, p, \quad & \sum_{i=1}^q w_{ij} = m_2^{\mathcal{X}}(F_j) \\ & w_{ij} > 0 \Rightarrow E_i \subseteq F_j. \end{aligned}$$

This relation is denoted by $m_1^{\mathcal{X}} \sqsubseteq m_2^{\mathcal{X}}$ and by $m_1^{\mathcal{X}} \sqsubset m_2^{\mathcal{X}}$ if there is at least a pair i, j such that $w_{ij} > 0$ and $E_i \subset F_j$.

This can be seen as a transfer of mass from each E_i to supersets F_j , w_{ij} denoting the part of $m_1(E_i)$ transferred to F_j . In other words, $m_1^{\mathcal{X}}$ is a specialization of $m_2^{\mathcal{X}}$ if the mass of any focal set F_j of $m_2^{\mathcal{X}}$ can be redistributed among subsets of E_i in $m_1^{\mathcal{X}}$. Let us recall that we have [11]

$$m_1^{\mathcal{X}} \sqsubseteq m_2^{\mathcal{X}} \Rightarrow pl_1^{\mathcal{X}}(x) \leq pl_2^{\mathcal{X}}(x), \quad \forall x \in \mathcal{X}. \quad (1)$$

C. Source behavioral states

Pichon *et al.* [32] framework integrating source behaviors is the following. Assume an agent wants to know the actual value taken by \mathbf{x} based on testimonies provided by several sources of information identified as \mathfrak{s}_i , $1 \leq i \leq K$. These testimonies can be of several forms: a value $x_i \in \mathcal{X}$, a set $A_i \subseteq \mathcal{X}$, a probability distribution p_i on \mathcal{X} , or in the most general form a mass function $m_i^{\mathcal{X}}$ on \mathcal{X} . To interpret those testimonies, the agent must have some knowledge (or make some assumptions) about the behavioral state (referred to as *meta-knowledge* in [32]) of the sources.

In the approach of Pichon *et al.*, the possible elementary behavioral states of a source \mathfrak{s}_i are formalized as a set $\mathcal{H}^i = \{h_1^i, \dots, h_N^i\}$. The set of elementary joint states on sources is therefore the Cartesian product $\mathcal{H}^K := \times_{i=1}^K \mathcal{H}^i$. The state space \mathcal{H}^i can be very general [32] and may include being unreliable, lying, etc. Two common assumptions for which we will use specific notations are the assumptions that a source \mathfrak{s}_i is relevant (R^i) or not ($\neg R^i$), and truthful (T^i) or not ($\neg T^i$). Together, they form the space of possible states $\mathcal{H}^i = \{(T^i, R^i), (T^i, \neg R^i), (\neg T^i, R^i), (\neg T^i, \neg R^i)\}$. Like the testimonies provided by the sources, the meta-knowledge of the agent can be of several forms, the most general one being a mass function defined over \mathcal{H}^K .

We can now detail how the notions of consistency introduced in [8] and of specificity (specialization) can be extended to include source behaviors, and how to then use these extensions to select an assumption about the sources.

III. CONSISTENCY AND SPECIFICITY: SINGLE SOURCE

A. Crisp testimony and sure meta-knowledge

The simplest situation is a source s delivering a testimony of the form $\mathbf{x} \in A$ with $A \subseteq \mathcal{X}$, and being known to be in a state $h \in \mathcal{H}$, with \mathcal{H} the state space of the source. The testimony $\mathbf{x} \in A$ should then be modified according to this state [32]. This transformation can be encoded by a multivalued mapping $\Gamma_A : \mathcal{H} \rightarrow \mathcal{X}$, where $\Gamma_A(h)$ indicates how to interpret the piece of information $\mathbf{x} \in A$ for each possible state h of the source. For instance, if $\mathcal{H} = \{(T, R), (T, \neg R), (\neg T, R), (\neg T, \neg R)\}$ are the possible states of the source, we have for all $A \subseteq \mathcal{X}$

$$\begin{aligned} \Gamma_A(R, T) &= A, & \Gamma_A(\neg R, T) &= \mathcal{X}, \\ \Gamma_A(R, \neg T) &= A^c, & \Gamma_A(\neg R, \neg T) &= \mathcal{X}, \end{aligned} \quad (2)$$

with A^c the complement of A . Eqs. (2) translate that if s is not relevant, it does not bring any information, while if it is not truthful, it declares the opposite of what it knows to be true [29]. If the knowledge about the source state is imprecise and given by $H \subseteq \mathcal{H}$, then the transformation is the image $\Gamma_A(H) := \bigcup_{h \in H} \Gamma_A(h)$ of H by Γ_A .

To measure consistency, we extend the work of Destercke and Burger [8], where any piece of knowledge $\mathbf{x} \in A$ about a variable \mathbf{x} is considered consistent if $A \neq \emptyset$, and inconsistent otherwise. This extends easily to the current framework, a transformed testimony yielding a consistent piece of knowledge on \mathcal{X} when $\Gamma_A(H) \neq \emptyset$, in which case $\mathbf{x} \in A$ is said *H-consistent*, and an inconsistent piece of knowledge when $\Gamma_A(H) = \emptyset$. We can then extend the measure of consistency of $\mathbf{x} \in A$ introduced in [8], to measure *H-consistency* of a testimony $\mathbf{x} \in A$ as the degree $\phi_H : 2^{\mathcal{X}} \rightarrow \{0, 1\}$ such that:

$$\phi_H(A) = \begin{cases} 1 & \text{if } \Gamma_A(H) \neq \emptyset, \\ 0 & \text{if } \Gamma_A(H) = \emptyset. \end{cases}$$

In some way, this consistency measure evaluates whether H is a valid assumption on the source when it provides the testimony $\mathbf{x} \in A$. Consider, for instance, the assumption $h = (R, \neg T)$ corresponding to a relevant and lying source. This assumption will be considered invalid only when the source provides the certainly true testimony $\mathbf{x} \in \mathcal{X}$, as $\Gamma_{\mathcal{X}}(h) = \emptyset$ and $\phi_h(\mathcal{X}) = 0$.

Meta-knowledge can also be characterized in terms of specificity: a piece of meta-knowledge $H_1 \subseteq \mathcal{H}$ will be said *at least as meta-specific* as another piece of meta-knowledge $H_2 \subseteq \mathcal{H}$ when $\Gamma_A(H_1) \subseteq \Gamma_A(H_2)$ for any $A \subseteq \mathcal{X}$, and we will denote it $H_1 \sqsubseteq_{\mathcal{H}} H_2$. For example, the assumption (R, T) is at least as meta-specific as the assumption $(\neg R, T)$. In general, this only induces a partial order over possible states, as for instance none of the two assumptions (R, T)

and $(R, \neg T)$ is more meta-specific than another. Note that we have the relation

$$H_1 \subseteq H_2 \Rightarrow H_1 \sqsubseteq_{\mathcal{H}} H_2, \quad (3)$$

since $H_1 \subseteq H_2 \Rightarrow \Gamma_A(H_2) = \Gamma_A(H_1) \cup (\bigcup_{h \in H_2 \setminus H_1} \Gamma_A(h))$ and thus $\Gamma_A(H_2) \supseteq \Gamma_A(H_1)$. We also have

$$H_1 \sqsubseteq_{\mathcal{H}} H_2 \Rightarrow \phi_{H_1}(A) \leq \phi_{H_2}(A), \quad (4)$$

since either $\Gamma_A(H_1) \neq \emptyset$ thus $\Gamma_A(H_2) \neq \emptyset$ (definition of $\sqsubseteq_{\mathcal{H}}$) and then $\phi_{H_1}(A) = \phi_{H_2}(A)$, or $\Gamma_A(H_1) = \emptyset$ and thus $\phi_{H_1}(A) = 0 \Rightarrow \phi_{H_1}(A) \leq \phi_{H_2}(A)$, $\forall \Gamma_A(H_2)$.

Remark 1. Relation (4) clearly indicates that reaching both consistency and specificity are somewhat opposite goals.

B. Uncertain testimony and meta-knowledge

More generally, both the testimony and the meta-knowledge of the agent may be uncertain. Let $m^{\mathcal{X}}$ be the uncertain testimony and $m^{\mathcal{H}}$ the uncertain meta-knowledge. The knowledge of the agent on \mathcal{X} is then given by the mass function $m[m^{\mathcal{H}}]^{\mathcal{X}}$ defined for all $B \subseteq \mathcal{X}$ as [32]:

$$m[m^{\mathcal{H}}]^{\mathcal{X}}(B) = \sum_{H \subseteq \mathcal{H}} m^{\mathcal{H}}(H) \sum_{A: \Gamma_A(H)=B} m^{\mathcal{X}}(A). \quad (5)$$

This definition is rather general and cover numerous cases, such as Shafer's discounting rule [35], as explained in [32].

The results of the previous section can be extended to this general setting: following [8], the mass function modeling the empty set ($m[m^{\mathcal{H}}]^{\mathcal{X}}(\emptyset) = 1$) can be associated to a complete inconsistent knowledge and a mass function $m[m^{\mathcal{H}}]^{\mathcal{X}}$ whose focal sets have a non-empty intersection can be associated to a totally consistent knowledge. That is, the testimony $m^{\mathcal{X}}$ is totally consistent under meta-knowledge $m^{\mathcal{H}}$ if and only if

$$\bigcap_{\substack{A \in \mathcal{F} \\ H \in \mathcal{F}_{\mathcal{H}}}} \Gamma_A(H) \neq \emptyset, \quad (6)$$

where \mathcal{F} and $\mathcal{F}_{\mathcal{H}}$ denote the sets of focal sets of $m^{\mathcal{X}}$ and $m^{\mathcal{H}}$, respectively. A mass function $m^{\mathcal{X}}$ is then said *$m^{\mathcal{H}}$ -consistent* if and only if (6) holds. Lemma 1 characterizes $m^{\mathcal{H}}$ -consistent testimonies in terms of the contour function.

Lemma 1. $\bigcap_{\substack{A \in \mathcal{F} \\ H \in \mathcal{F}_{\mathcal{H}}}} \Gamma_A(H) \neq \emptyset \Leftrightarrow \exists x \in \mathcal{X}$ such that $pl[m^{\mathcal{H}}]^{\mathcal{X}}(x) = 1$, where $pl[m^{\mathcal{H}}]^{\mathcal{X}}$ is the contour function associated to $m[m^{\mathcal{H}}]^{\mathcal{X}}$ obtained from (5).

