Idempotent conjunctive and disjunctive combination of belief functions by distance minimization

John Klein, Sébastien Destercke, Olivier Colot

Idempotent conjunctive and disjunctive combination of belief functions by distance minimization

John Klein[a], Sebastien Destercke[b] and Olivier Colot[a]
[a]Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France
[b]Technologic University of Compiegne, CNRS, UMR 7253 - Heudiasyc, Centre de Recherche de Royallieu, Compigne, France
{john.klein,olivier.colot}@univ-lille1.fr sebastien.destercke@hds.utc.fr

Abstract

Idempotence is a desirable property when cautiousness is wanted in an information fusion process, since in this case combining identical information should not lead to the reinforcement of some hypothesis. Idempotent operators also guarantee that identical information items are not counted twice in the fusion process, a very important property in decentralized applications where the information origin cannot always be tracked (ad-hoc wireless networks are typical examples). In the theory of belief functions, a sound way to combine conjunctively multiple information items is to design a combination rule that selects the least informative element among a subset of belief functions more informative than each of the combined ones. In contrast, disjunctive rules can be retrieved by selecting the most informative element among a subset of belief functions less informative than each of the combined ones. One interest of such approaches is that they provide idempotent rules by construction.

The notions of less and more informative are often formalized through partial orderings extending usual set-inclusion, yet the only two informative partial orders that provide a straightforward idempotent rule leading to a unique result are those based on the conjunctive and disjunctive weight functions. In this article, we show that other partial orders can achieve a similar goal when the problem is slightly relaxed into a distance optimization one. Building upon previous work, this paper investigates the use of distances compatible with informative partial orders to determine a unique solution to the combination problem. The obtained operators are conjunctive/disjunctive, idempotent and commutative, but lack associativity. They are, however, quasi-associative allowing sequential combinations at no extra complexity. Some experiments demonstrate interesting discrepancies as compared to existing approaches, notably with the aforementioned rules relying on weight functions.

Keywords: Belief functions, combination, distance, idempotence, partial order, convex optimization

1 Introduction

The theory of belief functions is a framework for reasoning under uncertainty. It was initially proposed to model imprecise statistical observation [6], and this initial work was then extended [22] to include subjective or epistemic uncertainty (*e.g.*, when a variable has a fixed, yet ill-known value). It has received a considerable attention in the soft computing community as it allows the combination of uncertain, imprecise or conflictual pieces of evidence. The flexibility of the theory of belief functions has led some people to think of it as a data fusion framework whereas its initial purpose is more general.

Combining pieces of evidence coming from different sources of information is one of the most frequently studied problem in the belief function theory. In particular, a rich literature exists (see for example [27, 11] and references therein) proposing alternatives to Dempster's rule when this latter does not apply, that is when sources of information are either unreliable or non-independent, or both. This paper deals with the second issue, that is the one concerning source independence, and more particularly with the case where this dependence is ill-known and hard to assess.

Under such an assumption, it is common to adopt a cautious approach, also known as least-commitment principle [14] (LCP). A natural consequence of this principle is that if all the sources provide the same mass function, then the result of the combination should be this very mass function, or in other words the combination should be idempotent. However, if idempotence is a consequence of the LCP, satisfying idempotence does not imply satisfying the LCP. As shown by Dubois and Yager [16], there is virtually an infinity of ways to derive idempotent combination rules, not all of them necessarily following a least-commitment principle. For instance, Cattaneo [3] provides an idempotent rule following a conflict-minimization approach, which may lead to non-least committed results [9].

So, to satisfy the LCP, we must add additional constraints on the combination rule. One such natural constraint is to consider a partial order over informative content of mass functions, and to require the combination result to be one of the maximal element of this partial order within the subset of possible combination results. Unfortunately, such an approach can present two shortcomings: it will very often lead to multiple solutions corresponding to all possible maximal elements [12], and estimating this set of solutions may be computationally challenging. Denœux [7] shows that using the canonical decomposition and the associated partial order leads to a unique LCP, idempotent solution, yet this solution has two limitations: the set of possible combination results is quite small, leading to a not so conservative behavior (as we will see on a simple example in Section 5, and as already pointed out in [9]), and the combination only apply to specific (i.e., non-dogmatic) mass functions.

In this paper, we take inspiration from some of our previous work [20] studying the consistency of distances with partial orders comparing informative contents to propose a new way to derive cautious combination rules. Our approach departs from previous ones, as it is formulated as an optimization problem (similarly to what is done by Cattaneo [3] for conflict minimization) that naturally satisfies the LCP principle. Our approach makes minimal assumptions about the shape of the combination result, in the sense that the only constraints it imposes on the combination result is to be more informative than each initial belief function in the conjunctive case, and less informative in the disjunctive case. This also contrasts with previous approaches [7, 4, 9], that considered the results to take specific forms (either in the form of a joint mass function with prescribed marginals [4, 9], or in a weight function combined through uninorms [7]). It is in fact in-line with the generic conjunctive operator described by Dubois *et al.* [11].

Our approach also solves the two problems of solution uniqueness and computability, since if the distance is chosen so as to minimize a strictly convex objective function, we are guaranteed to have a unique solution satisfying the LCP and computable by convex optimization. Section 2 recalls the basics needed in this paper. The bulk of the proposal is contained in Section 3, where we present the combination approach and study its properties in the conjunctive case. In section 4, we present equivalent results for the disjunctive case. Section 5 compares our proposal with respect to existing ones.

This paper is an extended version of [19] which was presented at the 4$^{\text{th}}$ international conference on belief functions, BELIEF'16 to which this special issue of IJAR is dedicated.

2 Preliminaries and problem statement

This section briefly sketches the basics of evidence theory and provides references for readers interested in further details. Like most of the belief function literature, this paper is limited to belief functions on finite spaces. The derivation of the results introduced in this paper in the continuous case is left for future work.

2.1 Basic concepts

A body of evidence $\mathcal{E}_i$ defined on the finite space $\Omega = \{\omega_1, \ldots, \omega_n\}$ will be modeled by a mass function $m_i : 2^\Omega \to [0,1]$ that sums up to one, *i.e.*, $\sum_{E \subseteq \Omega} m_i(E) = 1$. Following usual notation, 2^Ω denotes the power set of Ω. In evidence theory, this basic tool models our uncertainty about the true value of some quantity (parameter, variable) lying in Ω. The cardinality of 2^Ω is denoted by $N = 2^n$. The set $\mathcal{M}$ of mass functions on Ω is called **mass space**. A set A is a **focal element** of m iff $m(A) > 0$. The **complement** $\overline{m}$ of a mass function m is such that $\forall A \subseteq \Omega$, we have $\overline{m}(A) = m(A^c)$ where A^c denotes the complement of the set A in Ω.

A mass function assigning a unit mass to a single focal element A is called **categorical** and denoted by m_A: $m_A(A) = 1$. If $A \neq \Omega$, the mass function m_A is equivalent to providing the set A as information, while the **vacuous** mass function m_Ω represents ignorance. A function m_i such that $m_i = (1-\alpha)m_A + \alpha m_\Omega$ with $\alpha \in [0;1]$ is called a **simple** mass function and is regarded as an elementary evidence supporting the event A.

Besides, a mass function m_i such that $m_i(\Omega) = 0$, *i.e.* Ω is not a focal element of m_i, is called a **dogmatic** mass function. A mass function m_i such that $m_i(\emptyset) = 0$, *i.e.* $\emptyset$ is not a focal element of m_i, is called a **normalized** mass function while a function such that $m_i(\emptyset) > 0$ is called **subnormal**. Some authors [6, 22] consider that $\emptyset$ cannot be assigned a positive mass (closed world assumption) because it arises from inconsistencies and should be re-assigned to the remaining valid hypotheses. In contrast, other authors [23, 28] prefer to keep track of the mass assigned to $\emptyset$ (open world assumption) and consider that it is a valuable feature to assess the quality level of information encoded in a given mass function. Under the open world assumption, the mass assigned to $\emptyset$ can also be interpreted as the support given to the fact that the true value of the variable of interest does not belong to Ω, *i.e.* Ω is not exhaustive. Under the closed world assumption, Ω is assumed to be the exhaustive set of solutions and a positive mass for $\emptyset$ is induced by other factors such as errors in the model or unreliability of some pieces of evidence.

If the reliability of the evidence encoded in a mass function can be evaluated through a coefficient $\alpha \in [0,1]$, then a so-called **discounting** operation on m can be performed. A discounted mass function is denoted by m^α and we have :

$$m^\alpha = (1-\alpha)m + \alpha m_\Omega. \tag{1}$$

α is called the **discounting rate**. Consequently, setting $\alpha = 1$ turns a mass function into the vacuous mass function (as the source is totally unreliable), while $\alpha = 0$ leaves it untouched. Note that a discounted categorical mass function is a simple mass function. Several alternative set functions are commonly used in the theory of belief functions and encode the same information as a given mass function m_i. The **belief**, **plausibility** and **commonality** functions of a set A are defined as

$$bel_i(A) = \sum_{E \subseteq A, E \neq \emptyset} m_i(E), \tag{2}$$

$$pl_i(A) = \sum_{E \cap A \neq \emptyset} m_i(E), \tag{3}$$

$$q_i(A) = \sum_{E \supseteq A} m_i(E) \tag{4}$$

and respectively represent how much the event A is implied by, consistent with, and considered common by the actual evidence. Under the open world assumption, another representation is provided by the implicability function b_i. This function is closely related to the belief and plausibility functions through the following relations: $\forall A \in 2^\Omega$,

$$b_i(A) = bel_i(A) + m_i(\emptyset), \tag{5}$$

$$b_i(A) = 1 - pl_i\left(A^c\right). \tag{6}$$

In this paper, we will also use the **conjunctive weight function** denoted by w_i and introduced by Smets [25]. It is only defined for non-dogmatic mass functions ($m(\Omega) > 0$). This representation has its roots in a decomposition of a mass function into simple ones which are standing for elementary pieces of evidence supporting each event $A \subsetneq \Omega$ individually. The simplest transition relation allowing to compute a conjunctive weight function is obtained from the commonality function as follows:

$$w_i\left(A\right) = \prod_{E \supseteq A} q_i\left(E\right)^{(-1)^{|E|-|A|+1}}, \forall A \subseteq \Omega. \tag{7}$$

The non-dogmatic condition prevents division by zero to happen in Equation (7). In practice, when a dogmatic mass function m_i has to be turned into a conjunctive weight function, one can discount it with a very small discount rate as compared to the minimum positive mass of m_i. The discount rate can be chosen as small as necessary so that the values of w_i stabilize to some value up to a desired precision threshold.

