On consistent approximations of belief functions in the mass space

Fabio Cuzzolin
fabio.cuzzolin@brookes.ac.uk

Oxford Brookes University, Oxford, UK

Abstract. In this paper we study the class of consistent belief functions, as counterparts of consistent knowledge bases in classical logic. We prove that such class can be defined univocally no matter our definition of proposition implied by a belief function. As consistency can be desirable in decision making, the problem of mapping an arbitrary belief function to a consistent one arises, and can be posed in a geometric setup. We analyze here all the consistent transformations induced by minimizing L_p distances between belief functions, represented by the vectors of their basic probabilities.

1 Introduction

Belief functions (b.f.s) [1, 2] are complex objects, in which different and sometimes contradictory bodies of evidence may coexist, as they mathematically describe the fusion of possibly conflicting expert opinions and/or imprecise/corrupted measurements, et caetera. Indeed, conflict and combinability play a central role in the theory of evidence [3–5], and have been recently subject to novel analyses [6–8]. As a consequence, making decisions based on such objects can be misleading. This is a well known problem in classical logics, where the application of inference rules to inconsistent sets of assumptions or "knowledge bases" may lead to incompatible conclusions, depending on the set of assumptions we start reasoning from [9]. A set of formulas Φ is said consistent iff there does not exist another formula ϕ such that Φ implies both ϕ and $\neg\phi$.

As each formula ϕ can be put in correspondence with the set $A(\phi)$ of interpretations under which it holds, a straightforward extension of classical logic consists on assigning a probability value to such sets of interpretations, i.e, to each formula. If all possible interpretations are collected in a frame of discernment, we can easily define a belief function on such a frame, and attribute to each formula ϕ a belief value $b(\phi) = b(A(\phi))$ through the associated set of interpretations $A(\phi)$. A belief function can therefore be seen, in this context, as the generalization of a knowledge base [10].

A variety of approaches have been proposed in the context of classical logics to solve the problem of inconsistent knowledge bases, such as fragmenting the latter into maximally consistent subsets, limiting the power of the formalism, or adopting non-classical semantics [11, 12]. Even when a knowledge base is formally inconsistent, though, it may contain potentially useful information. Paris [9], for instance, tackles the problem by not assuming each proposition in the knowledge base as a fact, but by attributing to it a certain degree of belief in a probabilistic logic approach. This leads to something similar to a belief function.

To identify the counterparts of consistent knowledge bases in the theory of evidence we need to specify the notion of a belief function "implying" a certain proposition $A(\phi)$. As we show in this paper, under two different sensible definitions of such implication, the class of belief functions which generalize consistent knowledge bases is uniquely determined as the set of BFs whose non-zero mass "focal elements" have non-empty intersection. We are therefore allowed to call them *consistent* belief functions (cs.b.f.s).
Analogously to consistent knowledge bases, consistent b.f.s are characterized by null internal conflict. It may therefore be desirable to transform a generic belief function into a consistent one prior to making a decision, or picking a course of action. A similar transformation problem has been widely studied in both the probabilistic [13, 14] and possibilistic [15] case. A sensible approach, in particular, consists on studying the geometry [16, 17] of the class of b.f.s of interest and project the original belief function onto the corresponding geometric locus.

In [18] the author has indeed investigated the consistent transformation problem in the space of belief functions, represented by the vectors of their belief values. This paper further extends this line of research. Its goals are two-fold: 1. to formalize the notion of consistent belief functions as counterparts of consistent knowledge bases in belief calculus, and 2. to study the consistent transformation problem in the mass space of the basis probability vectors. We therefore introduce in Section 2 the notion of consistent belief function and prove that they generalize consistent knowledge bases under two distinct definitions of implication. Section 3 illustrates the consistent transformation problem in geometric terms. Finally, in Section 4 we solve the approximation problem in the mass space, using the classical L_1 (4.1), L_∞ (4.2) and L_2 (4.3) norms to measure distances between mass vectors. The results are interpreted and compared with those obtained in the belief space.

2 Semantics of consistent belief functions

A *basic probability assignment* on a finite set *(frame of discernment [1])* Θ is a function $m_b : 2^\Theta \rightarrow [0,1]$ on $2^\Theta \doteq \{A \subseteq \Theta\}$ s.t. $m_b(\emptyset) = 0$, $\sum_{A \subseteq \Theta} m_b(A) = 1$, $m_b(A) \geq 0$ for all $A \subseteq \Theta$. Subsets of Θ associated with non-zero values of m_b are called *focal elements*, and their intersection *core*: $\mathcal{C}_b \doteq \bigcap_{A \subseteq \Theta : m_b(A) \neq 0} A$.
The *belief function* $b : 2^\Theta \rightarrow [0,1]$ associated with a basic probability assignment m_b on Θ is defined as: $b(A) = \sum_{B \subseteq A} m_b(B)$. A dual mathematical representation of the evidence encoded by a belief function b is the *plausibility function* (pl.f.) $pl_b : 2^\Theta \rightarrow [0,1]$, $A \mapsto pl_b(A)$ where the plausibility value $pl_b(A)$ of an event A is given by $pl_b(A) \doteq 1 - b(A^c) = 1 - \sum_{B \subseteq A^c} m_b(B) = \sum_{B \cap A \neq \emptyset} m_b(B) \geq b(A)$ and expresses the amount of evidence *not against* A.