Proof. This follows directly from Lemma 1 of [8], when one recognizes that subsets $\Gamma_A(H) \subseteq \mathcal{X}$, such that $A \in \mathcal{F}$ and $H \in \mathcal{F}_{\mathcal{H}}$, are the focal sets of $m[m^{\mathcal{H}}]^{\mathcal{X}}$. \square

A source is thus *$m^{\mathcal{H}}$ -consistent* if it allows us to conclude that at least one value of \mathbf{x} is totally plausible under meta-knowledge $m^{\mathcal{H}}$. Following [8], this characterization of *$m^{\mathcal{H}}$ -consistency* suggests the following definition:

Definition 2 (*$m^{\mathcal{H}}$ -consistency measure*). The measure $\phi_{m^{\mathcal{H}}} : \mathcal{M}^{\mathcal{X}} \rightarrow [0, 1]$ of *$m^{\mathcal{H}}$ -consistency*, where $\mathcal{M}^{\mathcal{X}}$ denotes the set of all mass functions on \mathcal{X} , reads:

$$\phi_{m^{\mathcal{H}}}(m^{\mathcal{X}}) = \max_{x \in \mathcal{X}} pl[m^{\mathcal{H}}]^{\mathcal{X}}(x). \quad (7)$$

Equation (7) is relatively easy to evaluate. Indeed, in (5), the quantity $\sum_{A:\Gamma_A(H)=B} m^{\mathcal{X}}(A)$ is the mass allocated to $B \subseteq \mathcal{X}$ when the source is assumed to be in some state $H \subseteq \mathcal{H}$, i.e., we have

$$m[H]^{\mathcal{X}}(B) = \sum_{A:\Gamma_A(H)=B} m^{\mathcal{X}}(A).$$

Equation (5) can thus be rewritten:

$$m[m^{\mathcal{H}}]^{\mathcal{X}}(B) = \sum_{H \subseteq \mathcal{H}} m^{\mathcal{H}}(H) m[H]^{\mathcal{X}}(B). \quad (8)$$

As Eq. (8) is a convex mixture of mass functions (each $m[H]^{\mathcal{X}}$ is weighted by $m^{\mathcal{H}}(H)$ which is positive and sums up to one), and as the plausibility measure of a convex mixture is the convex mixture of plausibility measures, computing Eq. (7) only requires to compute the weighted average of the contour functions of $m[H]^{\mathcal{X}}$ for all $H \in \mathcal{F}_{\mathcal{H}}$.

Remark 2. Many generalizations of Shannon entropy (see [20, Ch. 3] for a review) have been defined for belief functions, such as the degree of dissonance [20, Sec. 6.5.], which corresponds to a mean value of conflict or inconsistency between the focal sets of a mass function. Using them instead of Def. 2 is tempting, however it should be noticed that most of them would not reach extremal values when $\bigcap_{\substack{A \in \mathcal{F} \\ H \in \mathcal{F}_{\mathcal{H}}}} \Gamma_A(H) \neq \emptyset$ nor when $m[m^{\mathcal{H}}]^{\mathcal{X}}(\emptyset) = 1$.

Remark 3. The more classical consistency measure $m(\emptyset)$ is also considered in [8]. We will not consider it here, as (1) it is argued in [8] that this measure is less adapted to a singular interpretation, (2) estimating $m(\emptyset)$ usually requires heavier computations, hence is of less practical interest and (3) most results presented here adapt easily to $m(\emptyset)$.

Meta-specificity may also be generalized to this setting:

Definition 3 (Meta-specificity). An uncertain piece of meta-knowledge $m_1^{\mathcal{H}}$ is said to be at least as meta-specific as another uncertain piece $m_2^{\mathcal{H}}$ when $m[m_1^{\mathcal{H}}]^{\mathcal{X}} \sqsubseteq m[m_2^{\mathcal{H}}]^{\mathcal{X}}$ for any $m^{\mathcal{X}} \in \mathcal{M}^{\mathcal{X}}$. This is denoted by $m_1^{\mathcal{H}} \sqsubseteq_{\mathcal{H}} m_2^{\mathcal{H}}$.

We may then show that in the general case we have relations extending (3) and (4); in particular that consistency and specificity are also at odds as shown by Proposition 2:

Proposition 1. Let $m_1^{\mathcal{H}}, m_2^{\mathcal{H}} \in \mathcal{M}^{\mathcal{H}}$ such that $m_1^{\mathcal{H}} \sqsubseteq m_2^{\mathcal{H}}$. We have $m_1^{\mathcal{H}} \sqsubseteq_{\mathcal{H}} m_2^{\mathcal{H}}$.

Proof. Consider a focal set A of a mass $m^{\mathcal{X}}$ and a focal set H_j of $m_2^{\mathcal{H}}$. The mass $m^{\mathcal{X}}(A)m_2^{\mathcal{H}}(H_j)$ is then affected to $\Gamma_A(H_j)$. Now, considering the states H_i of $m_1^{\mathcal{H}}$, we have that a fraction $w_{ij} > 0$ of the mass $m^{\mathcal{X}}(A)m_1^{\mathcal{H}}(H_i)$ affected to $\Gamma_A(H_i)$ will be transferred to $\Gamma_A(H_j)$, and $\Gamma_A(H_i) \subseteq \Gamma_A(H_j)$ since $H_i \subseteq H_j$ by definition. \square

Proposition 2. Let $m_1^{\mathcal{H}}, m_2^{\mathcal{H}} \in \mathcal{M}^{\mathcal{H}}$ such that $m_1^{\mathcal{H}} \sqsubseteq_{\mathcal{H}} m_2^{\mathcal{H}}$. We have $\phi_{m_1^{\mathcal{H}}}(m^{\mathcal{X}}) \leq \phi_{m_2^{\mathcal{H}}}(m^{\mathcal{X}})$ for all $m^{\mathcal{X}} \in \mathcal{M}^{\mathcal{X}}$.

Proof. From the definition of $\sqsubseteq_{\mathcal{H}}$, we have $m[m_1^{\mathcal{H}}]^{\mathcal{X}} \sqsubseteq m[m_2^{\mathcal{H}}]^{\mathcal{X}}$ for all $m^{\mathcal{X}} \in \mathcal{M}^{\mathcal{X}}$, which from (1) implies $pl[m_1^{\mathcal{H}}]^{\mathcal{X}}(x) \leq pl[m_2^{\mathcal{H}}]^{\mathcal{X}}(x)$ for all $x \in \mathcal{X}$. \square

The notions introduced in this section, and in particular Proposition 2, are illustrated in Example 1.

Example 1 (Inspired from Example 1 of [32]). Let $\mathcal{X} = \{x_1, x_2, x_3, x_4, x_5\}$ be an ordered space and consider $m^{\mathcal{X}}$ such that $m^{\mathcal{X}}(\{x_1, x_2\}) = 0.3$, $m^{\mathcal{X}}(\{x_4, x_5\}) = 0.3$ and $m^{\mathcal{X}}(\{x_3\}) = 0.4$. Now consider the following assumptions:

- h_1 “reliable” such that $\Gamma_A(h_1) = A$;
- h_3 “unreliable” such that $\Gamma_A(h_3) = \mathcal{X}$;
- and h_2 “approximately reliable” such that if $A = \{x_i, x_{i+1}, \dots, x_j\}$ then $\Gamma_A(h_2) = \{x_{i-1}\} \cup A \cup \{x_{j+1}\}$ with $x_0 = x_6 = \emptyset$ and meaning that the source is not totally reliable neither totally unreliable, but it is somewhere between these two extremes.

We have $h_1 \sqsubseteq_{\mathcal{H}} h_2 \sqsubseteq_{\mathcal{H}} h_3$ while $\phi_{h_1}(m^{\mathcal{X}}) = 0.4$, $\phi_{h_2}(m^{\mathcal{X}}) = 1$, $\phi_{h_3}(m^{\mathcal{X}}) = 1$.

This example allows us to lay bare some preliminary ideas on source behavior selection with a consistency-specificity trade-off. As observed, assumptions h_2 and h_3 are the most desirable in terms of consistency, since they both yield a totally consistent state of knowledge. However, as the state of knowledge obtained under h_2 is more specific (informative) than the one obtained under h_3 , h_2 appears preferable. This will be developed at length in Section V.

IV. CONSISTENCY AND SPECIFICITY: MULTIPLE SOURCES

Let us now consider the main case where multiple sources \mathfrak{s}_i , $i = 1, \dots, K$ provide information, each as a mass function $m_i^{\mathcal{X}}$. As recalled in the Introduction, the main problem in such a case is to combine these pieces of information in a sensible way. Since we focus in this paper on the problem of finding appropriate source behaviors, the sources will be assumed to rely on distinct evidences [38], [41], a usual assumption when merging belief functions.

Many combination rules have been proposed for belief functions [41]: the most usual is the unnormalized Dempster’s rule (or conjunctive rule), which applies when sources are assumed to rely on distinct evidences and are both relevant and truthful. It is denoted here by \odot . The mass $m_{1 \odot 2}^{\mathcal{X}}$ resulting from its application on $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$ is:

$$m_{1 \odot 2}^{\mathcal{X}}(A) = \sum_{B \cap C = A} m_1^{\mathcal{X}}(B) m_2^{\mathcal{X}}(C), \quad \forall A \subseteq \mathcal{X}. \quad (9)$$

The disjunctive rule \oplus [11], [39] is obtained by simply replacing \cap with \cup in (9). Both the conjunctive and disjunctive rules can be given a clear interpretation in terms of source behavior assumptions [32]. In this section, we extend our characterization of consistency and specificity to the theoretical results of Pichon *et al.* [32].

A. General case: uncertain testimonies and meta-knowledge

To simplify notations, we consider that all source \mathfrak{s}_i share the same¹ possible state space $\mathcal{H} = \{h_1, \dots, h_N\}$. For any state $\mathbf{h} = (h^1, \dots, h^k) \in \mathcal{H}^K$ we define a mapping for

¹The extension to particularized state spaces \mathcal{H}^i is straightforward.

any $\mathbf{A} = (A_1, \dots, A_K) \subseteq \mathcal{X}^K$ as $\Gamma_{\mathbf{A}}(\mathbf{h}) = \bigcap_{i=1}^K \Gamma_{A_i}(h^i)$. $\Gamma_{\mathbf{A}}(\mathbf{h})$ represents the information on \mathcal{X} deduced from testimonies (A_1, \dots, A_K) provided by sources $\mathfrak{s}_1, \dots, \mathfrak{s}_K$ when they are in states (h^1, \dots, h^K) [32]. For non-elementary hypotheses $H \subseteq \mathcal{H}^K$, we keep the previous notation, *i.e.*, $\Gamma_{\mathbf{A}}(H) := \bigcup_{\mathbf{h} \in H} \Gamma_{\mathbf{A}}(\mathbf{h})$, $\forall H \subseteq \mathcal{H}^K$, $\forall \mathbf{A} \subseteq \mathcal{X}^K$.