Unlike other representations, the codomain of these functions is $(0; +\infty)$ and not $[0, 1]$. Having $w_i(A) < 1$ is understood as the fact that some evidence has been collected allowing to support A being true. Having $w_i(A) > 1$ means that A is unlikely to the point that a significant amount of evidence needs to be collected before starting to support A being true. Finally, $w_i(A) = 1$ stands for a neutral opinion regarding event A. We refer to Denœux [7] for more details on the conjunctive weight function.

When mass function are unnormalized ($m(\emptyset) > 0$), a dual decomposition can be obtained using the **disjunctive weight function** denoted by v_i. It can be computed for instance from the implicability function as follows:

$$v_i(A) = \prod_{E \subseteq A} b_i(E)^{(-1)^{|A|-|E|+1}}, \forall A \subseteq \Omega. \quad (8)$$

In the same way as conjunctive weight functions, one turns a normalized mass function m_i into a disjunctive weight function by artificially assigning an infinitesimal mass value to $\emptyset$ and then renormalize so that $\sum_{E \subseteq \Omega} m(E) = 1$. Such a constraint may be perceived as less natural than $m(\Omega) > 0$, in particular under a closed-world assumption.

2.2 Mass function combination

In this subsection, we give a very brief presentation of how pieces of information are usually combined when they are represented by belief functions. Suppose two sources of information S_1 and S_2 have gathered pieces of evidence allowing them to define two mass functions m_1 and m_2 respectively. Let us further suppose that S_2 collected some certain but imprecise piece of evidence, *i.e.* $m_2 = m_B$ with $B \subsetneq \Omega$ and $|B| \geq 1$. Under such circumstances, a natural solution is to re-allocate the mass $m_1(A)$ to the set $A \cap B$. This combination is denoted by $m_{1|B}$ and is given by the following formula:

$$m_{1|B}(X) = \sum_{A \cap B = X} m_1(A). \quad (9)$$

This combination is called **conditioning** because when m_1 and m_2 are probability distributions, probabilistic conditioning is retrieved. The mass function $m_{1|B}$ can be understood as m_1 given that B is true. Note however that this corresponds to a revision step, which fits well with an information fusion setting, but that other conditioning rules fitted to other settings (e.g., focusing operation) could be devised. The interested reader can check [13].

Evidential combination rules address the same combination problem in a more general context, *i.e.* when m_2 is not a categorical mass function. The straightforward generalization of conditioning gives the **conjunctive rule** [23]. This rule is denoted by $\textcircled{\cap}$ and the result of the combination of m_1 with m_2 is denoted by $m_{1 \textcircled{\cap} 2} = m_1 \textcircled{\cap} m_2$. We have

$$m_{1 \textcircled{\cap} 2}(X) = \sum_{A \cap B = X} m_1(A) m_2(B). \tag{10}$$

The most widely used combination rule is Dempster's rule [6], denoted by $\oplus$ and obtained by normalizing the output of the conjunctive rule:

$$m_{1 \oplus 2}(X) = \begin{cases} \frac{m_{1 \textcircled{\cap} 2}(X)}{1 - m_{1 \textcircled{\cap} 2}(\emptyset)} \text{ if } X \neq \emptyset \\ 0 \text{ if } X = \emptyset \end{cases}. \tag{11}$$

Both the conjunctive and Dempster's rules transfer masses to intersections of focal sets, which is justified when the sources are reliable and consequently do not support incompatible events. In practice, this is not always true and one alternative is to transfer masses to unions of focal sets instead. The disjunctive rule [10] $\textcircled{\cup}$ relies on this idea. We have

$$m_{1 \textcircled{\cup} 2}(X) = \sum_{A \cup B = X} m_1(A) m_2(B). \tag{12}$$

Another important assumption that is a prerequisite to each of these rules is source independence. In [6], Dempster introduces belief functions in connection with probability theory. More precisely, Dempster derives a normalized belief function for each source S_i by coupling a probability space $(\mathcal{X}_i, \sigma_{\mathcal{X}_i}, \mu_i)$ with a multi-valued mapping Γ_i from $\mathcal{X}_i$ to Ω (which is actually a mapping $\mathcal{X}_i \longrightarrow 2^{\Omega}$). In this setting, independence between S_1 and S_2 is understood in the usual probabilistic way and the combined belief function is induced by the product measure $\mu_1 \otimes \mu_2$.

It is not easy to describe what practical situations fall in the independent source assumption. As an example, Dempster mentions that two belief functions derived from non overlapping statistical samples are obviously independent. Let alone this textbook case, what means independence (or distinctness [24]) is far less obvious. Yet, assuming independence when sources are likely to be correlated in some ways may lead to reinforcing some unwarranted assumptions (see [17] for example).

When independence cannot be reasonably assumed, a sound way to circumvent this difficulty is to require the combination rule to be idempotent. Indeed, idempotence will automatically prevent evidence pieces from being counted twice. Denoeux [7] introduced a conjunctive idempotent rule, known as the cautious rule $\textcircled{\wedge}$, as well as a disjunctive idempotent rule, known as the bold rule $\textcircled{\vee}$. For a pair of non-dogmatic mass functions (m_1, m_2), their cautious combination is defined as

$$m_{1 \textcircled{\wedge} 2} = \textcircled{\cap}_{A \subsetneq \Omega} m_A^{w_1(A) \wedge w_2(A)}, \tag{13}$$

where $\wedge$ is the minimum operator. For a pair of non-normalized mass functions (m_1, m_2), their bold combination is defined as

$$m_{1 \textcircled{\vee} 2} = \textcircled{\cup}_{A \neq \emptyset} \overline{m_{A^c}^{v_1(A) \wedge v_2(A)}}. \tag{14}$$

Finally, except for Dempster's rule, the normalized version of a given rule $\odot$ will be denoted by $\odot^*$.

2.3 Comparing mass functions with respect to informative content

When considering two mass functions m_1 and m_2 providing information about the same quantity, a natural question is to wonder if one of these two is more informative than the other. This question can be answered if the mass space $\mathcal{M}$ is endowed with a relevant partial order $\sqsubseteq$ with $m_1 \sqsubseteq m_2$ when m_1 is more informative than m_2.

Determining if a function m_1 is more informative than m_2 is not a trivial task except in specific circumstances. For instance, informative content related partial orders should extend set inclusion, since when $A \subseteq B$, A is more informative than B. Such partial orders are obtained by stating that m_1 is **f-included** in m_2, denoted $m_1 \sqsubseteq_f m_2$, if $f_1 \leq f_2$ where $\leq$ is the element-wise inequality, meaning that

$$f_1 \leq f_2 \Leftrightarrow f_1(A) \leq f_2(A), \forall A \subseteq \Omega,$$

for some $f \in \{pl, q, w\}$ denoting a given type of set functions. We consider that the partial order $\sqsubseteq_w$ is also valid for dogmatic mass functions using the infinitesimal discounting approximation as described in Section 2.1.

Besides, one could think of defining a partial order using implicability functions but from equation (6) this partial order is formally equivalent to $\sqsubseteq_{pl}$. It is also possible to use belief functions to derive the partial order $\sqsubseteq_{bel}$. Unlike the other partial orders, $\sqsubseteq_{bel}$ is defined as

$$m_1 \sqsubseteq_{bel} m_2 \Leftrightarrow bel_1(A) \geq bel_2(A), \forall A \subseteq \Omega.$$

Although $\sqsubseteq_{bel}$ is a valid partial order from a mathematical point of view, it turns out to have singular properties. For instance, for any subset $A \subsetneq \Omega$ such that $A \neq \emptyset$, we have

$$m_A \sqsubset_{bel} m_\Omega \sqsubset_{bel} m_\emptyset, \tag{15}$$

while for any $f \in \{pl, q, w\}$, we have

$$m_\emptyset \sqsubset_f m_A \sqsubset_f m_\Omega. \tag{16}$$

More precisely, $m_\emptyset$ and m_Ω are respectively the unique minimal and maximal elements of $(\mathcal{M}, \sqsubseteq_f)$, $f \in \{pl, q\}$. The function $m_\emptyset$ is the minimum of $(\mathcal{M}, \sqsubseteq_w)$ but this poset has several maximal elements. Concerning $(\mathcal{M}, \sqsubseteq_{bel})$ the maximal element is $m_\emptyset$ and any function $m_{\{x\}}$ is a minimal element (for any $x \in \Omega$). In particular, the non-uniqueness of minimal elements will be problematic for the approach that will be introduced in the next section. Since $\sqsubseteq_{bel}$ is very seldom used in the literature, we will not include it in our study but only comment on it in relevant parts of this article.

Similarly to the belief function, the disjuntive weight function induces a partial order with reversed inequalites:

$$m_1 \sqsubseteq_v m_2 \Leftrightarrow v_1(A) \geq v_2(A), \forall A \subseteq \Omega.$$

This latter will be mentioned only in the derivation of disjunctive operators (section 4).

Another widely used partial order relies on the concept of specialization. A function m_1 is a **specialization** of m_2, denoted $m_1 \sqsubseteq_s m_2$, if there is a non-negative $N \times N$ matrix $\mathbf{S} = [S(k,j)]$ such that

$$\text{for } j = 1, \ldots, N, \quad \sum_{k=1}^{N} S(k,j) = 1,$$
$$S(k,j) > 0 \Rightarrow E_k \subseteq E'_j,$$
$$\text{for } k = 1, \ldots, N, \quad \sum_{j=1}^{N} m_2(E'_j) S(k,j) = m_1(E_k).$$

The term $S(k,j) > 0$ is the proportion of the focal set E'_j that "flows down" to focal set E_k. The order in which subsets are indexed is arbitrary.

A subclass of specialization matrices are Dempsterian specialization matrices. A matrix $\mathbf{D}_i$ is a Dempsterian specialization matrix if it is a specialization matrix and if for any $E_k \subseteq E'_j$, one has $D_i(k,j) = m_{i|E'_j}(E_k)$ for some mass function m_i. Now, one writes $m_1 \sqsubseteq_d m_2$ if m_1 is a specialization of m_2 relying on a Dempsterian matrix $\mathbf{D}_0$ which actually means that $m_1 = m_0 \circledcirc m_2$.