Belief logic interpretation. Generalizations of classical logic in which propositions are assigned probability, rather than truth, values have been proposed in the past. As belief functions naturally generalize probability measures, it is quite natural to define non-classical logic frameworks in which propositions are assigned *belief values* instead. This approach has been brought forward in particular by Saffiotti [19], Haenni [10], and others.

In propositional logic, propositions or formulas are either true or false, i.e., their truth value is either 0 or 1 [20]. Formally, an *interpretation* or *model* of a formula ϕ is a valuation function mapping ϕ to the truth value "true" (1). Each formula can therefore be associated with the set of interpretations or models under which its truth value is 1. If we define the frame of discernment of all the possible interpretations, each formula ϕ is associated with the subset $A(\phi)$ of this frame which collects all its interpretations. If the available evidence allows to define a belief function on this frame of possible interpretations, to each formula $A(\phi) \subseteq \Theta$ it is then naturally assigned a degree of belief $b(A(\phi))$ between 0 and 1 [19, 10], measuring the total amount of evidence supporting the proposition "ϕ is true".

Consistent belief functions generalize consistent knowledge bases. In classical logic, a set Φ of formulas or "knowledge base" is said *consistent* if and only if there does not exist another formula ϕ such that the knowledge base implies both such formula and its negation: $\Phi \vdash \phi$, $\Phi \vdash \neg\phi$. In other words, it is impossible to derive incompatible conclusions from the set of propositions that form a consistent knowledge base. This is obviously crucial if we want to derive univocal, non-contradictory conclusions from a given body of evidence.

A knowledge base in propositional logic $\Phi = \{\phi : T(\phi) = 1\}$ corresponds in a belief logic framework [19] to a belief function, i.e., a set of propositions together with their non-zero belief values: $b = \{A \subseteq \Theta : b(A) \neq 0\}$. Therefore, to determine what consistency amounts to in such a framework, we need to formalize the notion of proposition implied by a belief function. One option is to decide that $b \vdash B \subseteq \Theta$ if B is implied by all the propositions supported by b:

$$b \vdash B \Leftrightarrow A \subseteq B \quad \forall A : b(A) \neq 0. \tag{1}$$

Alternatively, we could require the proposition B itself to receive non-zero support by the belief function b:

$$b \vdash B \Leftrightarrow b(B) \neq 0. \tag{2}$$

In both cases we can define the class of consistent belief functions as the set of b.f.s which cannot imply contradictory propositions.

Definition 1. *A belief function b is consistent if there exists no proposition A such that both A and its negation A^c are implied by b.*

When adopting the implication relation (1), it is easy to see that $A \subseteq B \ \forall A : b(A) \neq 0$ is equivalent to $\bigcap_{b(A)\neq 0} A \subseteq B$. Furthermore, as each proposition with non-zero belief value must by definition contain a focal element C s.t. $m_b(C) \neq 0$, the intersection of all non-zero belief propositions reduces to that of all focal elements of b, i.e., the core of b: $\bigcap_{b(A)\neq 0} A = \bigcap_{\exists C\subseteq A: m_b(C)\neq 0} A = \bigcap_{m_b(C)\neq 0} C = \mathcal{C}_b$. Indeed, regardless the chosen definition of implication, the class of consistent belief functions corresponds to the set of b.f.s whose core is not empty.

Definition 2. *A belief function is said to be* consistent *if its core is non-empty.*

We can prove that, under either definition (1) or definition (2) of the implication $b \vdash B$, Definitions 1 and 2 are equivalent.

Theorem 1. *A belief function $b : 2^\Theta \rightarrow [0,1]$ has non-empty core if and only if there do not exist two complementary propositions $A, A^c \subseteq \Theta$ which are both implied by b in the sense (1).*

Proof. We have seen above that a proposition A is implied (1) by b iff $\mathcal{C}_b \subseteq A$. Accordingly, in order for both A and A^c to be implied by b we would need $\mathcal{C}_b = \emptyset$.

Theorem 2. *A b.f. $b : 2^\Theta \rightarrow [0,1]$ has non-empty core if and only if there do not exist two complementary propositions $A, A^c \subseteq \Theta$ which both enjoy non-zero support from b, $b(A) \neq 0$, $b(A^c) \neq 0$ (i.e., they are implied by b in the sense (2)).*

Proof. By Definition 1, in order for a subset (or proposition, in a propositional logic interpretation) $A \subseteq \Theta$ to have non-zero belief value it has to contain the core of b: $A \supseteq \mathcal{C}_b$. In order to have both $b(A) \neq 0$, $b(A^c) \neq 0$ we need both to contain the core, but in that case $A \cap A^c \supseteq \mathcal{C}_b \neq \emptyset$ which is absurd as $A \cap A^c = \emptyset$.

It is worth noticing, however, that other authors have introduced a slightly different notion of consistency, by requiring simply that the mass of the empty set be null $m_b(\emptyset) = 0$, or equivalently that the core of probabilities dominating b be non-empty [21, 22].