If we have some joint meta-knowledge $m^{\mathcal{H}^K}$ over \mathcal{H}^K and if sources $\mathfrak{s}_1, \dots, \mathfrak{s}_K$ deliver distinct testimonies $m_i^{\mathcal{X}}$, $i = 1, \dots, K$, then the combined mass function $m[m^{\mathcal{H}^K}]^{\mathcal{X}}$ defined by (10) represents what can be inferred about \mathbf{x} from $\mathbf{m}^{\mathcal{X}} = (m_1^{\mathcal{X}}, \dots, m_K^{\mathcal{X}})$ [32]:

$$m[m^{\mathcal{H}^K}]^{\mathcal{X}}(B) = \sum_{H \subseteq \mathcal{H}^K} m^{\mathcal{H}^K}(H) \sum_{\substack{\mathbf{A} \subseteq \mathcal{X}^K \\ \Gamma_{\mathbf{A}}(H)=B}} \left[\prod_{i=1}^K m_i^{\mathcal{X}}(A_i) \right], \quad (10)$$

the product $\prod_{i=1}^K m_i^{\mathcal{X}}(A_i)$ coming from the distinctness assumption.

An interesting feature of using meta-knowledge is that all Boolean operators on sets $\mathbf{A} = (A_1, \dots, A_K) \subseteq \mathcal{X}^K$ can be obtained through particular assumptions on the behavior of the sources [32]. As a result, Equation (10) covers all combination rules based on Boolean operators. For instance, consider the assumption H_r^K on \mathcal{H}^K meaning the sources are truthful and “r-out-of-K” of them are relevant (an assumption that is common in interval analysis [17]). This amounts to

$$\Gamma_{\mathbf{A}}(H_r^K) = \bigcup_{\substack{\mathcal{A} \subseteq \{A_1, \dots, A_K\}, |\mathcal{A}|=r}} (\bigcap_{A \in \mathcal{A}} A). \quad (11)$$

When applying H_r^K to Eq. (10), the conjunctive and disjunctive rules are retrieved when $r = K$ and $r = 1$, respectively.

Keeping the same definition of complete inconsistent and consistent knowledge as in Section III-B, the counterpart of Lemma 1 suggests to use the following equation as a degree of $m^{\mathcal{H}^K}$ -consistency for the collection $\mathbf{m}^{\mathcal{X}}$

$$\phi_{m^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) = \max_{x \in \mathcal{X}} pl[m^{\mathcal{H}^K}]^{\mathcal{X}}(x), \quad (12)$$

where $pl[m^{\mathcal{H}^K}]^{\mathcal{X}}$ is the contour function of (10). Eq. (12) extends the conflict measure defined by Destercke and Burger [8] to any combination rule that can be obtained from (10), in particular to all rules based on Boolean operators. In addition, we have again that if $m_1^{\mathcal{H}^K} \sqsubseteq_{\mathcal{H}} m_2^{\mathcal{H}^K}$, then $\phi_{m_1^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \leq \phi_{m_2^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}})$ for $\mathbf{m}^{\mathcal{X}}$.

Those powerful results, *i.e.*, the extension of the conflict measure to any combination rule and the duality between specificity and consistency are heavily used in Section V where we introduce a method to select source behaviors.

B. Computation of the consistency degree

Equation (10) can be rewritten as

$$m[m^{\mathcal{H}^K}]^{\mathcal{X}}(B) = \sum_{H \subseteq \mathcal{H}^K} m^{\mathcal{H}^K}(H) m[H]^{\mathcal{X}}(B), \quad (13)$$

with $m[H]^{\mathcal{X}}$ the mass function representing the knowledge inferred on \mathcal{X} when the sources are assumed to be in some

state $H \subseteq \mathcal{H}^K$, defined by

$$m[H]^{\mathcal{X}}(B) = \sum_{\mathbf{A} \subseteq \mathcal{X}^K; \Gamma_{\mathbf{A}}(H)=B} \left[\prod_{i=1}^K m_i^{\mathcal{X}}(A_i) \right].$$

Hence, the computation of (12), which can be resource demanding, may be simplified as in the single source case, using the convexity property of plausibility measures. However, there are cases where it is easier to compute (13) and thus even easier to compute (12) than merely using this convexity property. In particular, when all focal elements of $m^{\mathcal{H}^K}$ are *behavior-separable* (or *b-separable*).

Definition 4 (Behavior-Separability). *A subset $H \subseteq \mathcal{H}^K$ is said b-separable iff $H = H^{\downarrow 1} \times \dots \times H^{\downarrow K}$ (where $H^{\downarrow i}$ is the projection of $H \subseteq \mathcal{H}^K$ on the i th source state space).*

Proposition 3. *When each focal set of $m^{\mathcal{H}^K}$ is b-separable, Equation (13) can be rewritten as:*

$$m[m^{\mathcal{H}^K}]^{\mathcal{X}}(B) = \sum_{H \subseteq \mathcal{H}^K} m^{\mathcal{H}^K}(H) [\bigcirc_{i=1}^K m[H^{\downarrow i}]^{\mathcal{X}}](B) \quad (14)$$

where $m[H^{\downarrow i}]^{\mathcal{X}}$ denotes $m_i^{\mathcal{X}}$ transformed according to $H^{\downarrow i}$.

Proof. Let H be a b-separable assumption on the sources. Such b-separable meta-knowledge H satisfies the property of so-called meta-independence in [32]. Therefore, from [32, Theorem 1], we have

$$m[H]^{\mathcal{X}}(B) = [\bigcirc_{i=1}^K m[H^{\downarrow i}]^{\mathcal{X}}](B). \quad (15)$$

If each focal set of a piece of meta-knowledge $m^{\mathcal{H}^K}$ is b-separable, it is direct to obtain (14) from (13) and (15). \square

That is, to compute $m[m^{\mathcal{H}^K}]^{\mathcal{X}}$ we first transform each $m_i^{\mathcal{X}}$ according to $H^{\downarrow i}$, apply unnormalized Dempster’s rule to them and compute the weighted sum according to $m^{\mathcal{H}^K}$. We can therefore make use of efficient algorithms to compute Dempster’s rule result [43].

This property simplifies the computation of the consistency measure (12) as follows. Consider the meta-knowledge $m^{\mathcal{H}^K}(H) = 1$ with H b-separable and let $pl[H]^{\mathcal{X}}$ be the corresponding contour function. Furthermore, let $pl[H^{\downarrow i}]^{\mathcal{X}}$ denote the contour function obtained by transforming $m_i^{\mathcal{X}}$ according to meta-knowledge $H^{\downarrow i}$. We have:

$$pl[H]^{\mathcal{X}}(x) = \prod_{i=1}^K pl[H^{\downarrow i}]^{\mathcal{X}}(x). \quad (16)$$

Now, let $m^{\mathcal{H}^K}$ be a piece of meta-knowledge with b-separable focal sets, we have then

$$pl[m^{\mathcal{H}^K}]^{\mathcal{X}}(x) = \sum_{H \subseteq \mathcal{H}^K} m^{\mathcal{H}^K}(H) \cdot \prod_{i=1}^K pl[H^{\downarrow i}]^{\mathcal{X}}(x).$$

In other words, when each focal set of a piece of meta-knowledge is b-separable, computing consistency measure (12) only requires to compute contour functions and to

take their weighted averaged products (hence not necessitating any combination). However, let us stress that in general,

$$\phi_{m^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \neq \sum_H m^{\mathcal{H}^K}(H) \phi_H(\mathbf{m}^{\mathcal{X}}), \quad (17)$$

even in the case where the focal sets of $m^{\mathcal{H}^K}$ are b-separable, as shown by Example 2.

Example 2. Let $\mathbf{m}^{\mathcal{X}} = (m_1^{\mathcal{X}}, m_2^{\mathcal{X}})$ with $\mathcal{X} = \{x_1, x_2\}$ and $m_1^{\mathcal{X}}(\{x_1\}) = 0.2, m_1^{\mathcal{X}}(\{x_2\}) = 0.3, m_1^{\mathcal{X}}(\mathcal{X}) = 0.5$ and $m_2^{\mathcal{X}}(\{x_1\}) = 0.4, m_2^{\mathcal{X}}(\{x_2\}) = 0.2, m_2^{\mathcal{X}}(\mathcal{X}) = 0.4$. Furthermore, sources \mathfrak{s}_1 and \mathfrak{s}_2 are assumed to be truthful and to have the following joint behavior in terms of relevance: with mass 0.7 \mathfrak{s}_1 is relevant and \mathfrak{s}_2 is not, and with mass 0.3 \mathfrak{s}_2 is relevant and \mathfrak{s}_1 is not. Let $m^{\mathcal{H}^2}$ denote such meta-knowledge on the sources:

$$\begin{aligned} m^{\mathcal{H}^2}(H_1 = \{((R, T), (-R, T))\}) &= 0.7, \\ m^{\mathcal{H}^2}(H_2 = \{((-R, T), (R, T))\}) &= 0.3. \end{aligned}$$

We have [32] $m[m^{\mathcal{H}^2}]^{\mathcal{X}} = 0.7 \cdot m_1^{\mathcal{X}} + 0.3 \cdot m_2^{\mathcal{X}}$. Computing the contour function associated to $m[m^{\mathcal{H}^2}]^{\mathcal{X}}$, we find:

$$\begin{aligned} pl[m^{\mathcal{H}^2}]^{\mathcal{X}}(x_1) &= 0.7 \cdot pl_1^{\mathcal{X}}(x_1) + 0.3 \cdot pl_2^{\mathcal{X}}(x_1) = 0.73, \\ pl[m^{\mathcal{H}^2}]^{\mathcal{X}}(x_2) &= 0.7 \cdot pl_1^{\mathcal{X}}(x_2) + 0.3 \cdot pl_2^{\mathcal{X}}(x_2) = 0.74, \end{aligned}$$

and thus $\phi_{m^{\mathcal{H}^2}}(\mathbf{m}^{\mathcal{X}}) = 0.74$. On the other hand, we have

$$\begin{aligned} m^{\mathcal{H}^2}(H_1) \cdot \phi_{\{H_1\}}(\mathbf{m}^{\mathcal{X}}) + m^{\mathcal{H}^2}(H_2) \cdot \phi_{H_2}(\mathbf{m}^{\mathcal{X}}) \quad (18) \\ = 0.7 \cdot 0.8 + 0.3 \cdot 0.8 = 0.8. \end{aligned}$$

Such a behavior can be easily explained by the fact that the sum operator of (17) and the maximum operator of (12) do not distribute over each other.