Each of these orders is partial in the sense that in general there are some incomparable pairs (m_1, m_2), *i.e.* $m_1 \not\sqsubseteq m_2$ and $m_2 \not\sqsubseteq m_1$. There are implications between them, as we have

$$\left.\begin{array}{c} m_1 \sqsubseteq_w m_2 \\ m_1 \sqsubseteq_v m_2 \end{array}\right\} \Rightarrow m_1 \sqsubseteq_d m_2 \Rightarrow m_1 \sqsubseteq_s m_2 \Rightarrow \left\{\begin{array}{c} m_1 \sqsubseteq_{pl} m_2 \\ m_1 \sqsubseteq_q m_2 \end{array}\right. . \tag{17}$$

2.4 Evidential distances and their compatibility with partial orders

Another way to compare mass functions is by measuring how distant they are. An **evidential distance** is a function $d : \mathcal{M} \times \mathcal{M} \to [0, \infty]$ that satisfies the symmetry, definiteness and triangle inequality properties. In [20], we have formalized the idea of compatibility between a distance and a partial order in the following way:

Definition 1. Given a partial order $\sqsubseteq_f$ defined on $\mathcal{M}$, an evidential distance d is said to be **$\sqsubset_f$-compatible (in the strict sense)** if for any mass functions m_1, m_2 and m_3 such that $m_1 \sqsubset_f m_2 \sqsubset_f m_3$, we have:

$$\max\{d(m_1, m_2); d(m_2, m_3)\} < d(m_1, m_3), \tag{18}$$

For some family of set-functions f that are in bijective correspondence with mass functions, an interesting distance $d_{f,k}$ is defined as

$$d_{f,k}(m_1, m_2) = \|f_1 - f_2\|_k = \left(\sum_{A \subseteq \Omega} |f_1(A) - f_2(A)|^k \right)^{\frac{1}{k}}. \tag{19}$$

In particular, we showed that for any $k \in \mathbb{N}^* \setminus \{\infty\}$, $d_{pl,k}$ is $\sqsubset_{pl}$-compatible and $d_{q,k}$ is $\sqsubset_q$-compatible (in the strict sense for all of them).

3 A distance-based cautious conjunctive aggregation

In Section 2.2 as well as in the belief function literature, all combination rules use some particular features of belief functions (masses m_i or weights w_i) to derive conjunctive rules. In [11], Dubois *et al.* consider the fusion problem from a more abstract point of view, and merely require conjunctive and disjunctive rules to satisfy the following principle: given items of information $\mathcal{I}_1, \ldots, \mathcal{I}_\ell$ and an information ordering $\sqsubseteq$ relation defined on them, a rule is conjunctive if its result $\mathcal{I}_\cap$ is such that

$$\mathcal{I}_\cap \sqsubseteq \mathcal{I}_i, \ \forall i \in \{1, \ldots, \ell\}$$

and is disjunctive if its result $\mathcal{I}_\cup$ is such that

$$\mathcal{I}_\cup \sqsupseteq \mathcal{I}_i, \ \forall i \in \{1, \ldots, \ell\}.$$

They then recommend (in absence of other information) to follow the LCP principle in the conjunctive case, and the "most committed principle" in the disjunctive case to pick a combination result. This view has the advantage that it makes no a priori assumption about the shape of the rule, nor about the dependence assumption it should satisfy. This section and the next introduce how we decline this view with belief functions to obtain computable idempotent operators, starting with the conjunctive operator in Section 3 and pursuing with the disjunctive one in Section 4.

3.1 Conjunctive combination using partial orders

Rather than seeing a conjunctive combination of $\mathcal{E}_1, \ldots, \mathcal{E}_\ell$ as a particular operator defined either on the mass functions $m_1, \ldots, m_\ell$ or on the weight functions $w_1, \ldots, w_\ell$, we simply consider that a mass function m^* resulting from a conjunction should be (i) more informative (in the sense of some partial order $\sqsubseteq_f$) than any $m_1, \ldots, m_\ell$ and (ii) should be among the least committed elements (in terms of information) among those, in accordance with the LCP. Formally speaking, if we denote by

$$\mathcal{S}_f\left(m_i\right) := \{m \in \mathcal{M} \mid m \sqsubseteq_f m_i\} \tag{20}$$

the set of all mass functions more informative than m_i, then m^* should be such that:

(i) $m^* \in \mathcal{S}_f\left(m_1\right) \cap \ldots \cap \mathcal{S}_f\left(m_\ell\right)$,

(ii) $\nexists m' \in \mathcal{S}_f\left(m_1\right) \cap \ldots \cap \mathcal{S}_f\left(m_\ell\right)$ such that $m^* \sqsubset_f m'$.

The first constraint expresses the conjunctive behavior of such an approach. The second constraint says that m^* is a maximal element (*i.e.* a least committed solution) for admissible solutions subject to the first constraint.

While this solution is generic and does not require any explicit model of dependence, it should be noted that the choice of the partial order to consider is not without consequence. Considering the orders mentioned in Section 2.3 and their relation, Equation (17) tells us that for a mass function m, $\mathcal{S}_w(m) \subseteq \mathcal{S}_d(m) \subseteq \mathcal{S}_s(m) \subseteq \mathcal{S}_{pl}(m)$, hence the space of solutions will be potentially much smaller when choosing $\sqsubseteq_w$ rather than $\sqsubseteq_{pl}$. In practice and in accordance with the LCP, it seems safer to choose the most conservative partial orders, *i.e.* $\sqsubseteq_{pl}$ or $\sqsubseteq_q$ in our case. We will see in Section 5 that it can have some impact on the combination results, even for simple examples.

While our definition of the cautious result of a conjunctive combination appears natural, it still faces the problem that many different solutions m^* could actually fit the two constraints, as $\sqsubseteq$ is a partial order. This means that to identify a unique solution, we need an additional criterion, that preferably leads to efficient computations. One idea to solve this problem that we explore here is to use distances that are compatible with $\sqsubseteq$.

3.2 New conjunctive operators from soft LCP

To derive new conjunctive operators, we consider a weaker form of least commitment principle which we call **soft LCP**. This principle states that when there are several candidate mass functions compliant with a set of constraints, the one with minimal distance value from the vacuous mass function should be chosen for some $\sqsubseteq$-compatible distance. We call this version of the LCP *soft* because the philosophy behind LCP is to guide us to the most uninformative solution, which may be ill-defined as requiring this solution to be unique is too strong. Our idea is therefore to soften the initial LCP requirements by adding constraints ensuring uniqueness.

The resulting conjunctive operator, denoted $\sqcap_{f,k}$, depends on the chosen distance $d_{f,k}$, and is defined as follows

Definition 2. for any set of ℓ functions $\{m_1, .., m_\ell\}$, we have

$$m_1 \sqcap_{f,k} .. \sqcap_{f,k} m_l = \underset{m \in \mathcal{S}_f(m_1) \cap .. \cap \mathcal{S}_f(m_\ell)}{\arg\min} d_{f,k}\left(m, m_\Omega\right). \tag{21}$$

According to [20, corollary 4], we know that the problem induced by the soft LCP is a convex optimization problem with a unique solution if the chosen distance $d_{f,k}$ is $\sqsubset_f$-compatible and if $2 \leq k < \infty$. Considering results in [20], the operator $\sqcap_{f,k}$ can be applied for $f \in \{pl, q\}$. To our knowledge, there is no evidential distance reported to be $\sqsubseteq_w$, $\sqsubseteq_d$ or $\sqsubseteq_s$-compatible in the literature, hence that $\sqcap_{f,k}$ can be easily applied for $f \in \{w, d, s\}$ is not guaranteed. One could probably derive a $\sqsubseteq_w$-compatible distance by computing $\|w_1 - w_2\|_k$. However, the minimization problem solution is already known from the cautious rule, *i.e.* $\owedge = \sqcap_{w,k}$, and is one of the rare exception[1] where the raw application of LCP leads to a unique solution. The absence (for the time being) of a practical mean to compute the solution of problem (21) when $f \in \{d, s\}$ is not a major drawback, as those partial orders limit the space of solutions by inducing more restrictive spaces $\mathcal{S}_f(m_i)$. For the sake of the paper readability, the $f = d$ or s cases are not further discussed in the main body of this paper but only commented in A.

Concerning $\sqcap_{bel,k}$, there is no theoretical impediment but a practical one. Indeed, one may have $\mathcal{S}_{bel}(m_1) \cap \mathcal{S}_{bel}(m_2) = \emptyset$. When $f \in \{pl, q\}$, we know that such intersections are not empty because they always contain $m_\emptyset$.

Operators $\sqcap_{q,k}$ and $\sqcap_{pl,k}$ can be easily implemented using standard solvers available in scientific programming libraries because they amount to a convex

[1] The inverse pignistic [15] being another one.

minimization problem. We give some details on how to implement those operators when $k = 2$ in B.

Definition 2 could have been formulated in a more general way by replacing distance $d_{f,k}$ by any distance such that the minimization problem has a unique solution. We chose to focus on L_k norm based distances as results are available to prove the existence of a unique solution for them. The most popular distance in evidence theory is Jousselme distance [18] which is not consistent with any of the partial order evoked in this paper, at least when we can have $m(\emptyset) \neq 0$. Note that [20, corollary 4] is a sufficient condition to obtain a unique solution therefore one cannot conclude about the ability of Jousselme distance to induce a bona fide operator in the same way as definition 2.

Finally, we would like also to stress that a soft LCP solution is an LCP solution to the problem presented in the previous subsection as long as m_Ω is the maximum of $(\mathcal{M}, \sqsubseteq_f)$. Indeed condition (i) is verified by construction. Concerning condition (ii), suppose $\exists m' \in \mathcal{S}_f(m_1) \cap \ldots \cap \mathcal{S}_f(m_\ell)$ such that $m^* \sqsubset_f m'$. As d is consistent with $\sqsubset_f$, $m^* \sqsubset_f m' \sqsubset m_\Omega$ implies $d(m', m_\Omega) < d(m^*, m_\Omega)$ which is in contradiction with the very definition of m^*, hence condition (ii) is verified as well. There is however little to no debate about the fact that m_Ω is the least informative mass function.

Just for a quick illustration, we provide the following example which is a continuation of [7, example 2].

Example 1. Let $\Omega = \{a, b, c\}$. Here are two non-dogmatic mass functions along with their combinations under our new operators and other standard approaches.

subset	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	Ω
m_1	0	0	0	0.3	0	0	0.5	0.2
m_2	0	0	0.3	0	0	0	0.4	0.3
$m_1 \sqcap_{q,2} m_2$	0	0	0.2	0.1	0	0	0.5	0.2
$m_1 \sqcap_{pl,2} m_2$	0	0	0.3	0	0	0	0.4	0.3
$m_1 \circledcirc m_2$	0	0	0.6	0.12	0	0	0.2	0.08
$m_1 \owedge m_2$	0	0	0.42	0.09	0	0	0.43	0.06

We see that the conjunctive and cautious rules transfer much more mass to $\{b\}$ than operators $\sqcap_{q,2}$ and $\sqcap_{pl,2}$ do. Also, observe that $m_1 \sqcap_{pl,2} m_2 = m_2$ because $m_2 \sqsubset_{pl} m_1$.