3 The L_p consistent approximation problem

The amount of *internal conflict* of a b.f. is typically defined as $c(b) \doteq 1 - \max_{x\in\Theta} pl_b(x)$ or, alternatively, $c(b) \doteq \sum_{A,B\subseteq\Theta: A\cap B=\emptyset} m_b(A)m_b(B)$. In both cases a belief function b is consistent if and only if its internal conflict is zero, $c(b) = 0$. As cs.b.f.s are the belief logic equivalent of consistent knowledge bases, it can be considered desirable to transform a generic belief function to a consistent one prior to drawing conclusions on the phenomenon at hand.

Consistent transformation. Consistent transformations can be built by solving a minimization problem of the form

$$cs[b] = \arg\min_{cs\in\mathcal{CS}} dist(b, cs). \quad (3)$$

where *dist* is some distance measure between belief functions, and $\mathcal{CS}$ denotes the collection of all consistent b.f.s. We call (3) the *consistent approximation problem.* Plugging in different distance functions in (3) we get different consistent transformations. In [18] we have studied transformations induced by norms of vectors of belief values $\boldsymbol{b}$ in the belief space $\mathcal{B}$. Similarly, we can measure distances between b.f.s via geometric norms between vectors of *mass values*. Here we focus in particular on what happens when using L_p norms in the space of basic probability assignments. This is supported by the fact that the contour function $pl_b(x)$ of a consistent belief function is a possibility distribution, which is in turn related to the L_p norm via $Pos(A) = \max_{x \in A} Pos(x)$. Note, however, that the plausibility measure of a b.f. is a necessity measure iff b is consonant, i.e., its focal elements are nested.

Mass space representation. To solve the consistent approximation problem (3) we need to understand the structure of the space in which consistent belief functions live. Each belief function is uniquely associated with the related set of mass values $\{m(A), \emptyset \subsetneq A \subseteq \Theta\}$ and can therefore be seen as a point of $\mathbb{R}^{N-2}$, $N = |2^{\Theta}|$, the vector $\boldsymbol{m}$ of its $N-1$ mass components minus the mass of Θ which is univocally determined by the normalization constraint:

$$\boldsymbol{m} = \sum_{\emptyset \subsetneq B \subsetneq \Theta} m_b(B) \boldsymbol{m}_B, \tag{4}$$

where $\boldsymbol{m}_B$ is the vector of mass values associated with the ("categorical") mass function $\boldsymbol{m}_A$ assigning all the mass to a single event A: $\boldsymbol{m}_A(A) = 1$, $\boldsymbol{m}_A(B) = 0$ $\forall B \neq A$. The collection $\mathcal{M}$ of points of $\mathbb{R}^{N-2}$ which are valid basic probability

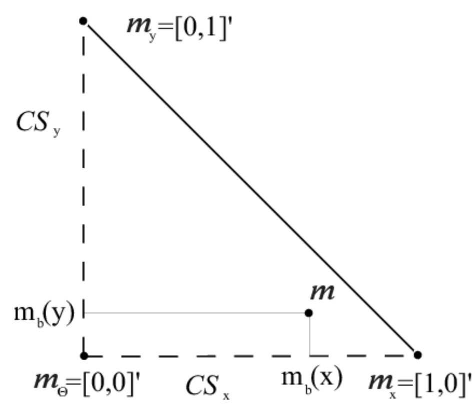

Fig. 1. The mass space $\mathcal{M}$ for a binary frame is a triangle of $\mathbb{R}^2$ whose vertices are the mass vectors associated with the categorical b.f.s focused on $\{x\}$, $\{y\}$ and Θ. Consistent b.f.s live in the union of the two segments $\mathcal{CS}_x = Cl(\boldsymbol{m}_\Theta, \boldsymbol{m}_x)$ and $\mathcal{CS}_y = Cl(\boldsymbol{m}_\Theta, \boldsymbol{m}_y)$.

assignments is a "simplex" (in rough words a higher-dimensional triangle), which we call *mass space*. $\mathcal{M}$ is the convex closure[1] $\mathcal{M} = Cl(\boldsymbol{m}_A, \emptyset \subsetneq A \subseteq \Theta)$.

Binary case. As an example let us consider a frame of discernment formed by just two elements, $\Theta_2 = \{x, y\}$. In this very simple case each belief function $b : 2^{\Theta_2} \rightarrow [0,1]$ is completely determined by its mass values $m_b(x)$, $m_b(y)$ as

[1] Here Cl denotes the convex closure operator: $Cl(\boldsymbol{m}_1, ..., \boldsymbol{m}_k) = \{\boldsymbol{m} \in \mathcal{M} : \boldsymbol{m} = \alpha_1 \boldsymbol{m}_1 + \cdots + \alpha_k \boldsymbol{m}_k, \sum_i \alpha_i = 1,\ \alpha_i \geq 0\ \forall i\}$.