V. SOURCE BEHAVIOR SELECTION APPROACH

When only little knowledge about the sources is available, it is not possible to estimate their behavior using learning procedures or multicriteria aggregations (see Section I), as information to do so is lacking. In such a case, a strategy is to consider a set of sensible behavior assumptions – hence a set of readable rules – to choose from, together with selection criteria. We provide guidelines to define such a set and selection criterion, based on previous section materials. Roughly speaking, we specify a set of behavior assumptions inducing decreasingly specific results, and take a minimal consistency threshold as a simple, single selection parameter to pick an assumption from this set. This leads to a general, yet practical and sensible, approach to select the behavior of the sources in poorly informed cases. Instances of Section VI illustrate the approach.

A. Initial meta-knowledge

We propose to consider a basic initial assumption $m_1^{\mathcal{H}^K}$ such that $m_1^{\mathcal{H}^K}(\mathbf{h}) = 1$, with $\mathbf{h} \in \mathcal{H}^K$ and $\Gamma_A(\mathbf{h}^{\downarrow i}) = A$ for all $A \subseteq \mathcal{X}$ and $i = 1, \dots, K$, i.e., an assumption that induces no transformation of the testimonies provided by the sources.

This assumption corresponds to not altering in any way the initial information, i.e., we accept the testimonies as they are. Most importantly, this assumption is the most classical one in information fusion in general and in belief function theory in particular, as it corresponds to unnormalized Dempster's rule. Hence, $m_1^{\mathcal{H}^K}$ is a natural default meta-knowledge.

Equation (12) gives an assessment of whether the assumption $m_1^{\mathcal{H}^K}$ applies to the current testimonies. As is classically advocated in belief function theory, we propose that assumption $m_1^{\mathcal{H}^K}$ should be used to combine the testimonies if the consistency derived from Eq. (12) induced by $m_1^{\mathcal{H}^K}$ is high enough, that is if it is above some threshold τ , and that $m_1^{\mathcal{H}^K}$ should be rejected as a valid assumption if the consistency is too low, i.e., below τ . In this latter case, other assumptions leading to higher consistency should be sought.

B. A specificity ordering approach

To define other assumptions that will result in more consistent results after merging, we use the counterpart of Proposition 2 in the multiple source case: choosing a meta-knowledge $m_2^{\mathcal{H}^K}$ such that $m_1^{\mathcal{H}^K} \sqsubseteq_{\mathcal{H}} m_2^{\mathcal{H}^K}$ will ensure a consistency increment. This leads us to propose the following two-step strategy to select the meta-knowledge to be used:

- 1) define a collection of meta-knowledge $\mathbf{m}^{\mathcal{H}^K} = (m_1^{\mathcal{H}^K}, \dots, m_M^{\mathcal{H}^K})$ such that for any $1 \leq j < M$, $m_j^{\mathcal{H}^K} \sqsubset_{\mathcal{H}} m_{j+1}^{\mathcal{H}^K}$, and with $m_1^{\mathcal{H}^K}$ as defined above;
- 2) test each $m_j^{\mathcal{H}^K}$ iteratively with $j = 1, \dots, M$, until $\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \geq \tau$.

This ensures that, at each iteration from j to $j+1$, specificity will decrease since $m_j^{\mathcal{H}^K} \sqsubset_{\mathcal{H}} m_{j+1}^{\mathcal{H}^K}$ and consistency will increase since $\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \leq \phi_{m_{j+1}^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}})$, the process stopping whenever one thinks the result is consistent (trustworthy) enough. In other words, this strategy gradually decreases specificity until a satisfactory consistency level is reached. Once a collection is defined (some interesting instances are given in Section VI), the only parameter to set is τ (any value over 0.5 seems reasonable, with 0.8 being a good compromise).

In itself, this idea is not entirely new, as it already appears in Dubois and Prade [12], where the foundations of source behavior assumptions are laid bare. More recently, one can find specificity-based comparisons of combination rules in [37], or strategies relaxing specificity to gain consistency [22], [21]. Yet, to our knowledge this is the first proposal to provide a so generic formal procedure allowing one to select interpretable fusion rules and to rely on consistency measured through contour functions.

Remark 4. The construction of $\mathbf{m}^{\mathcal{H}^K}$ should be based on common-sense and readability: pieces of meta-knowledge $m_j^{\mathcal{H}^K}$ should have a clear semantic given the application, and the space \mathcal{H} should be of reduced size, e.g., $\mathcal{H} = \{(T, R), (T, \neg R), (\neg T, R), (\neg T, \neg R)\}$.

TABLE I
MASS FUNCTIONS RESULTING FROM THE THREE DIFFERENT
ASSUMPTIONS.

A	$m_1^{\mathcal{X}}$	$m_2^{\mathcal{X}}$	$m_3^{\mathcal{X}}$	$m[H_1^3]^{\mathcal{X}}$	$m[H_2^3]^{\mathcal{X}}$	$m[H_3^3]^{\mathcal{X}}$
\emptyset	0	0	0	0	0	0.36
$\{x_1\}$	0.5	0	0	0	0.06	0.2
$\{x_1, x_2\}$	0	0.2	0	0	0.04	0.04
$\{x_3\}$	0	0	0.6	0	0	0.24
$\{x_1, x_3\}$	0	0	0	0	0.24	0
\mathcal{X}	0.5	0.8	0.4	1	0.66	0.16

VI. PRACTICAL INSTANCES

Before providing a real world case study in Section VII, we provide three peculiar instances illustrating the approach that are of practical and theoretical interest: the first one uses imprecise pieces of meta-knowledge; the second one is based on Bayesian pieces of meta-knowledge; and the third one explores an interesting link between our approach and the α -conjunctions [40]. Whenever possible, we relate to classical fusion strategies that our approach subsumes.

A. r -out-of- K relevant sources

The collection is defined by $m_j^{\mathcal{H}^K}(H_{K-j+1}^K) = 1$, with H_{K-j+1}^K being the assumption that all the sources are truthful and $r = K - j + 1$ out of them are relevant (see Eq. (11)). This collection satisfies our need:

Proposition 4. *If $m_j^{\mathcal{H}^K}(H_{K-j+1}^K) = 1$, then $m_j^{\mathcal{H}^K} \sqsubseteq_{\mathcal{H}} m_{j+1}^{\mathcal{H}^K}$ for $1 \leq j < K$.*

Proof. For each $j \in \{1, \dots, K-1\}$ and each $\mathbf{h} \in H_{K-j+1}^K$ (\mathbf{h} is a joint state assuming that the K sources are truthful and that $K-j+1$ specific sources are relevant), there exists an assumption $\mathbf{h}' \in H_{K-j}^K$, such that

$$\Gamma_{\mathbf{A}}(\mathbf{h}) \subseteq \Gamma_{\mathbf{A}}(\mathbf{h}') \quad (19)$$

for all $\mathbf{A} \subseteq \mathcal{X}^K$. Since $\Gamma_{\mathbf{A}}(H_{K-j+1}^K) = \bigcup_{\mathbf{h} \in H_{K-j+1}^K} \Gamma_{\mathbf{A}}(\mathbf{h})$ and $\Gamma_{\mathbf{A}}(H_{K-j}^K) = \bigcup_{\mathbf{h}' \in H_{K-j}^K} \Gamma_{\mathbf{A}}(\mathbf{h}')$ for all $\mathbf{A} \subseteq \mathcal{X}^K$, we have, using (19), $\Gamma_{\mathbf{A}}(H_{K-j+1}^K) \subseteq \Gamma_{\mathbf{A}}(H_{K-j}^K)$ for all $\mathbf{A} \subseteq \mathcal{X}^K$, i.e., H_{K-j+1}^K is at least as meta-specific as H_{K-j}^K . \square

Example 3. Consider the mass functions $m_1^{\mathcal{X}}$, $m_2^{\mathcal{X}}$ and $m_3^{\mathcal{X}}$ on $\mathcal{X} = \{x_1, x_2, x_3\}$ in the left part of Table I. Assume they were received from three distinct sources. Let $\mathbf{m}^{\mathcal{H}^3} = (m_1^{\mathcal{H}^3}, m_2^{\mathcal{H}^3}, m_3^{\mathcal{H}^3}) = (H_3^3, H_2^3, H_1^3)$ be three pieces of meta-knowledge we want to test on these sources. $m_1^{\mathcal{H}^3}$ corresponds to the use of the unnormalized Dempster's rule, while $m_3^{\mathcal{H}^3}$ corresponds to the use of the disjunctive rule. $m_2^{\mathcal{H}^3}$ corresponds to the assumption H_2^3 that the three sources are truthful and that two of them are relevant, but we do not know which ones.

The right part of Table I presents the mass functions on \mathcal{X} resulting from the three different assumptions. We have $\phi_{H_1^3}(\mathbf{m}^{\mathcal{X}}) = 1$, $\phi_{H_2^3}(\mathbf{m}^{\mathcal{X}}) = 1$ and $\phi_{H_3^3}(\mathbf{m}^{\mathcal{X}}) = 0.4$, hence our approach suggests to use H_2^3 to combine the pieces of information in this example.

Note that the assumption “ r -out-of- K ” is not b-separable in general. However, this assumption treats all sources in the same way, which seems interesting in absence of meta-knowledge about each individual source.

B. Partially relevant sources

Another interesting case is when we consider $\mathcal{H} = \{R, \neg R\}$ (relevant or not) and a vector $\mathbf{p} = (p_1, \dots, p_K)$ such that $m^{\mathcal{H}}(R^i) = p_i$, $m^{\mathcal{H}}(\neg R^i) = 1 - p_i$ and where $m^{\mathcal{H}^K}$ is obtained by considering the stochastic product of probabilities p_1, \dots, p_K . In such case, the assumption $m^{\mathcal{H}^K}$ amounts to discounting each source s_i according to reliability rate $1 - p_i$ and then combining the discounted sources using the conjunctive rule [32]. If we define a set $\mathbf{p}^1, \dots, \mathbf{p}^M$ of such vectors with $p_i^j \geq p_i^{j+1}$ with the inequality strict for at least one i , we get corresponding pieces of meta-knowledge $m_1^{\mathcal{H}^K}, \dots, m_M^{\mathcal{H}^K}$ that satisfy the property described by Proposition 5 below, the proof of which requires Lemmas 2 and 3.