3.3 Properties of new conjunctive operators

The commutativity of the set-intersection and the symmetry property of distance give that $\sqcap_{f,k}$ is commutative. Each operator $\sqcap_{f,k}$ is also idempotent: for any possible solution $m \in \mathcal{S}_f(m_1) \setminus \{m_1\}$, we have $d_{f,k}(m_1, m_\Omega) < d_{f,k}(m, m_\Omega)$ because $d_{f,k}$ is $\sqsubseteq_f$-compatible and $m \sqsubseteq_f m_1 \sqsubseteq_f m_\Omega$, hence $m_1 \sqcap_{f,k} m_1 = m_1$.

Each of these operators are also conjunctive by construction, in the sense that the output mass function is more informative than any of the initial mass functions. Indeed if m_i states that ω is not a possible value of the unknown quantity ($pl_i(\{\omega\}) = 0$), then any function in $\mathcal{S}(m_i)$ also states so. Since the combination result belongs to $\mathcal{S}(m_i)$, then this piece of information is propagated by $\sqcap_{f,k}$.

Except for the $f = w$ case[2], these operators are, however, not associative because we can have

$$\mathcal{S}_f\left(m_1 \sqcap_{f,k} m_2\right) \subsetneq \mathcal{S}_f\left(m_1\right) \cap \mathcal{S}_f\left(m_2\right).$$

The above remark is illustrated by the following example in the $f = q$ case.

Example 2. Let $\Omega = \{a, b, c\}$ denote some space. Let us introduce the following mass functions on Ω:

subset	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	Ω
m_1	0	0.1	0	0	0	0	0.1	0.8
m_2	0	0	0	0.1	0.1	0	0	0.8
$m_1 \sqcap_{q,2} m_2$	0	$1/15$	$1/15$	0	$1/15$	0	0	0.8
q_1	1	0.9	0.9	0.8	0.9	0.8	0.9	0.8
q_2	1	0.9	0.9	0.9	0.9	0.8	0.8	0.8
$q_1 \wedge q_2$	1	0.9	0.9	0.8	0.9	0.8	0.8	0.8
q_{12}	1	$1/15$	$1/15$	0.8	$1/15$	0.8	0.8	0.8

where q_{12} denote the commonality function in correspondence with $m_1 \sqcap_{q,2} m_2$. Let m_3 denote the following mass function

subset	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	Ω
m_3	0	0.1	0	0	0.3	0	0.2	0.4
q_3	1	0.5	0.6	0.4	0.9	0.4	0.6	0.4

We have $q_3 \leq q_1 \wedge q_2$ and thus $m_3 \in \mathcal{S}_q\left(m_1\right) \cap \mathcal{S}_q\left(m_2\right)$. However $m_1 \sqcap_{q,2} m_2 \not\sqsubseteq_q m_3$ and thus $m_3 \notin \mathcal{S}_q\left(m_1 \sqcap_{q,2} m_2\right)$.

Fortunately, when $f \in \{pl, q\}$, the constraints of the minimization problem can be stored and updated iteratively, meaning that the complexity of the combination does not increase with ℓ. In practice, one needs to be able to compute combinations iteratively without storing the whole set of mass functions $\{m_1, \ldots, m_\ell\}$ and restart the combination from scratch when a new function $m_{\ell+1}$ arrives. This property is often referred to as **quasi-associativity**. Let c denote a set function from 2^Ω to $[0; 1]$ which is meant to store the problem constraints. Algorithm 1 allows to compute combinations using $\sqcap_{q,k}$ sequentially. The same algorithm works for $\sqcap_{pl,k}$. In practice, what we simply do is storing, for each set A, the lowest commonality (resp. plausibility) value encountered in $\{m_1, \ldots, m_\ell\}$.

Algorithm 1 Sequential combination using $\sqcap_{q,k}$

entries : $\{m_1, .., m_\ell\}$, $k \geq 2$.
$c \leftarrow \min\{q_1; q_2\}$ (entrywise minimum).
$m \leftarrow m_1 \sqcap_{q,k} m_2$.
for i from 3 to ℓ **do**
  $c \leftarrow \min\{c; q_i\}$ (entrywise minimum).
  $m \leftarrow \arg\min_{m'} d_{q,k}\left(m', m_\Omega\right)$ subject to $q' \leq c$.
end for
return m.

[2]Remember that $\circledwedge = \sqcap_{w,k}$ and $\circledwedge$ is associative [7].

It should be noted that this quasi-associativity is induced by the associativity of the entrywise minimum.

It can be argued that the choice of screening distances from the least committed mass function in definition 2 is somewhat arbitrary. The following lemma shows that, for $\sqcap_{q,k}$ and $\sqcap_{pl,k}$, another relevant choice yields the same operators:

Lemma 1. *For $f \in \{q, pl\}$ and for any finite integer k such that $k \geq 2$, one has:*

$$m_1 \sqcap_{f,k} .. \sqcap_{f,k} m_l = \underset{m \in \mathcal{S}_f(m_1) \cap .. \cap \mathcal{S}_f(m_\ell)}{\arg\min} d_{f,k}(m, m_\Omega) = \underset{m \in \mathcal{S}_f(m_1) \cap .. \cap \mathcal{S}_f(m_\ell)}{\arg\max} d_{f,k}(m, m_\emptyset). \tag{22}$$

Proof. We give a proof for $\sqcap_q$. The one for $\sqcap_{pl}$ follows a similar scheme.

Let us denote by m^* the mass function yielded by $m_1 \sqcap_q m_2$. For any mass function $m \in \mathcal{S}_q(m_1) \cap .. \cap \mathcal{S}_q(m_\ell)$, we thus have

$$\begin{aligned} \|q - q_\Omega\|_k &\geq \|q^* - q_\Omega\|_k, \\ \Leftrightarrow \|1 - q\|_k &\geq \|1 - q^*\|_k. \end{aligned}$$

The above inequality comes from the fact that commonalities for the vacuous mass function are constant with value one. Observing that there is a symmetry relating function $g(\mathbf{x}) = \|\mathbf{1} - \mathbf{x}\|$ with function $h(\mathbf{x}) = \|\mathbf{x}\|$ for any vector $\mathbf{x}$ in the unit hypercube, we deduce

$$\begin{aligned} \|q\|_k &\leq \|q^*\|_k, \\ \Leftrightarrow \|q - q_\emptyset\|_k &\leq \|q^* - q_\emptyset\|_k. \end{aligned}$$

The above inequality is obtained by remembering that $q_\emptyset$ has null value for all non-empty set. It has value one for $\emptyset$ but this is tantamount to add the same constant term to both sides of the inequality. □

Getting nearer to the least committed state of belief is thus equivalent to drifting apart from the most committed one for these two operators. Another interesting property to investigate is the compatibility with Dempster's conditioning (9), recalled in Section 2.2. The next proposition shows that it is retrieved as a special case of the $\sqcap_{q,k}$ conjunctive rule.

Proposition 1. *Let m_0 denote a mass function. For any finite integer k such that $k \geq 2$ and any subset $A \subseteq \Omega$, we have*

$$m_0 \sqcap_{q,k} m_A = m_{0|A}. \tag{23}$$

Proof. The commonality function corresponding to the categorical mass function m_A is given by

$$q_A(B) = \begin{cases} 1 \text{ if } B \subseteq A \\ 0 \text{ otherwise} \end{cases}. \tag{24}$$

From this, one obviously has $q_0(B)\, q_A(B) = q_0(B) \wedge q_A(B)$. Remembering that entrywise product of two commonality functions is the commonality function of their conjunctive combination, we have

$$\mathcal{S}_q(m_0) \cap \mathcal{S}_q(m_A) = \mathcal{S}_q\left(m_{0|A}\right).$$

By definition of $\mathcal{S}_q\left(m_{0|A}\right)$, its unique maximal element is $m_{0|A}$, meaning that $\forall m \in \mathcal{S}_q\left(m_{0|A}\right)$, one has $m \sqsubseteq_q m_{0|A}$. Now since we also have that $m_{0|A} \sqsubseteq_q m_\Omega$ and $d_{q,k}$ is $\sqsubseteq_q$-compatible, then

$$\underset{m \in \mathcal{S}_q(m_0) \cap \mathcal{S}_q(m_A)}{\arg\min} \quad d_{q,k}\left(m, m_\Omega\right) \quad = \quad m_{0|A}. \tag{25}$$

□

Concerning $\sqcap_{pl,k}$, we have not been able to prove (or refute) an equivalent result. The only case in which the result can be easily proved is when A has unit cardinality. Indeed, after conditioning on a singleton, plausibility and commonality functions coincide and the result proved in proposition 1 applies. In order to derive a possible counter-example, we sampled uniformly mass functions m_0 and m_A using the sampling procedures presented in [2]. After 1e6 runs, no counter-example was found for $k = 2$, therefore the result may be conjectured.

Another property that can be sometimes interesting is invariance with respect to refinement. Given two spaces Ω and Θ, a refinement r of Ω into Θ is a mapping from Ω to $2^\Theta \setminus \emptyset$ such that the family $\{r(\omega)\}_{\omega \in \Omega}$ is a partition of Θ which is an abstract space with greater cardinality than Ω. For any mass function m defined on Ω, another one m' on Θ is induced by r as follows:

$$m'\left(\bigcup_{\omega \in A} r(\omega)\right) = m(A), \forall A \subseteq \Omega. \tag{26}$$

The idea behind refinement is that one may be interested in reasoning at different scales, starting with a coarse one and then transferring our beliefs on a finer one using mapping r. As shows the next example, the operators introduced in this article are not immune to such an operation, for the main reason that distances are in general not invariant with respect to refinements (for a discussion about this, see [8]).

Example 3. Let $\Omega = \{a, b, c\}$ and $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ denote two spaces. Suppose there exist a refinement r such that :

$$\begin{aligned} r(a) &= \{\theta_1, \theta_4\}, \\ r(b) &= \{\theta_2\}, \\ r(c) &= \{\theta_3\}. \end{aligned}$$

Let us introduce the following mass functions on Ω:

subset	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	Ω
m_1	0.1	0	0.1	0.5	0.1	0	0.1	0.1
m_2	0	0	0	0.3	0.1	0.3	0	0.3

Let us denote by m'_1 and m'_2 the mass functions on Θ induced by r from m_1 and m_2, respectively. The mass function $m'_{1\sqcap_{q,2}2}$ induced by r from $m_1 \sqcap_{q,2} m_2$ is not equal to $m'_1 \sqcap_{q,2} m'_2$. In particular, we have $m'_1 \sqcap_{q,2} m'_2(\{\theta_3\}) = 0.2$ while $m'_{1\sqcap_{q,2}2}(\{\theta_3\}) = 0.1$.