$m_b(\Theta) = 1 - m_b(x) - m_b(y)$, $m_b(\emptyset) = 0 \ \forall b$. We can then represent each b.f. b as the vector of its basic probabilities (masses) $\boldsymbol{m} = [m_b(x), m_b(y)]'$ of $\mathbb{R}^{N-2} = \mathbb{R}^2$ (since $N = 2^2 = 4$). Since $m_b(x) \geq 0, \ m_b(y) \geq 0, \ m_b(x) + m_b(y) \leq 1$ the set $\mathcal{M}_2$ of all the possible belief functions on Θ_2 is the triangle of Figure 1, whose vertices are the points $\boldsymbol{m}_\Theta = [0,0]'$, $\boldsymbol{m}_x = [1,0]'$, $\boldsymbol{m}_y = [0,1]'$ which correspond respectively to the vacuous belief function b_Θ ($m_{b_\Theta}(\Theta) = 1$), the Bayesian b.f. b_x with $m_{b_x}(x) = 1$, and the Bayesian b.f. b_y such that $m_{b_y}(y) = 1$.
In the binary case consistent belief functions can have as list of focal elements either $\{\{x\}, \Theta_2\}$ or $\{\{y\}, \Theta_2\}$. Therefore the space of cs.b.f.s $\mathcal{CS}_2$ is the union of two line segments: $\mathcal{CS}_2 = \mathcal{CS}_x \cup \mathcal{CS}_y = Cl(\boldsymbol{m}_\Theta, \boldsymbol{m}_x) \cup Cl(\boldsymbol{m}_\Theta, \boldsymbol{m}_y)$.

The consistent complex. In the general case the geometry of consistent belief functions can be described by resorting to the notion of *simplicial complex* [23]. A simplicial complex is a collection Σ of simplices of arbitrary dimensions possessing the following properties: 1. if a simplex belongs to Σ, then all its faces of any dimension belong to Σ; 2. the intersection of any two simplices is a face of both the intersecting simplices. It has been proven that [24, 18] the region $\mathcal{CS}$ of consistent belief functions in the belief space is a simplicial complex, the union $\mathcal{CS}_\mathcal{B} = \bigcup_{x \in \Theta} Cl(\boldsymbol{b}_A, A \ni x)$. It is not difficult to see that the same holds in the mass space, where the consistent complex is the union $\mathcal{CS} = \bigcup_{x \in \Theta} Cl(\boldsymbol{m}_A, A \ni x)$ of maximal simplices $Cl(\boldsymbol{m}_A, A \ni x)$ formed by the mass vectors associated with all the belief functions with core containing a particular element x of Θ.

Why use L_p norms. A close relation exists between consistent belief functions and L_p norms, in particular the L_∞ one. As the plausibility of all the elements in their core is 1 ($pl_b(x) = \sum_{A \supseteq \{x\}} m_b(A) = 1 \ \forall x \in \mathcal{C}_b$), the region of consistent b.f.s can be expressed as $\mathcal{CS} = \big\{ b : \max_{x \in \Theta} pl_b(x) = 1 \big\} = \big\{ b : \|\bar{pl}_b\|_{L_\infty} = 1 \big\}$, i.e., the set of b.f.s for which the L_∞ norm of the *plausibility distribution* $\bar{pl}_b(x) = pl_b(\{x\})$ is equal to 1. This argument is strengthened by the observation that cs.b.f.s relate to possibility distributions, and possibility measures Pos are inherently related to L_∞ as $Pos(A) = \max_{x \in A} Pos(x)$. It makes therefore sense to conjecture that a consistent transformation obtained by picking as distance function in the approximation problem (3) one of the classical L_p norms

$$\|\boldsymbol{m} - \boldsymbol{m}'\|_{L_1} = \sum_{A \subseteq \Theta} |m_b(A) - m'_b(A)|, \|\boldsymbol{m} - \boldsymbol{m}'\|_{L_2} = \sqrt{\sum_{A \subseteq \Theta} (m_b(A) - m'_b(A))^2}, \\ \|\boldsymbol{m} - \boldsymbol{m}'\|_{L_\infty} = \max_{A \subseteq \Theta} \{|m_b(A) - m'_b(A)|\} \tag{5}$$

would be meaningful. In the probabilistic case, $p[b] = \arg\min_{p \in \mathcal{P}} dist(b, p)$, the use of L_p norms leads indeed to quite interesting results. On one side, the L_2 approximation induces the so-called "orthogonal projection" of b onto $\mathcal{P}$ [14]. On the other, the set of L_1/L_∞ probabilistic approximations of b (in the belief space) coincides with the set of probabilities dominating b: $\{p : p(A) \geq b(A)\}$ (at least in the binary case).

Distance of a point from a simplicial complex. As the consistent complex $\mathcal{CS}$ is a collection of linear spaces (better, simplices which generate a linear

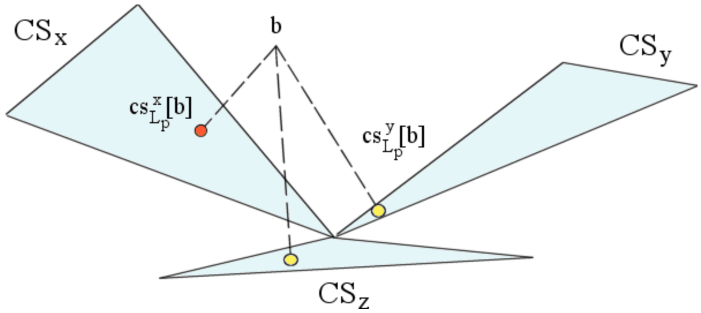

Fig. 2. To minimize the distance of a point from a simplicial complex, we need to find all the partial solutions on all the maximal simplices in the complex (empty circles), and compare these partial solutions to select a global optimum (black circle).

space) solving the problem (3) involves finding a number of partial solutions: $cs^x_{L_p}[b] = \arg\min_{\boldsymbol{cs}\in\mathcal{CS}_x} \|\boldsymbol{m}-\boldsymbol{cs}\|_{L_p}$ (see Figure 2). Afterwards, the distance of b from all such partial solutions has to be assessed in order to select a global optimal approximation. We will apply here this scheme to the approximation problems associated with the L_1, L_2, and L_∞ norms in the mass space.