Lemma 2. [6, Proposition 2] *The conjunctive rule is monotonic with respect to \sqsubseteq , i.e., for all mass functions $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$ on \mathcal{X} such that $m_1^{\mathcal{X}} \sqsubseteq m_2^{\mathcal{X}}$, we have*

$$m_1^{\mathcal{X}} \odot m_3^{\mathcal{X}} \sqsubseteq m_2^{\mathcal{X}} \odot m_3^{\mathcal{X}}, \quad \forall m_3^{\mathcal{X}}.$$

Lemma 3. *Let $m_k^{\mathcal{X}}$, $m'_k{}^{\mathcal{X}}$ ($k = 1, \dots, K$) be $2K$ mass functions on \mathcal{X} such that $m_k^{\mathcal{X}} \sqsubseteq m'_k{}^{\mathcal{X}}$. We have*

$$\bigodot_{k=1}^K m_k^{\mathcal{X}} \sqsubseteq \bigodot_{k=1}^K m'_k{}^{\mathcal{X}}. \quad (20)$$

Proof. Since $m_k^{\mathcal{X}} \sqsubseteq m'_k{}^{\mathcal{X}}$, $k = 1, \dots, K$, Equation (20) holds for $K = 1$. Assume now that (20) holds for $K = N$. To prove this lemma, it suffices then to show that (20) holds for $K = N + 1$. From Lemma 2, we have

$$\begin{aligned} & \bigodot_{k=1}^N m_k^{\mathcal{X}} \sqsubseteq \bigodot_{k=1}^N m'_k{}^{\mathcal{X}} \\ \Rightarrow & (\bigodot_{k=1}^N m_k^{\mathcal{X}}) \odot m_{N+1}^{\mathcal{X}} \sqsubseteq (\bigodot_{k=1}^N m'_k{}^{\mathcal{X}}) \odot m_{N+1}^{\mathcal{X}}, \end{aligned} \quad (21)$$

as well as

$$\begin{aligned} & m_{N+1}^{\mathcal{X}} \sqsubseteq m'_{N+1}{}^{\mathcal{X}} \\ \Rightarrow & (\bigodot_{k=1}^N m'_k{}^{\mathcal{X}}) \odot m_{N+1}^{\mathcal{X}} \sqsubseteq (\bigodot_{k=1}^N m'_k{}^{\mathcal{X}}) \odot m'_{N+1}{}^{\mathcal{X}}. \end{aligned} \quad (22)$$

Eq. (21) and (22) lead to $\bigodot_{k=1}^{N+1} m_k^{\mathcal{X}} \sqsubseteq \bigodot_{k=1}^{N+1} m'_k{}^{\mathcal{X}}$ \square

Proposition 5. *Let $m_j^{\mathcal{H}^K}$, $j = 1, \dots, M$, be the mass functions defined using $\mathbf{p}^1, \dots, \mathbf{p}^M$. We have $m_j^{\mathcal{H}^K} \sqsubseteq_{\mathcal{H}} m_{j+1}^{\mathcal{H}^K}$, for $1 \leq j < M$.*

Proof. Let $m[p_i^j]$ denote the mass function $m_i^{\mathcal{X}}$ on \mathcal{X} discounted according to reliability rate $1 - p_i^j$ (see [35], [39]). If $p_i^j \geq p_i^{j+1}$, then it is immediate that

$$m[p_i^j] \sqsubseteq m[p_i^{j+1}]. \quad (23)$$

From Lemma 3 and Equation (23), we obtain

$$\bigodot_{i=1}^K m[p_i^j] \sqsubseteq \bigodot_{i=1}^K m[p_i^{j+1}]. \quad (24)$$

As shown in [32], we have

$$\bigodot_{i=1}^K m[p_i^j] = m[m_j^{\mathcal{H}^K}]^{\mathcal{X}}, \quad (25)$$

TABLE II
MASS FUNCTIONS RESULTING FROM THREE RELIABILITY VECTORS OF
EXAMPLE 4.

A	$m_1^{\mathcal{X}}$	$m_2^{\mathcal{X}}$	$m_3^{\mathcal{X}}$	$m[\mathbf{p}^3]^{\mathcal{X}}$	$m[\mathbf{p}^2]^{\mathcal{X}}$	$m[\mathbf{p}^1]^{\mathcal{X}}$
\emptyset	0	0	0	0.04	0.18	0.36
$\{x_1\}$	0.5	0	0	0.26	0.26	0.2
$\{x_1, x_2\}$	0	0.2	0	0.08	0.06	0.04
$\{x_3\}$	0	0	0.6	0.08	0.18	0.24
\mathcal{X}	0.5	0.8	0.54	1	0.32	0.16

with $m_j^{\mathcal{H}^K}$ the meta-knowledge obtained by taking the stochastic product of probabilities p_1^j, \dots, p_K^j . Finally, from (24) and (25), we have $m[m_j^{\mathcal{H}^K}]^{\mathcal{X}} \sqsubseteq m[m_{j+1}^{\mathcal{H}^K}]^{\mathcal{X}}$. \square

A useful feature of such $\mathbf{m}^{\mathcal{H}^K}$ is that the pieces of meta-knowledge $m_j^{\mathcal{H}^K}$ are b-separable, and therefore $\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}})$ can be computed efficiently using the results of Section IV-B.

In practice, we will have $\mathbf{p}^1 = \mathbf{1}$ (a vector of 1s), and specifying $\mathbf{p}^2, \dots, \mathbf{p}^M$ can be done in two main ways: by directly giving them, or by specifying their decreasing rate according to, e.g., expert opinions. For example, the statement " ε_ℓ is less reliable than ε_k " implies that p_ℓ should decrease more quickly than p_k .

Example 4. Consider the same three masses $m_i^{\mathcal{X}}$, $i = 1, \dots, 3$ as in Example 3, and the fact that according to some experts, the third source is likely to be far less reliable than the two others. This could be modeled by the successive vectors $\mathbf{p}^1 = (1, 1, 1)$, $\mathbf{p}^2 = (0.8, 0.8, 0.6)$, $\mathbf{p}^3 = (0.6, 0.6, 0.4)$ and so on (until $\mathbf{p}^M = (0, 0, 0)$), to which are associated the meta-knowledges $m_i^{\mathcal{H}^K}$, $i = 1, \dots, 3$.

Table II summarizes the obtained results (showing only the focal elements). We have $\phi_{\mathbf{p}^1}(\mathbf{m}^{\mathcal{X}}) = 0.4$, $\phi_{\mathbf{p}^2}(\mathbf{m}^{\mathcal{X}}) = 0.64$ and $\phi_{\mathbf{p}^3}(\mathbf{m}^{\mathcal{X}}) = 0.88$, hence if we fix $\tau = 0.8$, then \mathbf{p}^3 is selected. Note that the resulting masses are quite different from those obtained in Example 3.

As indicate the next remarks, existing fusion schemes already use this idea.

Remark 5. The sequential discounting approaches proposed in [19]², [34], and [45] and based, respectively, on the degrees of dissent, falsity and disagreement, can all be included in the present approach. For instance, if we associate p_i^j with the discounting rate used at step j in Schubert's sequential discounting [34], then $m[m_j^{\mathcal{H}^K}]^{\mathcal{X}}$ is nothing else but the mass function on \mathcal{X} obtained at step j in Schubert's scheme, showing that it is included in our approach.

Remark 6. Consider K input data o_i in \mathbb{R}^N , $i = 1, \dots, K$ with outputs in space \mathcal{X} . To predict the output of a new data o , the evidential k -nearest neighbor method [5] first orders o_i , $i = 1, \dots, K$ by increasing distance to o . Without loss of generality, let us assume that data are indexed such that

$d(o, o_i) \leq d(o, o_{i+1})$ with d some distance. Let $f(d(o, o'))$ be some discounting function that increases with d (we refer to [5] for details). Then, if we define p_i^j as

$$p_i^j = (1 - f(d(o, o_i))) \times \delta(\{i \leq K - j + 1\})$$

with $\delta(\{i \leq K - j + 1\})$ the indicator function of $\{i \leq K - j + 1\}$ ($= 1$ if $i \leq K - j + 1$, 0 else), $m[m_j^{\mathcal{H}^K}]^{\mathcal{X}}$ is the mass function obtained by the evidential k -nn rule with $k = K - j + 1$. Using Proposition 5, we have that the lower k is, the less specific is the prediction of the output, and therefore the higher its consistency, that is choosing k in the evidential k -nn rule can be seen as choosing a consistency-specificity trade-off.

C. Similarly untruthful sources

In [40], Smets introduces a family of combination rules, the α -conjunctions, depending on a parameter $\alpha \in [0, 1]$. This family represents the set of associative, commutative and linear operators for belief functions with the vacuous mass ($m(\mathcal{X}) = 1$) as neutral element. Let $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$ be two mass functions on \mathcal{X} and let $m_{1 \odot \alpha}^{\mathcal{X}}$ denote the mass function resulting from the α -conjunction of $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$. We have, for all $D \subseteq \mathcal{X}$ [30]:

$$m_{1 \odot \alpha}^{\mathcal{X}}(D) = \sum_{(A \cap B) \cup (\bar{A} \cap \bar{B}) = D} m_1^{\mathcal{X}}(A) m_2^{\mathcal{X}}(B) m_\alpha^{\mathcal{X}}(C), \quad (26)$$

$$\text{where } m_\alpha^{\mathcal{X}}(A) = \alpha^{|\bar{A}|} (1 - \alpha)^{|A|}, \quad \forall A \subseteq \mathcal{X}. \quad (27)$$

Similarly to the assumption "r-out-of-K" that allows one to go from the unnormalized Dempster's rule to the disjunctive rule as a function of r , the α -junctions also allows one to move between two rules based on Boolean operators as α decreases: it includes the unnormalized Dempster's rule (for $\alpha = 1$) and the so-called equivalence rule [29] (for $\alpha = 0$), based on the operator $A \sqcap B = (A \cap B) \cup (\bar{A} \cap \bar{B})$ expressing logical equivalence.

Smets derived these rules in [40] from axiomatic requirements, but admitted that they lacked a clear interpretation for $\alpha \in (0, 1)$. Recently, in [29], such an interpretation was provided in terms of truthfulness of the sources: it was shown that they correspond to assuming that either both³ sources tell the truth or they commit the same lie⁴ with some particular mass depending on α . Let $m_\alpha^{\mathcal{H}^2}$ denote this meta-knowledge on the sources (we refrain from providing the definition of $m_\alpha^{\mathcal{H}^2}$ since it is not needed in this paper; see [29] for the definition). We may then show Proposition 6 concerning $m_\alpha^{\mathcal{H}^2}$, using Lemmas 4 and 5.

Lemma 4. [6, Proposition 2] Disjunctive rule is monotonic with respect to \sqsubseteq .

²Klein and Colot [19] approach amounts basically to an iterative use of Martin et al. approach [26], which is one of the methods (see also, e.g., [24]) to estimate the reliability of a source from its dissimilarity from the other ones.

³The interpretation in [29] was only provided for the case where $K = 2$.

⁴In [29], different forms of lack of truthfulness are considered besides the crudest one, that is, telling the opposite of what one knows.