Although informative partial orders are preserved after refinement, the sets of more informative functions $\mathcal{S}_f(m_i)$ are different. In example 3, the hypothesis a is refined into two elements: θ_1 and θ_4. This implies increased freedom in the selection of the mass function minimizing the distance from the vacuous function. In general, there is no reason why this solution should be in correspondence (through mapping r) with the solution obtained without refining.

A last point that deserves investigation is the presence of a neutral element, *i.e.* a function m_{e} such that $m \sqcap_{f,k} m_{\mathrm{e}} = m, \forall m \in \mathcal{M}$.

Proposition 2. *For any finite integer k such that $k \geq 2$ and any $f \in \{q, pl\}$, the unique neutral element of operator $\sqcap_{f,k}$ is the vacuous mass function m_Ω.*

Proof. The vacuous mass function m_Ω is the maximum of $(\mathcal{M}, \sqsubseteq_f)$ for $f \in \{q, pl\}$ which implies that $\mathcal{S}_f(m_\Omega) = \mathcal{M}$. Consequently, the feasible set of $m \sqcap_{f,k} m_\Omega$ is $\mathcal{S}_f(m)$. By defintion of $\mathcal{S}_f(m)$, we have $(m \sqcap_{f,k} m_\Omega) \sqsubseteq_f m$. Since distance d is consistent with $\sqsubseteq_f$, $(m \sqcap_{f,k} m_\Omega) \sqsubseteq_f m \sqsubseteq_f m_\Omega$ implies $d(m \sqcap_{f,k} m_\Omega, m_\Omega) \geq d(m, m_\Omega)$. But $m \sqcap_{f,k} m_\Omega$ is by definition the unique minimizer in $\mathcal{S}_f(m)$ of the distance from m_Ω, hence $m \sqcap_{f,k} m_\Omega = m$.

Furthermore, suppose $m_{\mathrm{e}} \neq m_\Omega$ is a neutral element. Since m_Ω is neutral $m_{\mathrm{e}} \sqcap_{f,k} m_\Omega = m_{\mathrm{e}}$ but since m_{e} is neutral as well then $m_{\mathrm{e}} \sqcap_{f,k} m_\Omega = m_\Omega$, hence a contradiction. □

This property is desirable because the vacuous mass function represents the absence of information and in a conjunctive combination context, we expect the absence of information to have no impact on the combination result.

4 Disjunctive combination using partial orders

In the same fashion as the conjunctive case, one can consider that a mass function m^* resulting from a disjunction should be (i) less informative (in the sense of some partial order $\sqsubseteq_f$) than any $m_1, \ldots, m_\ell$ and (ii) should be among the most committed elements (in terms of information) among those. This is a dual reasoning as LCP. Formally speaking, if we denote by

$$\mathcal{G}_f(m_i) := \{m \in \mathcal{M} \mid m_i \sqsubseteq_f m\} \tag{27}$$

the set of all mass functions less informative than m_i, then m^* should be such that:

(i) $m^* \in \mathcal{G}_f(m_1) \cap \ldots \cap \mathcal{G}_f(m_\ell)$,

(ii) $\nexists m' \in \mathcal{G}_f(m_1) \cap \ldots \cap \mathcal{G}_f(m_\ell)$ such that $m' \sqsubset_f m^*$.

Again, such a procedure does not lead to a unique solution in general (except when $f = v$). One way to circumvent this issue is to select the mass function in $\mathcal{G}_f(m_1) \cap \ldots \cap \mathcal{G}_f(m_\ell)$ with minimal distance from the minimum of $(\mathcal{M}, \sqsubseteq_f)$ (if it exists) as long as the chosen metric is $\sqsubset_f$-compatible.

Let us focus on $f \in \{pl, q\}$, where a unique minimal element exists and is $m_\emptyset$. According to corollary 3 in [20], we know that this problem is a convex optimization problem with a unique solution if the chosen distance $d_{f,k}$ is $\sqsubset_f$-compatible and if $2 \leq k < \infty$. Let $\sqcup_{f,k}$ denote this operator which is formally defined as follows:

Definition 3. for any set of ℓ functions $\{m_1, .., m_\ell\}$, we have

$$m_1 \sqcup_{f,k} .. \sqcup_{f,k} m_l = \underset{m \in \mathcal{G}_f(m_1) \cap .. \cap \mathcal{G}_f(m_\ell)}{\arg\min} d_{f,k}(m, m_\emptyset). \tag{28}$$

The fact that $m_\emptyset$ is the most committed mass function for $\sqsubseteq_q$ and $\sqsubseteq_{pl}$ is not very intuitive. The following lemma delivers a better intuition as to what operators $\sqcup_{q,k}$ and $\sqcup_{pl,k}$ consist of, as it shows that they can be understood as the maximization of the distance from the vacuous mass function which is more intuitive.

Lemma 2. *For $f \in \{q, pl\}$ and for any finite integer k such that $k \geq 2$, one has:*

$$m_1 \sqcup_{f,k} .. \sqcup_{f,k} m_l = \underset{m \in \mathcal{G}_f(m_1) \cap .. \cap \mathcal{G}_f(m_\ell)}{\arg\min} d_{f,k}(m, m_\emptyset) = \underset{m \in \mathcal{G}_f(m_1) \cap .. \cap \mathcal{G}_f(m_\ell)}{\arg\max} d_{f,k}(m, m_\Omega). \quad (29)$$

The proof of this lemma is identical to that of Lemma 1.

The combination operators $\sqcup_{q,k}$ and $\sqcup_{pl,k}$ have similar properties as their conjunctive counterparts. They are commutative, idempotent and quasi-associative[3] but not invariant to refinement. Quasi-associativity is achieved using algorithm 2 which is almost the same as algorithm 1.

Algorithm 2 Sequential combination using $\sqcup_{q,k}$

```
entries : {m_1, .., m_ℓ}, k ≥ 2.
c ← max {q_1; q_2} (entrywise maximum).
m ← m_1 ⊓_{q,k} m_2.
for i from 3 to ℓ do
   c ← max {c; q_i} (entrywise maximum).
   m ← arg min_{m'} d_{q,k} (m', m_∅) subject to q' ≥ c.
end for
return m.
```

The neutral element of some of the disjunctive operators is given by the following proposition.

Proposition 3. *For any finite integer k such that $k \geq 2$ and any $f \in \{q, pl\}$, the unique neutral element of operator $\sqcup_{f,k}$ is the total conflict mass function $m_\emptyset$.*

The proof of proposition 3 is very similar as the one of proposition 2 and is thus omitted. The key point is that $m_\emptyset$ is the minimum of $(\mathcal{M}, \sqsubseteq_f)$ for $f \in \{q, pl\}$.

Just for a quick illustration we provide the following example which is a continuation of [7, example 7].

Example 4. Let $\Omega = \{a, b, c\}$. Here are two subnormal mass functions along with their combinations under our new operators and other standard approaches.

subset	$\emptyset$	$\{a\}$	$\{b\}$	$\{a,b\}$	$\{c\}$	$\{a,c\}$	$\{b,c\}$	Ω
m_1	0.1	0	0	0.3	0	0	0.6	0
m_2	0.1	0	0.5	0	0	0	0.4	0
$m_1 \sqcup_{q,2} m_2$	0.1	0	0	0.3	0	0	0.6	0
$m_1 \sqcup_{pl,2} m_2$	0.1	0	0	0.3	0	0	0.6	0
$m_1 \bigcirc\!\!\!\!\cup\, m_2$	0.01	0	0.05	0.18	0	0	0.64	0.12
$m_1 \bigcirc\!\!\!\!\vee\, m_2$	0.006	0	0.0298	0.1071	0	0	0.2143	0.6429

[3] Again, when $f = v$, $\sqcup_{f,k} = \bigcirc\!\!\!\!\vee$ and this rule is associative [7].

Table 1: Basic properties of operators $\textcircled{\cap}$, $\oplus$, $\textcircled{\wedge}$ and $\sqcap_{f,k}$.

operator	condition for use	commutativity	associativity	idempotence	invariance w.r.t. refinement	neutral element
$\textcircled{\cap}$	none	yes	yes	no	yes	m_Ω
$\oplus$	$m_{1\textcircled{\cap}2}(\emptyset) < 1$	yes	yes	no	yes	m_Ω
$\textcircled{\wedge}$	$m_1(\Omega) > 0$ and $m_2(\Omega) > 0$	yes	yes	yes	yes	none
$\sqcap_{q,k}$	none	yes	quasi	yes	no	m_Ω
$\sqcap_{pl,k}$	none	yes	quasi	yes	no	m_Ω

We see that the disjunctive and bold rules transfer much more mass to Ω than operators $\sqcup_{q,2}$ and $\sqcup_{pl,2}$ do. Also, observe that $m_1 \sqcup_{q,2} m_2 = m_1 \sqcup_{pl,2} m_2 = m_1$ because $m_2 \sqsubset_q m_1$ and $m_2 \sqsubset_{pl} m_1$.

5 Related works: discussion and experiments

This section studies the relation between the current work and the main operators used to combine belief functions, both in terms of basic properties and experiments. They demonstrate that our distance-based operators allow to redistribute masses more gradually than standard approaches. We also discuss the influence of parameter k and f in the mass function returned by our operators.

5.1 A comparison with related works in the conjunctive case

As said earlier, there are many works that have addressed the problem of deriving alternatives to Dempster's rule or the conjunctive rule that do not rely on independence assumptions.

A principled and common approach is to rely on a set of axiomatic properties [11] or to adapt existing rules from other frameworks [9]. In practice, such axioms seldom lead to a unique solution, and it is then necessary to advocate more practical solutions. Our rule can be seen as an instance of such an approach, where the axiom consists in using the LCP over sets of f-included masses, and the practical solution is to use a distance compliant with such an axiom. Cattaneo's solution [3] as well as Denoeux [7] cautious rules can also be seen as instances of the same principle. The former proposes to solve a conflict minimization problem rather than minimizing the informative content (thus not strictly following an LCP principle), while the latter focuses on using the set $\mathcal{S}_w(m_1) \cap \ldots \cap \mathcal{S}_w(m_\ell)$ and the order $\sqsubseteq_w$, and demonstrates that in this case there is a unique LCP solution known in closed form. Finally, an idea of combination by distance minimization is suggested but not studied in [4]. The author pursues a different goal anyway as the constraints are on marginal mass functions.

We compare our approach with rules $\textcircled{\cap}$, $\oplus$ and $\textcircled{\wedge}$ which are the most frequently used. Table 1 summarizes some basic theoretical properties satisfied by operators $\textcircled{\cap}$, $\oplus$, $\textcircled{\wedge}$ and $\sqcap_{f,k}$.