4 Consistent approximation in $\mathcal{M}$

Using the notation $\boldsymbol{cs} = \sum_{B\supseteq\{x\},B\neq\Theta} m_{cs}(B)\boldsymbol{m}_B$, $\boldsymbol{m} = \sum_{B\subsetneq\Theta} m_b(B)\boldsymbol{m}_B$ (as in $\mathbb{R}^{N-2}$ $m_b(\Theta)$ is not included by normalization) the difference vector is

$$\boldsymbol{m}-\boldsymbol{cs} = \sum_{B\supseteq\{x\},B\neq\Theta} (m_b(B)-m_{cs}(B))\boldsymbol{m}_B + \sum_{B\not\supset\{x\}} m_b(B)\boldsymbol{m}_B \qquad (6)$$

so that its classical L_p norms read as

$$\begin{aligned}
\|\boldsymbol{m}-\boldsymbol{cs}\|_1^{\mathcal{M}} &= \sum_{B\supseteq\{x\},B\neq\Theta} |m_b(B)-m_{cs}(B)| + \sum_{B\not\supset\{x\}} |m_b(B)|,\\
\|\boldsymbol{m}-\boldsymbol{cs}\|_2^{\mathcal{M}} &= \sqrt{\sum_{B\supseteq\{x\},B\neq\Theta} |m_b(B)-m_{cs}(B)|^2 + \sum_{B\not\supset\{x\}} |m_b(B)|^2},\\
\|\boldsymbol{m}-\boldsymbol{cs}\|_\infty^{\mathcal{M}} &= \max\Big\{\max_{B\supseteq\{x\},B\neq\Theta} |m_b(B)-m_{cs}(B)|, \max_{B\not\supset\{x\}} |m_b(B)|\Big\}.
\end{aligned} \qquad (7)$$

4.1 L_1 approximation

Let us tackle first the L_1 case. After introducing the auxiliary variables $\beta(B) \doteq m_b(B) - m_{cs}(B)$ we can write the L_1 norm of the difference vector as

$$\|\boldsymbol{m}-\boldsymbol{cs}\|_1^{\mathcal{M}} = \sum_{B\supseteq\{x\},B\neq\Theta} |\beta(B)| + \sum_{B\not\supset\{x\}} |m_b(B)|, \qquad (8)$$

which is obviously minimized by $\beta(B) = 0$, for all $B \supseteq \{x\}$, $B \neq \Theta$. Therefore:

Theorem 3. *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$ and an element $x \in \Theta$ of the frame, its unique L_1 consistent approximation $cs^x_{L_1,\mathcal{M}}[b]$ in $\mathcal{M}$ with core containing x is the consistent b.f. whose mass distribution coincides with that of b on all the subsets containing x:*

$$m_{cs^x_{L_1,\mathcal{M}}[b]}(B) = \begin{cases} m_b(B) & \forall B \supseteq \{x\}, B \neq \Theta \\ m_b(\Theta) + b(\{x\}^c) & B = \Theta. \end{cases} \tag{9}$$

The mass value for $B = \Theta$ comes from normalization.
The mass of all the subsets not in the desired "principal ultrafilter" $\{B \supseteq \{x\}\}$ is simply re-assigned to Θ. A similarity emerges with the case of L_1 conditional belief functions [25], when we recall that the set of L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is the simplex whose vertices are each associated with a subset $\emptyset \subsetneq B \subseteq A$ of the conditional event A, and have b.p.a.:

$$m'(B) = m_b(B) + 1 - b(A), \; m'(X) = m_b(X) \quad \forall \emptyset \subsetneq X \subsetneq A, X \neq B.$$

In the L_1 conditional case, each vertex of the set of solutions is obtained by re-assigning the mass *not in the conditional event* A to a single subset of A, just as in L_1 consistent approximation all the mass *not in the principal ultrafilter* $\{B \supseteq \{x\}\}$ is re-assigned to the top of the ultrafilter, Θ.

Global approximation. The global L_1 consistent approximation in $\mathcal{M}$ is the partial approximation (9) at minimal distance from the original b.p.a. $\boldsymbol{m}$. By (8) the partial approximation focussed on x has distance $b(\{x\}^c) = \sum_{B \not\supset \{x\}} m_b(B)$ from $\boldsymbol{m}$. The global L_1 approximation $m_{cs_{L_1,\mathcal{M}}[b]}$ is therefore the partial approximation associated with the maximal plausibility singleton: $\hat{x} = \arg\min_x b(x^c) = \arg\max_x pl_b(x)$.

4.2 L_∞ approximation

In the L_∞ case $\|\boldsymbol{m} - \boldsymbol{cs}\|^{\mathcal{M}}_\infty = \max\Big\{ \max_{B \supseteq \{x\}, B \neq \Theta} |\beta(B)|, \max_{B \not\supset \{x\}} m_b(B) \Big\}$. The L_∞ norm of the difference vector is obviously minimized by $\{\beta(B)\}$ such that: $|\beta(B)| \leq \max_{B \not\supset \{x\}} m_b(B)$ for all $B \supseteq \{x\}, B \neq \Theta$, i.e., $-\max\limits_{B \not\supset \{x\}} m_b(B) \leq m_b(B) - m_c(B) \leq \max\limits_{B \not\supset \{x\}} m_b(B) \quad \forall\, B \supseteq \{x\}, B \neq \Theta$.