Lemma 5. Let $m_k^{\mathcal{X}}, m'_k{}^{\mathcal{X}}, k = 1, \dots, K$ be $2K$ mass functions on \mathcal{X} such that $m_k^{\mathcal{X}} \sqsubseteq m'_k{}^{\mathcal{X}}, k = 1, \dots, K$. Let $0 \leq v_k \leq 1$, such that $\sum_{k=1}^K v_k = 1$. We have

$$m_{\Sigma}^{\mathcal{X}} = \sum_{k=1}^K v_k \cdot m_k^{\mathcal{X}} \sqsubseteq \sum_{k=1}^K v_k \cdot m'_k{}^{\mathcal{X}} = m'_{\Sigma}{}^{\mathcal{X}}. \quad (28)$$

Proof. Let $\mathcal{F}_{\Sigma} = \cup_{i=1}^K \mathcal{F}_i$ with \mathcal{F}_i the focal sets of $m_i^{\mathcal{X}}$ denote the set of focal elements of all $m_i^{\mathcal{X}}, i = 1, \dots, K$. Similarly, let us denote $\mathcal{F}'_{\Sigma} = \cup_{i=1}^K \mathcal{F}'_i$ the set of focal elements of all $m'_i{}^{\mathcal{X}}, i = 1, \dots, K$. We also know that for each $k \in \{1, \dots, K\}$, there is a matrix $W^k = [w_{ij}^k]$ satisfying Definition 1. Now, for any $E_i \in \mathcal{F}_{\Sigma}$, we can write

$$m_{\Sigma}^{\mathcal{X}}(E_i) = \sum_{k=1}^K v_k m_k^{\mathcal{X}}(E_i) = \sum_{k=1}^K v_k \sum_{j=1}^q w_{ij}^k = \sum_{j=1}^q \sum_{k=1}^K v_k w_{ij}^k$$

The last equality following by distributing the v_k over $\sum_{j=1}^q w_{ij}^k$. Similarly, for any $F_j \in \mathcal{F}'_{\Sigma}$, we can write

$$m'_{\Sigma}{}^{\mathcal{X}}(F_j) = \sum_{k=1}^K v_k m'_k{}^{\mathcal{X}}(F_j) = \sum_{k=1}^K v_k \sum_{i=1}^p w_{ij}^k = \sum_{i=1}^p \sum_{k=1}^K v_k w_{ij}^k$$

From the above equality, the matrix W^{Σ} with elements w_{ij}^{Σ} satisfies the two first conditions of Definition 1, and we also have $w_{ij}^{\Sigma} > 0$ only if $E_i \subseteq F_j$, since $w_{ij}^k = 0$ when $E_i \not\subseteq F_j$ for any $k \in \{1, \dots, K\}$. This means that W^{Σ} satisfies Definition 1. \square

Proposition 6. $m_{\alpha}^{\mathcal{H}^2} \sqsubseteq_{\mathcal{H}} m_{\alpha'}^{\mathcal{H}^2}$ holds with $1 \geq \alpha \geq \alpha' \geq 0$.

Proof. Equation (26) can be equivalently rewritten:

$$m_{1 \odot \alpha 2}^{\mathcal{X}}(D) = \sum_{A, B} m_1^{\mathcal{X}}(A) m_2^{\mathcal{X}}(B) m_{\alpha}^{\mathcal{X}}(D|A, B), \quad (29)$$

with $m_{\alpha}^{\mathcal{X}}(\cdot|A, B)$ the mass function defined $\forall D \subseteq \mathcal{X}$ as:

$$m_{\alpha}^{\mathcal{X}}(D|A, B) = \sum_{(A \cap B) \cup (\bar{A} \cap \bar{B} \cap C) = D} m_{\alpha}^{\mathcal{X}}(C), \quad (30)$$

where $m_{\alpha}^{\mathcal{X}}$ is the mass function defined by (27).

We may remark that $m_{\alpha}^{\mathcal{X}}(\cdot|A, B)$ is actually the mass function obtained on \mathcal{X} after combining by \odot^{α} the (certain) testimonies $\mathbf{x} \in A$ and $\mathbf{x} \in B$, i.e., we have

$$m_{\alpha}^{\mathcal{X}}(\cdot|A, B) = A \odot^{\alpha} B.$$

From the definition (30) of $m_{\alpha}^{\mathcal{X}}(\cdot|A, B)$, we can furthermore remark that $A \odot^{\alpha} B$ can be obtained by combining $m_{\alpha}^{\mathcal{X}}$ with the information $\mathbf{x} \in \bar{A} \cap \bar{B}$ using the conjunctive rule, and then combining the result of this combination with the information $\mathbf{x} \in A \cap B$ using the disjunctive rule, i.e., we have: $A \odot^{\alpha} B = (m_{\alpha}^{\mathcal{X}} \odot (\bar{A} \cap \bar{B})) \odot (A \cap B)$. Now, as shown by Lemma F.1 of [28], we have

$$m_{\alpha}^{\mathcal{X}} = \bigodot_{x \in \mathcal{X}} m_{\alpha, x}^{\mathcal{X}}, \quad (31)$$

with $m_{\alpha, x}^{\mathcal{X}}$ the mass function defined by $m_{\alpha, x}^{\mathcal{X}}(\{\mathcal{X} \setminus x\}) = \alpha$ and $m_{\alpha, x}^{\mathcal{X}}(\mathcal{X}) = 1 - \alpha$.

Let $1 \geq \alpha \geq \alpha' \geq 0$. We have clearly, $\forall x \in \mathcal{X}$, $m_{\alpha, x}^{\mathcal{X}} \sqsubseteq m_{\alpha', x}^{\mathcal{X}}$. From Lemma 3 and Equation (31), we may

TABLE III
MASS FUNCTIONS RESULTING FROM THE THREE DIFFERENT
HYPOTHESES.

A	$m_1^{\mathcal{X}}$	$m_2^{\mathcal{X}}$	$m[\alpha = 0]^{\mathcal{X}}$	$m[\alpha = 0.5]^{\mathcal{X}}$	$m[\alpha = 1]^{\mathcal{X}}$
\emptyset	0	0	0	0.09	0.18
$\{x_1\}$	0.2	0	0.2	0.17	0.14
$\{x_2\}$	0.4	0	0.22	0.25	0.28
$\{x_1, x_2\}$	0	0.3	0.12	0.12	0.12
$\{x_3\}$	0	0.3	0.12	0.12	0.12
$\{x_1, x_3\}$	0	0	0.06	0.03	0
$\{x_2, x_3\}$	0	0	0.12	0.06	0
\mathcal{X}	0.4	0.4	0.16	0.12	0.16

then conclude that $m_{\alpha}^{\mathcal{X}} \sqsubseteq m_{\alpha'}^{\mathcal{X}}$. Using Lemmas 2 and 4, we find

$$(m_{\alpha}^{\mathcal{X}} \odot (\bar{A} \cap \bar{B})) \odot (A \cap B) \sqsubseteq (m_{\alpha'}^{\mathcal{X}} \odot (\bar{A} \cap \bar{B})) \odot (A \cap B),$$

that is, we have

$$A \odot^{\alpha} B \sqsubseteq A \odot^{\alpha'} B, \quad \forall A, B. \quad (32)$$

Using Lemma 5 together with Equations (29) and (32), we can prove that

$$m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}} \sqsubseteq m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}}, \quad \forall m_1^{\mathcal{X}}, m_2^{\mathcal{X}}, \quad (33)$$

and thus $m_{\alpha}^{\mathcal{H}^2} \sqsubseteq_{\mathcal{H}} m_{\alpha'}^{\mathcal{H}^2}$. \square

Thanks to Proposition 6, one may define a collection $\mathbf{m}^{\mathcal{H}^2}$ based on $m_{\alpha}^{\mathcal{H}^2}$, with α decreasing, that satisfies our proposed source behavior selection approach. An example of such a collection is provided in Example 5.

Example 5. Consider the two mass functions $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$ on $\mathcal{X} = \{x_1, x_2, x_3\}$ in the left part of Table III. Let $\mathbf{m}^{\mathcal{H}^2} = (m_1^{\mathcal{H}^2}, m_{0.5}^{\mathcal{H}^2}, m_0^{\mathcal{H}^2})$ be three meta-knowledge we want to test on these sources, corresponding respectively to $\alpha = 1$, $\alpha = 0.5$ and $\alpha = 0$. The right part of Table III presents the mass functions on \mathcal{X} resulting from the three different assumptions. We have $\phi_{\alpha=0}(\mathbf{m}^{\mathcal{X}}) = 0.62$, $\phi_{\alpha=0.5}(\mathbf{m}^{\mathcal{X}}) = 0.59$ and $\phi_{\alpha=1}(\mathbf{m}^{\mathcal{X}}) = 0.56$. Depending on the chosen value for τ , different meta-knowledge can be selected.

The above results provide practical means to exploit α -conjunctions. This may be regarded as important, as despite the fact that α -conjunctions represent an important theoretical family of combination rules, they are seldom exploited in practice. As shown by our results, these rules can be used to design new strategies to deal with conflicting situations in a principled and meaningful manner. Yet, we must note that the appeal of these rules with respect to conflict resolution is limited, as Proposition 6 which means that for all $m_1^{\mathcal{X}}, m_2^{\mathcal{X}}$,

$$m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}} \sqsubseteq m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}}, \quad 1 \geq \alpha \geq \alpha' \geq 0.$$

does not extend to the general case of $K > 2$ sources, i.e., we do not have in general:

$$\bigodot_{k=1}^K m_k^{\mathcal{X}} \sqsubseteq \bigodot_{k=1}^K m_k^{\mathcal{X}}, \quad 1 \geq \alpha \geq \alpha' \geq 0,$$

as shown by Example 6 below. Hence, it is not possible to use α -conjunction meta-knowledge to more than two sources in

the present framework, unless we relax the assumption of the first step that meta-knowledges $m_j^{\mathcal{H}^K}$ of the collection must form a complete order with respect to $\sqsubseteq_{\mathcal{H}}$. In Section VI-D, we suggest some ways to deal with this case, considering partial rather than complete orders.