From a practical point of view, let us stress that combinations using $\sqcap_{f,k}$ for $f \in \{pl, q\}$ and $k = 2$ are really easy to compute. Indeed, quadratic programming techniques can solve equation (21) in a very few iterations. The function $m_\emptyset$ can be used to initialize the minimization as we are sure that it belongs to

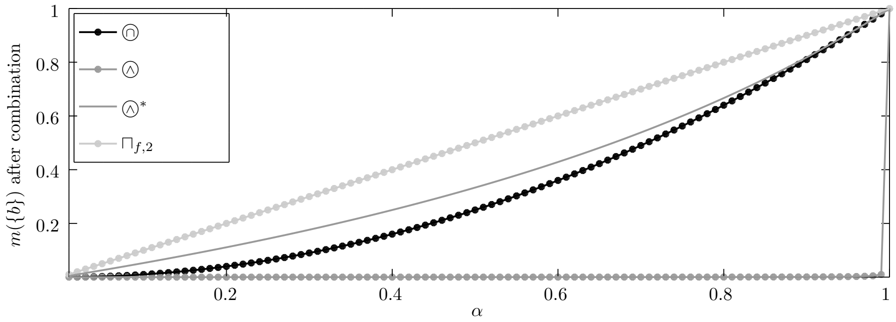

Figure 1: Mass assigned to $\{b\}$ after combination of $m_1 = \alpha m_{\{b\}} + (1-\alpha)\, m_{\{a\}}$ and $m_2 = \alpha m_{\{b\}} + (1-\alpha) m_{\{c\}}$ with $\textcircled{\cap}$, $\textcircled{\wedge}$ and $\sqcap_{f,2}$.

$\mathcal{S}_f\,(m_1) \cap \ldots \cap \mathcal{S}_f\,(m_\ell)$. See also B for more details on how to implement this quadratic program.

Let us illustrate the operator discrepancies on a simple situation inspired from Zadeh's counter-example [29]. Suppose $m_1 = \alpha m_{\{b\}} + (1-\alpha)\, m_{\{a\}}$ and $m_2 = \alpha m_{\{b\}} + (1-\alpha)\, m_{\{c\}}$ are two mass functions on a frame $\Omega = \{a, b, c\}$. Figure 1 shows the mass assigned to $\{b\}$, the commonly supported element of m_1 and m_2, after combination by $\textcircled{\cap}$, $\textcircled{\wedge}$ and $\sqcap_{f,2}$. The same masses are obtained for $f \in \{pl, q\}$. A very small mass $\epsilon = 1e-4$ was assigned to Ω while a mass $\frac{\epsilon}{2}$ was removed from each focal element of each input mass function when using $\textcircled{\wedge}$ so as to circumvent the non-dogmatic constraint.

As could be expected, our rule tries to maintain as much evidence on $\{b\}$ as possible. A striking fact is that we have obviously $m_1 \sqcap_{f,2} m_2\,(\{b\}) = \alpha$. More precisely, we have $m_1 \sqcap_{f,2} m_2 = (1-\alpha)\, m_\emptyset + \alpha m_{\{b\}}$. This result can be proved for any finite $k \geq 2$ when $f = q$. Let $q_{1\wedge 2}$ denote the entrywise minimum of functions q_1 and q_2. In this particular setting, $q_{1\wedge 2}$ happens to be a valid commonality function. Consequently, $m_{1\wedge 2} \in \mathcal{S}_q\,(m_1) \cap \mathcal{S}_q\,(m_2)$. By definition of the partial order $\sqsubseteq_q$, for any function $m \in \mathcal{S}_q\,(m_1) \cap \mathcal{S}_q\,(m_2)$, we have $m \sqsubseteq_q m_{1\wedge 2}$. Since we also have $m_{1\wedge 2} \sqsubseteq_q m_\Omega$ and $d_{q,k}$ is $\sqsubseteq_q$-compatible, then $m_1 \sqcap_{q,k} m_2 = m_{1\wedge 2}$. In other words, our approach coincides with the minimum rule applied to commonalities in this case. When $f = pl$, the result can also be proved. For any $m \in \mathcal{S}_{pl}\,(m_1) \cap \mathcal{S}_{pl}\,(m_2)$, the constraints $pl\,(\{a\}) = pl\,(\{c\}) = 0$ imply that only $\{b\}$ and $\emptyset$ are possible focal sets for m. More precisely, this actually implies that $\mathcal{S}_{pl}\,(m_1) \cap \mathcal{S}_{pl}\,(m_2)$ is the segment $(1-\beta)\, m_\emptyset + \beta m_{\{b\}}$ in $\mathcal{M}$ parametrized by $\beta \in [0; \alpha]$. $\sqsubseteq_{pl}$ is a total order for this segment. From relation (18), we obtain $m_1 \sqcap_{pl,k} m_2 = (1-\alpha)\, m_\emptyset + \alpha m_{\{b\}}$.

A closed form expression for the other rules can also be obtained. It is easy to see that $m_1 \textcircled{\cap} m_2 = \left(1-\alpha^2\right) m_\emptyset + \alpha^2 m_{\{b\}}$. Concerning the cautious rule, taking the limit $\epsilon \to 0$, we obtain

$$m_1 \textcircled{\wedge} m_2 = \begin{cases} m_\emptyset & \text{if } \alpha < 1 \\ m_{\{b\}} & \text{if } \alpha = 1 \end{cases}.$$

This example shows also that the behavior of Denœux's cautious rule $\textcircled{\wedge}$ may not be so cautious, as it keeps no mass on $\{b\}$ except when $\alpha = 1$. This is a

Table 2: Basic properties of operators $\textcircled{\cup}$, $\textcircled{\vee}$ and $\sqcup_{f,k}$.

operator	condition for use	commutativity	associativity	idempotence	invariance w.r.t. refinement	neutral element
$\textcircled{\cup}$	none	yes	yes	no	yes	$m_\emptyset$
$\textcircled{\vee}$	$m_1(\emptyset) > 0$ and $m_2(\emptyset) > 0$	yes	yes	yes	yes	none
$\sqcup_{q,k}$	none	yes	quasi	yes	no	$m_\emptyset$
$\sqcup_{pl,k}$	none	yes	quasi	yes	no	$m_\emptyset$

quite bold behavior, due mainly to the fact that $\mathcal{S}_w$ induces stronger constraints than $\mathcal{S}_{pl}$ or $\mathcal{S}_q$. This clearly shows that while idempotence is a pre-requisite to have a cautious attitude towards source (in)dependence, it is not sufficient to guarantee a really cautious behavior when mass functions are not identical. Even the conjunctive rule appears to have an intermediate behavior as compared to the two others, hence could be termed as more cautious than Denœux's rule. In [7, example 3], Denœux actually shows that using $\textcircled{\wedge}$ can yield a mass function that is w-included in the result of the conjunctive combination. It should also be stressed that this is a limit use case for $\textcircled{\wedge}$, hence arguably an unfavourable one.

The normalized versions of these three rules deserve also some comments. This time, we obtain $m_1 \sqcap^*_{q,k} m_2 = m_1 \sqcap^*_{pl,k} m_2 = m_1 \oplus m_2 = m_{\{b\}}$ which is the result criticized by Zadeh. In contrast, the normalized cautious rule achieves a progressive reduction of the support given to $\{b\}$ as α decreases. The normalized cautious rule appears to offer an intermediate behavior as compared to the conjunctive rule and either of the unnormalized operator $\sqcap_{q,k}$ or $\sqcap_{pl,k}$. In particular, when $\alpha = \frac{1}{2}$, $m_1 \textcircled{\wedge}^* m_2$ is the uniform Bayesian mass function whereas operators $\sqcap_{q,k}$ and $\sqcap_{pl,k}$ are still giving some support to $\{b\}$ solely. This time, the rule $\textcircled{\wedge}^*$ appears indeed more cautious than ours, but could be argued to no longer be really conjunctive, as it supports every element whereas each source respectively discarded one as totally impossible.

5.2 A comparison with related works in the disjunctive case

Similarly as in the conjunctive case, we also give a comparison with popular disjunctives rules: $\textcircled{\cup}$ and $\textcircled{\vee}$. Table 2 summarizes some basic theoretical properties satisfied by operators $\textcircled{\cup}$, $\textcircled{\vee}$ and $\sqcup_{f,k}$.

Let us illustrate the disjunctive operator discrepancies on a simple situation analogous to the experiment presented in the conjunctive case. Suppose $m_1 = \alpha m_{\{a,b\}} + (1-\alpha) m_{\{a,c\}}$ and $m_2 = \alpha m_{\{a,b\}} + (1-\alpha) m_{\{b,c\}}$ are two mass functions on a frame $\Omega = \{a, b, c\}$. Figure 2 shows the mass assigned to $\{a, b\}$ after combination by $\textcircled{\cup}$, $\textcircled{\vee}$ and $\sqcup_{f,2}$. The same masses are obtained for $f \in \{pl, q\}$. A very small mass $\epsilon = 1e-4$ was assigned to $\emptyset$ while a mass $\frac{\epsilon}{2}$ was removed from each focal element of each input mass function when using $\textcircled{\vee}$ so as to circumvent the normalization constraint.

The aspect of figure 2 is remarkably similar to the conjunctive case but the conclusions that we will draw from it are different. As could be expected, our rule tries to maintain as much evidence on $\{a, b\}$ as possible. A striking fact is that we have obviously $m_1 \sqcup_{f,2} m_2(\{a, b\}) = \alpha$. More precisely, we have $m_1 \sqcup_{f,2} m_2 = (1-\alpha) m_\Omega + \alpha m_{\{a,b\}}$.

This result can be proved for any finite $k \geq 2$ when $f = q$. Let $q_{1\vee 2}$

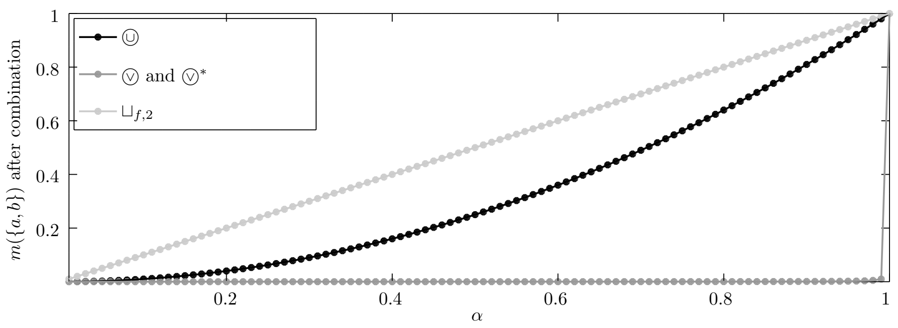

Figure 2: Mass assigned to $\{a,b\}$ after combination of $m_1 = \alpha m_{\{a,b\}} + (1-\alpha)\, m_{\{a,c\}}$ and $m_2 = \alpha m_{\{a,b\}} + (1-\alpha) m_{\{b,c\}}$ with $\textcircled{\cup}$, $\textcircled{\vee}$ and $\sqcup_{f,2}$.

denote the entrywise maximum of functions q_1 and q_2. We have $q_{1\vee 2}(\{a\}) = q_{1\vee 2}(\{b\}) = 1$ which implies that for any $m \in \mathcal{G}_q(m_1) \cap \mathcal{G}_q(m_2)$, only supersets of $\{a,b\}$ can be focal sets of m. In this example, this means that $m = \beta m_{\{a,b\}} + (1-\beta)\, m_\Omega$ with $\beta \in [0;1]$. Observing that if q denotes the commonality function in correspondence with m, we also have

$$
\begin{aligned}
q_{1\vee 2}(\{c\}) &\leq q(\{c\}), \\
\Leftrightarrow 1-\alpha &\leq \sum_{B \subseteq \{c\}} m(B).
\end{aligned}
$$

Since Ω is the only set that is a superset of both $\{a,b\}$ and $\{c\}$, we deduce that $m(\Omega) \geq 1-\alpha$ or equivalently $\beta \leq \alpha$.