Theorem 4. *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$ and an element $x \in \Theta$ of the frame, its L_∞ consistent approximations $cs^x_{L_\infty,\mathcal{M}}[b]$ with core containing x in $\mathcal{M}$ are those whose mass values on all the subsets containing x differ from the original ones by the maximum mass of the subsets not in the ultrafilter: for all $B \supset \{x\}$, $B \neq \Theta$*

$$m_b(B) - \max_{C \not\supset \{x\}} m_b(C) \leq m_{cs^x_{L_\infty,\mathcal{M}}[b]}(B) \leq m_b(B) + \max_{C \not\supset \{x\}} m_b(C). \tag{10}$$

Clearly this set of solutions can include pseudo belief functions, i.e., b.f.s whose mass function is not necessarily non-negative.

Global approximation. Once again, the global L_∞ consistent approximation in $\mathcal{M}$ coincides with the partial approximation (10) at minimal distance from the original b.p.a. $\boldsymbol{m}$. The partial approximation focussed on x has distance $\max_{B \not\supset \{x\}} m_b(B)$ from $\boldsymbol{m}$. The global L_∞ approximation $m_{cs_{L_\infty,\mathcal{M}}[b]}$ is therefore the partial approximation associated with the singleton such that: $\hat{x} = \arg\min_x \max_{B \not\supset \{x\}} m_b(B)$.

4.3 L_2 approximation

To find the L_2 consistent approximation in $\mathcal{M}$ we need to minimize the L_2 norm of the difference vector $\|\boldsymbol{m} - \boldsymbol{cs}\|_2^{\mathcal{M}}$, or, equivalently, impose a condition of orthogonality between the difference vector itself $\boldsymbol{m} - \boldsymbol{cs}$ and the vector space associated with consistent mass functions focused on $\{x\}$. Clearly the generators of such linear space are the vectors in $\mathcal{M}$: $\boldsymbol{m}_B - \boldsymbol{m}_{\{x\}}$, for all $B \supsetneq \{x\}$. The desired orthogonality condition reads therefore as $\langle \boldsymbol{m} - \boldsymbol{cs}, \boldsymbol{m}_B - \boldsymbol{m}_{\{x\}} \rangle = 0$ where $\boldsymbol{m} - \boldsymbol{cs}$ is given by Equation (6), while $\boldsymbol{m}_B - \boldsymbol{m}_{\{x\}}(C) = 1$ if $C = B$, $= -1$ if $C = \{x\}$, 0 elsewhere. Therefore, using once again the variables $\{\beta(B)\}$, the condition simplifies as follows:

$$\langle \boldsymbol{m} - \boldsymbol{cs}, \boldsymbol{m}_B - \boldsymbol{m}_{\{x\}} \rangle = \begin{cases} \beta(B) - \beta(\{x\}) = 0 & \forall B \supsetneq \{x\}, B \neq \Theta; \\ -\beta(x) = 0 & B = \Theta. \end{cases} \quad (11)$$

Notice that, when using vectors $\boldsymbol{m}$ of $\mathbb{R}^{N-1}$ (including $B = \Theta$) to represent b.f.s, the orthogonality condition reads instead as:

$$\langle \boldsymbol{m} - \boldsymbol{cs}, \boldsymbol{m}_B - \boldsymbol{m}_{\{x\}} \rangle = \beta(B) - \beta(\{x\}) = 0 \quad \forall\, B \supsetneq \{x\}. \quad (12)$$

Theorem 5. *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$ and an element $x \in \Theta$ of the frame, its unique L_2 partial consistent approximation $cs^x_{L_2,\mathcal{M}}[b]$ with core containing x in $\mathcal{M}$ coincides with its partial L_1 approximation $cs^x_{L_1,\mathcal{M}}[b]$. However, when using the mass representation in $\mathbb{R}^{N-1}$, the partial L_2 approximation is obtained by equally redistributing to each element of the ultrafilter $\{B \supseteq \{x\}\}$ an equal fraction of the mass of focal elements not in it:*

$$m_{cs^x_{L_2,\mathcal{M}}[m_b]}(B) = m_b(B) + \frac{b(\{x\}^c)}{2^{|\Theta|-1}} \quad \forall B \supseteq \{x\}. \quad (13)$$

Proof. In the $N-2$ representation, by (11) we have that $\beta(B) = 0$, i.e., $m_{cs}(B) = m_b(B)\ \forall B \supseteq \{x\}$, $B \neq \Theta$. By normalization we get $m_{cs}(\Theta) = m_b(\Theta) + m_b(x^c)$: but this is exactly the L_1 approximation (9).
In the $N-1$ representation, by (12) we have that $m_{cs}(B) = m_{cs}(x) + m_b(B) - m_b(x)$ for all $B \supsetneq \{x\}$. By normalizing we get $\sum_{\{x\} \subseteq B \subseteq \Theta} m_{cs}(B) = m_{cs}(x) + \sum_{\{x\} \subsetneq B \subseteq \Theta} m_{cs}(B) = 2^{|\Theta|-1} m_{cs}(x) + pl_b(x) - 2^{|\Theta|-1} m_b(x) = 1$, i.e., $m_{cs}(x) = m_b(x) + (1 - pl_b(x))/2^{|\Theta|-1}$, as there are $2^{|\Theta|-1}$ subsets in the ultrafilter containing x. By replacing the value of $m_{cs}(x)$ into the first equation we get (13). The partial L_2 approximation in $\mathbb{R}^{N-1}$ redistributes the mass equally to all the elements of the ultrafilter.