Example 6. Consider three mass functions $m_i^{\mathcal{X}}$, $i \in \{1, 2, 3\}$ on $\mathcal{X} = \{x_1, x_2\}$ defined by $m_1^{\mathcal{X}}(\emptyset) = m_2^{\mathcal{X}}(\emptyset) = 1$, and $m_3^{\mathcal{X}}(\{x_1\}) = 1$. Let $1 > \alpha > \alpha' = 0$. We have $(m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}})(\mathcal{X}) = 1$, and thus $((m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}}) \odot^{\alpha'} m_3^{\mathcal{X}}) = m_3^{\mathcal{X}}$. Besides, we have

$$(m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}})(A) = m_{\alpha}^{\mathcal{X}}(A), \quad \forall A \subseteq \mathcal{X}.$$

According to (26), the quantity

$$(m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}})(\{x_1\}) m_3^{\mathcal{X}}(\{x_1\}) m_{\alpha}^{\mathcal{X}}(\mathcal{X}) = m_{\alpha}^{\mathcal{X}}(\{x_1\}) m_{\alpha}^{\mathcal{X}}(\mathcal{X})$$

is transferred to \mathcal{X} . Therefore, we have

$$((m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}}) \odot^{\alpha} m_3^{\mathcal{X}})(\mathcal{X}) \geq m_{\alpha}^{\mathcal{X}}(\{x_1\}) m_{\alpha}^{\mathcal{X}}(\mathcal{X}).$$

Since $m_{\alpha}^{\mathcal{X}}(\{x_1\}) m_{\alpha}^{\mathcal{X}}(\mathcal{X}) > 0$ and

$$((m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}}) \odot^{\alpha'} m_3^{\mathcal{X}})(\mathcal{X}) = m_3^{\mathcal{X}}(\mathcal{X}) = 0,$$

$m_1^{\mathcal{X}} \odot^{\alpha} m_2^{\mathcal{X}} \odot^{\alpha} m_3^{\mathcal{X}}$ cannot be a specialization of $m_1^{\mathcal{X}} \odot^{\alpha'} m_2^{\mathcal{X}} \odot^{\alpha'} m_3^{\mathcal{X}}$.

D. Considering Partially ordered $\mathbf{m}^{\mathcal{H}^K}$

Although general enough to include many existing fusion strategies, as well as to propose new ones, the approach described so far still relies on assumptions that we may want to question or to relax. In this section, we extend the previous proposal by considering that pieces of meta-knowledge that are of interest may be only partially ordered; hence the mass functions resulting from their applications may also be partially ordered with respect to specialization, and the problem of selecting a meta-knowledge and its associated assumption then becomes more complex.

The approach to select the source behavior proposed in Section V requires to choose a particular collection $\mathbf{m}^{\mathcal{H}^K} = (m_1^{\mathcal{H}^K}, \dots, m_M^{\mathcal{H}^K})$ of totally ordered pieces of meta-knowledge. This total order easily allows us getting a unique solution, *i.e.*, the one for which $\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \geq \tau$.

However, in some cases, we may want to define a collection $\mathbf{m}^{\mathcal{H}^K} = (m_1^{\mathcal{H}^K}, \dots, m_M^{\mathcal{H}^K})$ that is only partially ordered with respect to meta-specificity, that is we may have i, j with $m_i^{\mathcal{H}^K} \not\sqsubseteq_{\mathcal{H}} m_j^{\mathcal{H}^K}$ and $m_j^{\mathcal{H}^K} \not\sqsubseteq_{\mathcal{H}} m_i^{\mathcal{H}^K}$. In such a case, we can have $\min\{\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}), \phi_{m_i^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}})\} \geq \tau$ (both assumptions have sufficient consistency) with $m[m_i^{\mathcal{H}^K}]^{\mathcal{X}} \not\sqsubseteq m[m_j^{\mathcal{H}^K}]^{\mathcal{X}}$ and $m[m_j^{\mathcal{H}^K}]^{\mathcal{X}} \not\sqsubseteq m[m_i^{\mathcal{H}^K}]^{\mathcal{X}}$, in which case the consistency-specificity trade-off cannot be used any more.

As an example, consider two sources $m_1^{\mathcal{X}}$ and $m_2^{\mathcal{X}}$ such that $\mathcal{H} = \{R, AR, NR\}$, where R stands for reliable, AR for approximately reliable (see Example 1) and NR for non-reliable. We can then consider the collection of certain pieces of meta-knowledge where each source can be in one state. Such partial order is illustrated on Fig. 1.

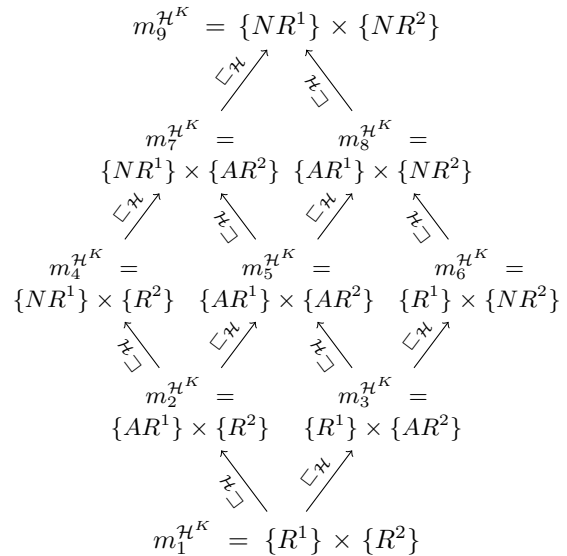


Fig. 1. Illustration of partially ordered collection $\mathbf{m}^{\mathcal{H}^K}$

A possible procedure to deal with such partially ordered collection of hypotheses is the following: first retrieve the set $C \subseteq \mathbf{m}^{\mathcal{H}^K}$ of pieces of meta-knowledge such that

$$C = \left\{ m_i^{\mathcal{H}^K} \mid \phi_{m_i^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \geq \tau \wedge \nexists m_j^{\mathcal{H}^K}, j \neq i \right. \\ \left. \text{s.t. } (\phi_{m_j^{\mathcal{H}^K}}(\mathbf{m}^{\mathcal{X}}) \geq \tau \wedge m[m_j^{\mathcal{H}^K}]^{\mathcal{X}} \sqsubset m[m_i^{\mathcal{H}^K}]^{\mathcal{X}}) \right\}$$

C corresponds to the pieces of meta-knowledge that induce sufficient consistency and that are incomparable with respect to specialization. According to Figure 1, we could have for example the set $C = \{m_7^{\mathcal{H}^K}, m_4^{\mathcal{H}^K}\}$ but not $M = \{m_3^{\mathcal{H}^K}, m_5^{\mathcal{H}^K}\}$ as $m_3^{\mathcal{H}^K} \sqsubseteq_{\mathcal{H}} m_5^{\mathcal{H}^K}$. The next step is to select an element from C . We see at least three ways to do so:

First, the most obvious is to select the element in C with the highest degree of $m^{\mathcal{H}}$ -consistency. As any element in C has a consistency greater than τ while being minimal w.r.t. specificity, such selection makes sense.

Second, if the previous strategy is not applicable because at least two elements have very close degree of $m^{\mathcal{H}}$ -consistency, one may use some information measures [20], [2] refining the specialization partial ordering (see *e.g.*, [1]) to compare the induced specificity of the elements in C .

A last strategy is to consider an additional criteria on top of consistency and specificity, for instance a minimal change principle (see, for example [25], [15], [16]), favoring the piece of meta-knowledge whose fusion result is the closest to the original testimonies. This change may be measured, *e.g.*, by some distances between belief functions [18].

VII. APPLICATION TO CASE STUDY

Our case study concerns the analysis of results issued from BEMUSE international exercise [3]. It consisted in comparing results of uncertainty analysis performed on nuclear computer codes to estimate (among other things)

s_1		s_2		s_3		s_4		s_5		s_6		s_7		s_8		s_9		s_{10}	
E	m	E	m	E	m	E	m	E	m	E	m	E	m	E	m	E	m	E	m
$[x_4, x_5]$.25	$[x_5, x_6]$.5	$[x_4, x_5]$.25	$[x_4, x_5]$.25	$[x_4, x_5]$.25	$[x_4, x_5]$.5	$[x_3, x_5]$.25	$\{x_4\}$.5	$[x_3, x_4]$.25	$[x_5, x_6]$.75
$[x_3, x_5]$.25	$[x_4, x_6]$.25	$[x_3, x_5]$.5	$[x_3, x_5]$.25	$[x_3, x_5]$.25	$[x_3, x_6]$.25	$[x_2, x_5]$.25	$[x_3, x_4]$.25	$[x_3, x_5]$.25	\mathcal{X}	.25
$[x_2, x_6]$.25	\mathcal{X}	.25	\mathcal{X}	.25	\mathcal{X}	.5	$[x_2, x_5]$.25	\mathcal{X}	.25	$[x_2, x_6]$.25	\mathcal{X}	.25	$[x_2, x_5]$.25	\mathcal{X}	.25
\mathcal{X}	.25							\mathcal{X}	.25			\mathcal{X}	.25			\mathcal{X}	.25		

TABLE IV

FOCAL SETS AND ASSOCIATED MASSES OF BEMUSE PARTICIPANTS FOR PCT2. THE DOMAIN $\mathcal{X} = \{x_1, \dots, x_6\}$ IS A UNIFORM DISCRETIZATION OF THE INTERVAL $[592, 1228]$, *i.e.*, $x_i, i = 1, \dots, 6$, CORRESPONDS TO THE INTERVAL $[592 + (i - 1) \cdot \frac{1228 - 592}{6}, 592 + i \cdot \frac{1228 - 592}{6}]$.

pl	x_1	x_2	x_3	x_4	x_5	x_6
$r = 8$	0.001	0.008	0.216	1	1	0.031
$r = 9$	0	0.001	0.0511	0.625	0.813	0.004
$r = 10$	0	0	0.005	0.125	0.188	0

TABLE V

r -OUT-OF- K STRATEGY, WITH $K = 10$ AND $r \in \{10, 9, 8\}$.

peak temperatures during transient conditions of nuclear power plant. Ten participants gave analysis results that were compared to measurements coming from the experiment L2-5 performed on the loss-of-fluid test (LOFT) facility, to test their assessment abilities.

The final results of all these uncertainty analysis are usually difficult to combine and analyze. Moreover, given the lack (and cost) of experimental data together with the complexity of the phenomena involved, there are no reliable means to know the source reliabilities. Also, another aspect of such exercises is that the analysis of the results is as important as the results themselves; it enhances the importance of performing a readable and interpretable merging (if only to be able to explain the results to non-computer scientists).

We focus on the most important of the measured variables, *i.e.*, the second peak clad temperature (PCT2), a critical value of the reactor. The different values are summarized in Table IV, where focal elements are presented as intervals (*i.e.*, $[x_i, x_j]$ with $i < j$ represents the set $\{x_i, x_{i+1}, \dots, x_j\}$, unless it is a singleton). We applied r -out-of- K strategy, with $K = 10$; the results on the contour function (for $r \in \{10, 9, 8\}$) are displayed in Table V. The value $\phi_{H_{10}^{10}} \simeq 0.2$ shows that the sources are globally disagreeing, but the values $\phi_{H_9^{10}} \simeq 0.81$ and $\phi_{H_8^{10}} \simeq 1$ show that assuming 9 out of 10 sources to be reliable ensures an important agreement, and that assuming 8 out of 10 sources to be reliable ensures a totally coherent answer. As a conclusion, our method allows delivering results with the three following desirable properties: (1) They are consistent; (2) They are informative (x_4, x_5 are definitely more plausible); (3) They can be provided through a readable format (for instance, the end-user receiving a message such as ‘‘A fully consistent result can be obtained by assuming that 8 sources out of the 10 available are reliable’’).