More precisely, this actually implies that $\mathcal{G}_q(m_1) \cap \mathcal{G}_q(m_2)$ is the segment $(1-\beta)\, m_\Omega + \beta m_{\{a,b\}}$ in $\mathcal{M}$ parametrized by $\beta \in [0;\alpha]$. $\sqsubseteq_q$ is a total order for this segment. This segment can also be seen as the set of mass functions obtained by discounting $\alpha m_{\{a,b\}} + (1-\alpha)\, m_\Omega$. From relation (18), we obtain $m_1 \sqcup_{q,k} m_2 = (1-\alpha)\, m_\Omega + \alpha m_{\{a,b\}}$. When $f = pl$, the same reasoning applies.

A closed form expression for the other rules can also be obtained. It is easy to see that $m_1 \textcircled{\cup} m_2 = \left(1-\alpha^2\right) m_\Omega + \alpha^2 m_{\{a,b\}}$. Concerning the bold rule, taking the limit $\epsilon \to 0$, we obtain

$$
m_1 \textcircled{\vee} m_2 = \begin{cases} m_\Omega & \text{if } \alpha < 1 \\ m_{\{a,b\}} & \text{if } \alpha = 1 \end{cases}.
$$

Like in the conjunctive example, the behavior of the bold rule $\textcircled{\vee}$ is symptomatic of the fact that $\mathcal{G}_w$ induces stronger constraints than $\mathcal{G}_{pl}$ or $\mathcal{G}_q$. The bold rule keeps no mass on $\{a,b\}$ except when $\alpha = 1$. Finally, the disjunctive rule appears to have an intermediate behavior as compared to the two others. Also, this time all normalized versions of these rules coincide with their unnormalized counterparts.

5.3 Sensitivity with respect to parameters

In this subsection, we investigate the influence of the parameters of the newly introduced operators starting with the norm integer parameter k on the re-

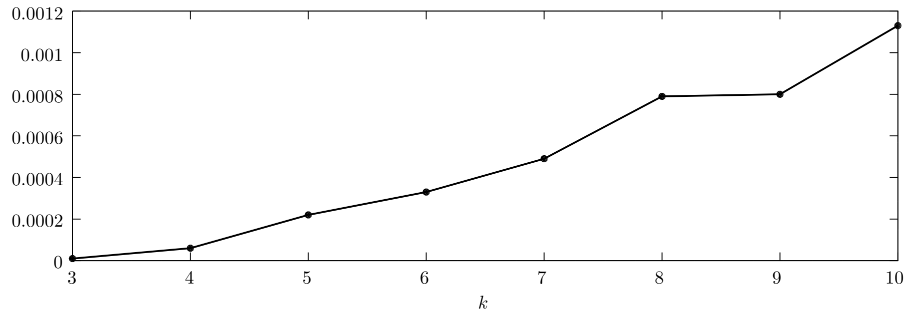

Figure 3: Sensitivity w.r.t. k: estimated probability to observe a mass discrepancy higher than $1e-5$ when using $\sqcap_{q,2}$ as compared to $\sqcap_{q,k}$ ($k \in \{3, .., 10\}$).

sults produced by the introduced operators. We restrict this analysis to the conjunctive operators relying on $\sqsubset_q$.

Figure 3 gives the frequentist probability estimates that a discrepancy (maximal absolute difference between mass functions) greater than $1e-5$ is observed between $m_1 \sqcap_{q,2} m_2\,(A)$ and $m_1 \sqcap_{q,k} m_2\,(A)$ (for any A) and with $k \in \{3, ..10\}$. The input mass functions m_1 and m_2 are sampled uniformly in $\mathcal{M}$ [2]. The estimated probabilities are computed after $1e5$ runs.

According to this figure, discrepancies are more and more likely to occur as k increases. It can also be concluded that choosing values of $k+1$ instead of k has a very limited impact on the returned results. Even for a quite large gap of values for k, significant discrepancies are rare events: the probability to witness a discrepancy from $k=2$ as compared to $k=10$ is less than $1.2e-3$.

Another parameter to choose from is the partial order. We have argued that $f \in \{q, pl\}$ should be preferred because the corresponding partial orders are the finest ones. It is however more arbitrary to choose between q and pl-inclusion. As already illustrated in example 1, the operators do not coincide and can lead to significantly different results. However, we could hardly expect to have equivalent results when using different partial orders (hence different sets of solutions) and distances (different optimisation criteria).

The frequentist probability (estimated from $1e5$ trials[4]) to observe a discrepancy greater than $1e-5$ between $m_1 \sqcap_{q,2} m_2$ and $m_1 \sqcap_{pl,2} m_2$ is 0.8807 which is pretty high. Figure 4 shows how the discrepancies are distributed. The most probable case (with probability of 0.12) is to obtain identical results. When the operators leads to different results, the maximal absolute difference between masses is concentrated around 0.09, which is significant but also not too overwhelmingly high.

[4]Mass functions are sampled uniformly in the mass space.

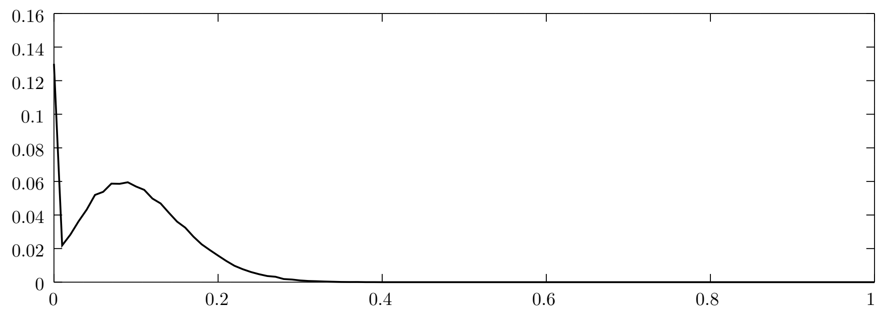

Figure 4: Sensitivity w.r.t. $f \in \{q, pl\}$: estimated probability distribution of the observed discrepancy (in norm L_∞) when using $\sqcap_{q,2}$ as compared to $\sqcap_{pl,2}$.

6 Conclusion

This paper introduces cautious conjunctive combination operators for mass functions by relying on constraints inducing a more informative mass function than the combined ones on one hand, and on the minimization of distances to total ignorance on the other hand. The metrics used in the minimization procedure must be compatible with partial orders comparing informative contents. A dual idea is also developed to introduce disjunctive rules. It relies on constraints inducing a less informative mass function than the combined ones on one hand, and on the maximization of distances to total ignorance on the other hand.

These procedures give rise to several commutative, idempotent and quasi-associative combination operators. It is also noteworthy that conjunctive ones are in line with the philosophy of the LCP principle. Our distance optimization approach allows these new operators to be easily interpretable and to rely on sound justifications.

Simple experiments demonstrate that the introduced operators allow one to redistribute masses more gradually as compared to standard approaches and thus comply with some user's expectations, in contrast with some other well-known rules.

We believe distance based approaches offer promising perspectives in many problems within the belief function theory framework or the imprecise probabilities framework, and provides a new view of combination relying on an optimization standpoint. For example, we could try to extend our framework to n-monotone capacities, or to rules combining conjunctive and disjunctive behaviors.

Another question that deserves investigations is the connection of our conjunctive operators with conflict minimization ones. Distances and conflict are not unrelated notions [21] although they cannot be directly interchanged [1]. Indeed, although conflict minimization is never explicitely sought in the proposed approach, lemma 1 suggests that the result of the combination is the farthest possible to the total conflict mass function. Consequently, there are some situations in which both approaches coincide and those situations should be identified.

A Comments on distance minimization in the d and s ordering cases

In this appendix, we comment on the properties of rules that could potentially be induced by problem (21) or (28) if an evidential distance was proved to be $\sqsubset_d$ or $\sqsubset_s$-compatible.

Concerning quasi-associativity of an operator $\sqcap_f$ or $\sqcup_f$ ($f \in \{d, s\}$), the picture is not as simple as for $f \in \{q, pl\}$. Indeed, to our knowledge the characterization of the set of inner approximations of a mass function does not translate in compact constraints for these partial orders. For instance, we know from Cuzzolin [5] that $\mathcal{S}_d\left(m_i\right)$ is a simplex with at most N vertices. Yet, in general the intersection of simplices is not a simplex but a polytope whose vertices are not easily derived and can increase in number after each iteration. Whether there is an easy way to characterize these intersections is a topic for further research.

For other properties, the same reasoning as in the $f \in \{q, pl\}$ cases apply as well when $f \in \{d, s\}$. In particular, commutativity and the conjunctive/disjunctive nature are proved identically. Also, m_Ω is necessarily the neutral element of operators $\sqcap_d$ or $\sqcap_s$ as m_Ω is the maximum of both $(\mathcal{M}, \sqsubseteq_d)$ and $(\mathcal{M}, \sqsubseteq_s)$. Likewise, $m_\emptyset$ is the minimum of both $(\mathcal{M}, \sqsubseteq_d)$ and $(\mathcal{M}, \sqsubseteq_s)$ and thus $m_\emptyset$ is the neutral element of operators $\sqcup_d$ and $\sqcup_s$.

B Quadratic programming implementation

In this appendix, we show how the problem specified by equation (21) can be reshaped as a traditional quadratic programming one when $k = 2$. This reshaping relies on matrix calculus with belief functions. In short, any set function used for evidence representation can be seen as a vector in $\mathbb{R}^N$. Under this setting, many belief function operations are translated into dot products between matrices and vectors. We refer to Smets' paper [26] for an exemplified presentation on this topic. The vector version of any evidential function will be denoted in bold letters, *e.g.* $\mathbf{m}_k$ is the vector version of mass function m_k.