Global approximation. The global L_2 consistent approximation in $\mathcal{M}$ is again given by the partial approximation (13) at minimal L_2 distance from $\boldsymbol{m}_b$. In the $N-2$ representation, by definition of L_2 norm in $\mathcal{M}$ (7), the partial approximation focussed on x has distance from $\boldsymbol{m}_b$: $(b(x^c))^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 = \big(\sum_{B \not\supset \{x\}} m_b(B)\big)^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2$. The latter is minimized by the element(s) $\hat{x} \in \Theta$ such that $\hat{x} = \arg\min_x \sum_{B \not\supset \{x\}} (m_b(B))^2$, which in turn determines the global L_2 approximation(s). In the $N-1$-dimensional case, instead, we get

$$\sum_{B \supseteq \{x\}, B \neq \Theta} \Big[m_b(B) - \Big(m_b(B) + \frac{b(x^c)}{2^{|\Theta|-1}} \Big) \Big]^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 =$$
$$\sum_{B \supseteq \{x\}, B \neq \Theta} \Big(\frac{b(x^c)}{2^{|\Theta|-1}} \Big)^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 = \frac{(\sum_{B \not\supset \{x\}} m_b(B))^2}{2^{|\Theta|-1}} + \sum_{B \not\supset \{x\}} (m_b(B))^2$$

which is minimized by the same singleton(s). In any case, even though (in the $N-2$ representation) the partial L_1 and L_2 approximations coincide, the global approximations in general may fall on different components of the complex.

5 Comparison with approximation in the belief space

It is interesting to compare the above results with those obtained in the belief space [18]. The partial L_1/L_2 consistent approximations of b (in the belief space $\mathcal{B}$) focused on a given element x coincide, and have b.p.a.:

$$m_{cs^x_{L_1}}(A) = m_{cs^x_{L_2}}(A) = m_b(A) + m_b(A \setminus \{x\}) \tag{14}$$

for all events A such that $\{x\} \subseteq A \subsetneq \Theta$. To get a consistent b.f. focused on a singleton x, the mass contribution of all the events B such that $B \cup \{x\} = A$ coincide is assigned indeed to A. They coincide with Dubois and Prade's "focused consistent transformations" [15].

It can be useful to illustrate the different approximations in the toy case of a ternary frame, $\Theta = \{x, y, z\}$. Assuming we want the consistent approximation to focus on x, by (14) the partial L_1/L_2 approximations in $\mathcal{B}$ are given by (using the simplified notation $m'(A)$): $m'(x) = m_b(x)$, $m'(x, y) = m_b(y) + m_b(x, y)$, $m'(x, z) = m_b(z) + m_b(x, z)$ $m'(\Theta) = 1 - b(x, y)$. The partial approximations induced by L_p norms in $\mathcal{M}$ can be computed via (9), (13) and (10). Figure 3 illustrates the different partial consistent approximations in the simplex $Cl(\boldsymbol{m}_x, \boldsymbol{m}_{x,y}, \boldsymbol{m}_{x,z}, \boldsymbol{m}_\Theta)$ of consistent belief functions focussed on x in a ternary frame, for the belief function with masses $m_b(x) = 0.2, m_b(y) = 0.1, m_b(z) = 0, m_b(x, y) = 0.4, m_b(x, z) = 0, m_b(y, z) = 0.3$. This is a tetrahedron with four vertices, delimited by dark solid edges.

The set of partial L_∞ approximations in $\mathcal{M}$ is depicted in Figure 3 as a green cube. As expected, it does not entirely fall inside the tetrahedron of admissible consistent belief functions. Its barycenter (the green star) coincides with

the L_1 partial consistent approximation in $\mathcal{M}$. The L_2^{N-2} approximation does also coincide, as expected, with the L_1 approximation. It remains to be seen if this holds for general frames of discernment as well. Regardless, there seems to exist a strong case for the latter transformation, which possesses a natural interpretation in terms of mass assignment: all the mass outside the ultrafilter is reassigned to Θ, increasing the overall uncertainty of the belief state.
The L_2 partial approximation in the $N-1$ representation (blue star) is distinct from the previous ones, but still falls inside the polytope of L_∞ partial approximations and is admissible, as it falls in the interior of the simplicial component $Cl(\boldsymbol{m}_x, \boldsymbol{m}_{x,y}, \boldsymbol{m}_{x,z}, \boldsymbol{m}_\Theta)$. Its interpretation is rather compelling, as it splits the mass not in the ultrafilter focused on x equally among all the subsets in the ultrafilter. Finally, the unique L_1/L_2 partial approximation in $\mathcal{B}$ is shown (red square). It has something in common with the $L_2^{N-2,M} = L_1^M = \overline{L_\infty}^M$ approximation (green star), as they both fall exactly on the border of admissible consistent b.f.s (the face highlighted in yellow): they assign zero mass to $\{x, z\}$, which fails to be supported by any focal element of the original belief function.