VIII. CONCLUSION

When little is known about the sources, that is when one does not have access to observed (training) data or to very accurate expert assessments, then traditional source behavior

estimation approaches cannot be used. We have proposed a practical and sensible method to select a source behavior assumption, and thus incidentally an interpretable rule, in such poorly informed environment.

Our approach relies on a general framework for modeling source behavior [32]. It proposes to pick, from a set of sensible behavior assumptions allowed by this latter framework, the assumption that achieves the best trade-off between specificity and consistency. Of particular interest is the fact that we have extended the notions of inconsistency [8] and specificity to this framework, in order to introduce our approach. It should also be noticed that, up to now, Pichon *et al.* [32] framework has remained mainly theoretical, and this is a first proposal to apply it. We have illustrated our approach by different practical instances, and have provided an illustrative case-study extracted from a real-world problem. In addition, we have related the instances to classical fusion strategies, showing that our framework is quite general and subsumes some existing proposals.

Our method also opens some interesting related questions, which fall outside the scope of the paper.

First, while we allowed behaviors to be dependent, *e.g.*, s_1 is truthful iff source s_2 is, information was assumed to be distinct. It would be interesting to address the case of non-distinctness, as in [8]. However, while formally this can be easily done, studying the interplay of meta-knowledge (in)dependency and of source (non)-distinctness is non-trivial; we leave this interesting topic for further researches.

It would also be interesting to take advantage of extra information, if such information is available. One means is to extend the current framework so that it integrates other approaches such as Smets expert system [41] or Mercier *et al.* [27] contextual discounting. Another issue is how to take account of previous experiences or data concerning the sources. Solving such issues would mean bridging the gap between two extremes: no knowledge on the sources (our approach) and a refined knowledge on the sources issued from learning or experts [27], [14], [9].

Finally, the spirit of our approach, which proposes to lower the inconsistency of fusion results by modifying source behavior assumptions, is quite different from techniques redistributing the resulting mass $m(\emptyset)$ to non-empty focal sets [44], [36]. An interesting future work would be to make a general practical or theoretical comparison of the performances of these different approaches, yet it is not entirely clear how such generic and systematic comparisons (*i.e.*, not

relying on very specific examples) could be done, as the two kinds of techniques rely on different basic assumptions.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous referees for their comments that helped to improve this paper. Special thanks also to Dr. M. Florea for his support for some of the programming.

This work was supported in part by Thales Research and Technology France, in part by the ANR funding ANR-11-IDEX-0004-02 (Labex MS2T, “Investissements d’Avenir” call), in part by the ANR funding ANR-10-INBS-08 (ProFI project, “Infrastructures Nationales en Biologie et Santé”; “Investissements d’Avenir” call), and in part by the Prospector project of the Mastodons 2012 challenge (CNRS).

REFERENCES

- [1] J. Abellán and S. Gomez. Measures of divergence on credal sets. *Fuzzy Sets and Systems*, 157:1514–1531, 2006.
- [2] J. Abellán and S. Moral. Difference of entropies as a non-specificity function on credal sets. *Int. J. of General Systems*, 34:201–214, 2005.
- [3] A. De Crécy, P. Bazin, H. Glaeser, T. Skorek, J. Joucla, P. Probst, K. Fujioka, B.D. Chung, D.Y. Oh, M. Kyncl, et al. Uncertainty and sensitivity analysis of the loft I2-5 test: Results of the bemuse programme. *Nuclear Engin. & Design*, 238(12):3561–3578, 2008.
- [4] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. of Math. Stat.*, 38:325–339, 1967.
- [5] T. Denœux. A k -nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE Trans. on Syst., Man and Cyb.*, 25(5):804–213, 1995.
- [6] T. Denœux. Inner and outer approximation of belief structures using a hierarchical clustering approach. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(4):437–460, 2001.
- [7] S. Destercke, P. Buche, and B. Charnomordic. Evaluating data reliability: An evidential answer with application to a web-enabled data warehouse. *IEEE Trans. on Knowledge and Data Engineering*, 25(1):92–105, 2013.
- [8] S. Destercke and T. Burger. Toward an axiomatic definition of conflict between belief functions. *IEEE Trans. on Syst., Man and Cyb. - B*, 43(2):585–596, 2013.
- [9] S. Destercke and E. Chojnacki. Methods for the evaluation and synthesis of multiple sources of information applied to nuclear computer codes. *Nuclear Engin. and Design*, 238(9):2484–2493, 2008.
- [10] S. Destercke and D. Dubois. Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. *Information Sciences*, 181(18):3925–3945, 2011.
- [11] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *Int. J. of General Systems*, 12(3):193–226, 1986.
- [12] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4(3):244–264, 1988.
- [13] D. Dubois, H. Prade, and Ph. Smets. A definition of subjective possibility. *Int. J. Approx. Reasoning*, 48(2):352–364, 2008.
- [14] Z. Elouedi, E. Lefevre, and D. Mercier. Discountings of a belief function using a confusion matrix. In *Proc. ICTAI*, pages 287–294, 2010.
- [15] J. Grant and A. Hunter. Measuring consistency gain and information loss in stepwise inconsistency resolution. In W. Liu, editor, *Proc. ECSQARU*, volume 6717 of *LNCS*, pages 362–373. Springer Berlin Heidelberg, 2011.
- [16] J. Grant and A. Hunter. Measuring the good and the bad in inconsistent information. In T. Walsh, editor, *Proc. IJCAI*, pages 2632–2637. IJCAI/AAAI, 2011.
- [17] L. Jaulin. Robust set-membership state estimation; application to underwater robotics. *Automatica*, 45(1):202–206, 2009.
- [18] A.-L. Jousselme and P. Maupin. Distances in evidence theory: Comprehensive survey and generalizations. *Int. J. Approx. Reas.*, 53(2):118–145, 2012.
- [19] J. Klein and O. Colot. Automatic discounting rate computation using a dissent criterion. In *Proc. Belief*, pages 1–6, 2010.
- [20] G. J. Klir. *Uncertainty and information: foundations of generalized information theory*. John Wiley & Sons, Inc., 2006.
- [21] E. Lefevre, O. Colot, and P. Vannoorenberghe. Knowledge modeling methods in the framework of evidence theory: an experimental comparison for melanoma detection. In *Proc. SMC*, volume 4, pages 2806–2811, 2000.
- [22] E. Lefevre, P. Vannoorenberghe, and O. Colot. About the use of dempster-shafer theory for color image segmentation. In *Proc. CGIP*, pages 164–169, 2000.
- [23] W. Liu. Analyzing the degree of conflict among belief functions. *Artificial Intelligence*, 170(11):909–924, 2006.
- [24] Z.-G. Liu, J. Dezert, Q. Pan, and G. Mercier. Combination of sources of evidence with different discounting factors based on a new dissimilarity measure. *Decision Support Systems*, 52:133–141, 2011.
- [25] J. Ma, W. Liu, D. Dubois, and H. Prade. Bridging Jeffrey’s rule, AGM revision and Dempster conditioning in the theory of evidence. *Int. J. on Artif. Intell. Tools*, 20(4):691–720, 2011.
- [26] A. Martin, A.-L. Jousselme, and C. Osswald. Conflict measure for the discounting operation on belief functions. In *Proc. Fusion*, pages 1003–1010, 2008.
- [27] D. Mercier, B. Quost, and T. Denœux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9:246–258, 2008.
- [28] F. Pichon. *Belief functions: canonical decompositions and combination rules*. PhD thesis, Université de Technologie de Compiègne, Compiègne, France, 2009.
- [29] F. Pichon. On the α -conjunctions for combining belief functions. In T. Denœux and M.-H. Masson, editors, *Proc. Belief*, volume 164 of *Advances in Intelligent and Soft Computing*, pages 285–292. Springer Berlin / Heidelberg, 2012.
- [30] F. Pichon and T. Denœux. Interpretation and computation of α -junctions for combining belief functions. In *Proc. ISIPTA*, 2009.
- [31] F. Pichon, S. Destercke, and T. Burger. Selecting source behavior in information fusion on the basis of consistency and specificity. In L. C. van der Gaag, editor, *Proc. ECSQARU*, volume 7958 of *LNCS*, pages 473–484. Springer Berlin Heidelberg, 2013.
- [32] F. Pichon, D. Dubois, and T. Denœux. Relevance and truthfulness in information correction and fusion. *Int. J. Approx. Reas.*, 53(2):159–175, 2012.
- [33] F. Pichon, C. Labreuche, B. Duqueroie, and T. Delavallade. Multi-dimensional approach to reliability evaluation of information sources. In P. Capet and T. Delavallade, editors, *Information Evaluation*, pages 129–156. Eds. Hoboken, NJ, USA: Wiley-ISTE, 2014.
- [34] J. Schubert. Conflict management in Dempster-Shafer theory using the degree of falsity. *Int. J. Approx. Reasoning*, 52(3):449–460, 2011.
- [35] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.
- [36] F. Smarandache and J. Dezert. Information fusion based on new proportional conflict redistribution rules. In *Proc. Fusion*, 2005.
- [37] F. Smarandache, A. Martin, and C. Osswald. Contradiction measures and specificity degrees of basic belief assignments. In *Proc. Fusion*, 2011.
- [38] P. Smets. The concept of distinct evidence. In *Proc. IPMU*, pages 789–794, 1992.
- [39] P. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approx. Reas.*, 9(1):1–35, 1993.
- [40] P. Smets. The α -junctions: combination operators applicable to belief functions. In D. M. Gabbay et al., editor, *Proc. ECSQARU-FAPR*, volume 1244 of *LNCS*, pages 131–153. Springer, 1997.
- [41] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4):387–412, 2007.
- [42] P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–243, 1994.
- [43] N. Wilson. Algorithms for Dempster-Shafer theory. *Handbook of defeasible reasoning and uncertainty management systems*, 5:421–475, 2000.
- [44] J.-B. Yang and D.-L. Xu. Evidential reasoning rule for evidence combination. *Artificial Intelligence*, 205:1–29, 2013.
- [45] Y. Yang, D. Han, and C. Han. Discounted combination of unreliable evidence using degree of disagreement. *Int. J. Approx. Reasoning*, 54(8):1197–1216, 2013.