Most solvers available in scientific programming libraries allow to solve quadratic programming problems defined as follows:

$$\begin{aligned}
\mathbf{x}^* = \quad & \arg\min_{\mathbf{x}} \ \tfrac{1}{2}\mathbf{x}^t\mathbf{Q}\mathbf{x} + \mathbf{x}^t\mathbf{a} \quad , \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{c}_1 \quad , \\
& \mathbf{c}_2^- \leq \mathbf{x} \leq \mathbf{c}_2^+ \quad , \\
& \mathbf{c}_3^- \leq \mathbf{B}\mathbf{x} \leq \mathbf{c}_3^+ \quad ,
\end{aligned} \tag{30}$$

where $\mathbf{Q}$ is a symmetric matrix, $\mathbf{a}$ is a vector with the same dimensionality as $\mathbf{x}$ and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{c_1}$, $\mathbf{c}_2^-$, $\mathbf{c}_2^+$, $\mathbf{c}_3^-$ and $\mathbf{c}_3^+$ are known matrices and vectors depicting linear constraints.

Let us now explain how our distance based combination operators can be reshaped in such a problem. We start with $\sqcap_q$ and comments on other orders and for the disjunctive case will follow. There exists a definite positive matrix $\mathbf{M}$ that maps mass vectors to commonality vectors. Matrix $\mathbf{M}$ is called the inclusion matrix because

$$M_{ij} = \begin{cases} 1 \text{ if } E_i \subseteq E_j \\ 0 \text{ otherwise} \end{cases},$$

where M_{ij} is the entry of $\mathbf{M}$ where i encodes set E_i and j encodes set E_j. For any mass function m_k, we have :

$$\mathbf{q}_k = \mathbf{M}\mathbf{m}_k.$$

Following matrix notations, problem$\tilde{}$(21) for a pair of mass functions and when $k = 2$ writes as:

$$\begin{aligned}
m_1 \sqcap_{q,2} m_2 &= \underset{m \in \mathcal{S}_q^+(m_1) \cap \mathcal{S}_q^+(m_2)}{\arg\min} \; d_{q,2}\left(m, m_\Omega\right), \\
&= \underset{m \in \mathcal{S}_q^+(m_1) \cap \mathcal{S}_q^+(m_2)}{\arg\min} \; \left\| \mathbf{M}\mathbf{m} - \mathbf{M}\mathbf{m}_\Omega \right\|_2, \\
&= \underset{m \in \mathcal{S}_q^+(m_1) \cap \mathcal{S}_q^+(m_2)}{\arg\min} \; \left(\mathbf{M}\mathbf{m} - \mathbf{M}\mathbf{m}_\Omega \right)^t \left(\mathbf{M}\mathbf{m} - \mathbf{M}\mathbf{m}_\Omega \right), \\
&= \underset{m \in \mathcal{S}_q^+(m_1) \cap \mathcal{S}_q^+(m_2)}{\arg\min} \; \mathbf{m}^t\mathbf{M}^t\mathbf{M}\mathbf{m} + \mathbf{m}_\Omega^t\mathbf{M}^t\mathbf{M}\mathbf{m}_\Omega - 2\mathbf{m}^t\mathbf{M}^t\mathbf{M}\mathbf{m}_\Omega, \\
&= \underset{\substack{\mathbf{m} \in \mathbb{R}^N \text{ subject to:} \\ \mathbf{1}^t\mathbf{m} = 1 \\ \mathbf{0} \leq \mathbf{m} \leq \mathbf{1} \\ \mathbf{0} \leq \mathbf{M}\mathbf{m} \leq \min\{\mathbf{q}_1; \mathbf{q}_2\}}}{\arg\min} \; \frac{1}{2}\mathbf{m}^t\mathbf{M}^t\mathbf{M}\mathbf{m} + \mathbf{m}^t\left(-\mathbf{M}^t\mathbf{M}\mathbf{m}_\Omega \right).
\end{aligned}$$

Now the solver described by (30) can be employed by choosing:

- $\mathbf{Q} \leftarrow \mathbf{M}^t\mathbf{M}$,
- $\mathbf{a} \leftarrow -\mathbf{M}^t\mathbf{M}\mathbf{m}_\Omega$,
- $\mathbf{A} \leftarrow \mathbf{1}^t$,
- $\mathbf{c}_1 \leftarrow 1$,
- $\mathbf{c}_2^- \leftarrow \mathbf{0}$,
- $\mathbf{c}_2^+ \leftarrow \mathbf{1}$,
- $\mathbf{B} \leftarrow \mathbf{M}$,
- $\mathbf{c}_3^- \leftarrow \mathbf{0}$,
- $\mathbf{c}_3^+ \leftarrow \min\{\mathbf{q}_1, \mathbf{q}_2\}$,

where $\mathbf{0}$ is the null vector and $\mathbf{1}$ is the all-one vector. The solver will return the mass vector $\mathbf{m}_1 \sqcap_{q,2} \mathbf{m}_2$. The optimization can start at point $\mathbf{m}_\emptyset$ since we are sure that it is compliant with all constraints.

For the plausibility based operator $\sqcap_{pl,2}$, one just has to replace matrix $\mathbf{M}$ with $1 - \mathbf{J}\mathbf{M}^t$ where $\mathbf{J}$ is the binary anti-diagonal matrix. Of course, $\min\{\mathbf{pl}_1, \mathbf{pl}_2\}$ is assigned to constraint $\mathbf{c}_3^+$.

Finally, in the disjunctive case, the same settings as in the conjunctive case are used for operators $\sqcup_{q,2}$ and $\sqcup_{pl,2}$ respectively. Only constraints $\mathbf{c}_3^-$ and $\mathbf{c}_3^+$ are adapted to problem (28) as well as the initial mass vector to start the optimization which is now $\mathbf{m}_\Omega$.

GNU Octave and Matlab implementations are available at this link: `https://github.com/john-klein/Conjunctive-and-Disjunctive-combination-by-distance-minimization`.

References

[1] Thomas Burger. Geometric interpretations of conflict: A viewpoint. In Fabio Cuzzolin, editor, *Belief Functions: Theory and Applications*, volume 8764 of *Lecture Notes in Computer Science*, pages 412–421. Springer International Publishing, 2014.

[2] Thomas Burger and Sebastien Destercke. How to randomly generate mass functions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, pages 645–673, 2013.

[3] M. E. G. V. Cattaneo. Combining belief functions issued from dependent sources. In J.M. Bernad, T. Seidenfeld, and M. Zaffalon (Eds.), editors, *Third International Symposium on Imprecise Probabilities and Their Applications (SIPTA'03)*, pages 133–147. Carleton Scientific, Lugano (Switzerland), 2003.

[4] Marco E.G.V. Cattaneo. Belief functions combination without the assumption of independence of the information sources. *International Journal of Approximate Reasoning*, 52(3):299 – 315, 2011. Dependence Issues in Knowledge-Based Systems.

[5] Fabio Cuzzolin. Geometry of Dempster's rule of combination. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, 34(2):961 – 977, 2004.

[6] A.P. Dempster. Upper and lower probabilities induced by a multiple valued mapping. *Annals of Mathematical Statistics*, 38:325 – 339, 1967.

[7] Thierry Denœux. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, 172:234 – 264, 2008.

[8] S. Destercke and T. Burger. Toward an axiomatic definition of conflict between belief functions. *Cybernetics, IEEE Transactions on*, 43(2):585–596, April 2013.

[9] Sébastien Destercke and Didier Dubois. Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. *Information Sciences*, 181(18):3925–3945, 2011.

[10] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12(3):193–226, 1986.

[11] Didier Dubois, Weiru Liu, Jianbing Ma, and Henri Prade. The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks. *Information Fusion*, 32:12–39, 2016.

[12] Didier Dubois and Henri Prade. Consonant approximations of belief functions. *International Journal of Approximate Reasoning*, 4(56):419 – 449, 1990.

[13] Didier Dubois and Henri Prade. Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge. In *Qualitative and quantitative practical reasoning*, pages 96–107. Springer, 1997.

[14] Didier Dubois, Henri Prade, and Philippe Smets. A definition of subjective possibility. *International Journal of Approximate Reasoning*, 48(2):352 – 364, 2008. In Memory of Philippe Smets (19382005).

[15] Didier Dubois, Henri Prade, and Philippe Smets. A definition of subjective possibility. *International Journal of Approximate Reasoning*, 48(2):352–364, 2008.

[16] Didier Dubois and Ronald R Yager. Fuzzy set connectives as combinations of belief structures. *Information Sciences*, 66(3):245–276, 1992.

[17] B Ducourthial and V Cherfaoui. Experiments with self-stabilizing distributed data fusion. In *Reliable Distributed Systems (SRDS), 2016 IEEE 35th Symposium on*, pages 289–296. IEEE, 2016.

[18] Anne-Laure Jousselme, Dominic Grenier, and Eloi Bossé. A new distance between two bodies of evidence. *Information Fusion*, 2:91–101, 2001.

[19] John Klein, Sebastien Destercke, and Olivier Colot. *Idempotent Conjunctive Combination of Belief Functions by Distance Minimization*, pages 156–163. Springer International Publishing, Cham, 2016.

[20] John Klein, Sebastien Destercke, and Olivier Colot. Interpreting evidential distances by connecting them to partial orders: Application to belief function approximation. *International Journal of Approximate Reasoning*, 71:15 – 33, 2016.

[21] Frédéric Pichon and Anne-Laure Jousselme. A norm-based view on conflict. In *Proceedings of the Rencontres Francophones sur la Logique Floue et ses Applications (LFA 2016)*, pages 153–159, La Rochelle, France, November 3-4 2016. Cépaduès.

[22] Glenn Shafer. *A mathematical theory of evidence.* Princeton University Press, 1976.

[23] Philippe Smets. The combination of evidence in the transferable belief model. *IEEE Transactions on On Pattern Analysis And Machine Intelligence.*, 11(5):447–458, 1990.

[24] Philippe Smets. The concept of distinct evidence. In *IPMU 92 Proceedings. pg*, pages 789–794, 1992.

[25] Philippe Smets. The canonical decomposition of a weighted belief. *14th international joint conference on Artificial intelligence*, 2:1896–1901, 1995.

[26] Philippe Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31:1–30, 2002.

[27] Philippe Smets. Analysing the combination of conflicting belief functions. *Information Fusion*, 8:387–412, 2007.

[28] Philippe Smets and Robert Kennes. The transferable belief model. *Artificial Intelligence*, 66(2):191 – 234, 1994.

[29] L.A. Zadeh. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *the Artificial Intelligence Magazine*, 7:85–90, 1986.