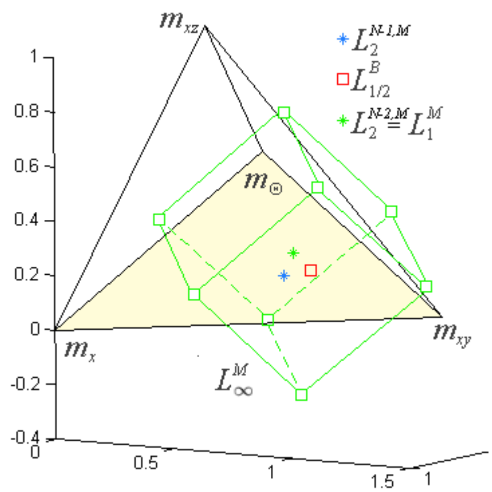

Fig. 3. The simplex (solid black tetrahedron) $Cl(\boldsymbol{m}_x, \boldsymbol{m}_{x,y}, \boldsymbol{m}_{x,z}, \boldsymbol{m}_\Theta)$ of consistent belief functions focussed on x in $\Theta = \{x, y, z\}$, and the related L_p partial consistent approximations of the b.f. with mass assignment (14).

6 Conclusions

In this paper we proved that consistent belief functions are the counterparts of consistent knowledge bases in belief calculus, analyzed consistent transformations induced by L_p norms in the mass space, and compared them with analogous transformations obtained in the belief space. The open-world scenario in which the current frame of discernment does not necessarily cover all possible alternatives, represented by the assumption $m_b(\emptyset)$ was not covered in this paper, but will be explored in the near future.

References

1. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
2. Dempster, A.P.: Upper and lower probabilities induced by a multivariate mapping. Annals of Mathematical Statistics **38** (1967) 325–339
3. Yager, R.R.: On the dempster-shafer framework and new combination rules. Information Sciences **41** (1987) 93–138
4. Smets, P.: The degree of belief in a fuzzy event. Inf. Sciences **25** (1981) 1–19
5. Ramer, A., Klir, G.J.: Measures of discord in the Dempster-Shafer theory. Information Sciences **67**(1-2) (1993) 35–50
6. Liu, W.: Analyzing the degree of conflict among belief functions. Artif. Intell. **170**(11) (2006) 909–924
7. Hunter, A., Liu, W.: Fusion rules for merging uncertain information. Information Fusion **7**(1) (2006) 97–134
8. Lo, K.C.: Agreement and stochastic independence of belief functions. Mathematical Social Sciences **51**(1) (2006) 1–22
9. Paris, J.B., Picado-Muino, D., Rosefield, M.: Information from inconsistent knowledge: A probability logic approach. In: Advances in Soft Computing. Volume 46. Springer-Verlag (2008) 291–307
10. Haenni, R.: Towards a unifying theory of logical and probabilistic reasoning. In: Proceedings of ISIPTA'05. (2005) 193–202
11. Priest, G., Routley, R., Norman, J.: Paraconsistent logic: Essays on the inconsistent. Philosophia Verlag (1989)
12. Batens, D., Mortensen, C., Priest, G.: Frontiers of paraconsistent logic. In: Studies in logic and computation. Volume 8. Research Studies Press (2000)
13. Daniel, M.: On transformations of belief functions to probabilities. International Journal of Intelligent Systems **21**(6) (2006) 261–282
14. Cuzzolin, F.: Two new Bayesian approximations of belief functions based on convex geometry. IEEE Tr. SMC-B **37**(4) (2007) 993–1008
15. Dubois, D., Prade, H.: Consonant approximations of belief functions. International Journal of Approximate Reasoning **4** (1990) 419–449
16. Black, P.: An examination of belief functions and other monotone capacities. PhD dissertation, Department of Statistics, Carnegie Mellon University (1996)
17. Cuzzolin, F.: A geometric approach to the theory of evidence. IEEE Transactions on Systems, Man and Cybernetics part C **38**(4) (2008) 522–534
18. Cuzzolin, F.: Consistent approximations of belief functions. In: Proceedings of ISIPTA'09. (2009) 139–148
19. Saffiotti, A.: A belief-function logic. In: Universit Libre de Bruxelles, (MIT Press) 642–647
20. Mates, B.: Elementary Logic. Oxford University Press (1972)
21. Cattaneo, M.E.G.V.: Combining belief functions issued from dependent sources. In: ISIPTA. (2003) 133–147
22. de Cooman, G.: Belief models: An order-theoretic investigation. Annals of Mathematics and Artificial Intelligence **45**(1-2) (2005) 5–34
23. Dubrovin, B.A., Novikov, S.P., Fomenko, A.T.: Sovremennaja geometrija. Metody i prilozenija. Nauka, Moscow (1986)
24. Cuzzolin, F.: An interpretation of consistent belief functions in terms of simplicial complexes. In: Proc. of ISAIM'08. (2008)
25. Cuzzolin, F.: Geometric conditioning of belief functions. In: Proceedings of BELIEF'10, Brest, France. (2010)