Consistent transformations of belief functions

Fabio Cuzzolin [a]

[a] *Department of Computing and Communication Technologies*
Oxford Brookes University
Wheatley campus, Oxford OX33 1HX, United Kingdom

Abstract

Consistent belief functions represent collections of coherent or non-contradictory pieces of evidence, but most of all they are the counterparts of consistent knowledge bases in belief calculus. The use of consistent transformations $cs[\cdot]$ in a reasoning process to guarantee coherence can therefore be desirable, and generalizes similar techniques in classical logics. Transformations can be obtained by minimizing an appropriate distance measure between the original belief function and the collection of consistent ones. We focus here on the case in which distances are measured using classical L_p norms, in both the "mass space" and the "belief space" representation of belief functions. While mass consistent approximations reassign the mass not focussed on a chosen element of the frame either to the whole frame or to all supersets of the element on an equal basis, approximations in the belief space do distinguish these focal elements according to the "focussed consistent transformation" principle. The different approximations are interpreted and compared, with the help of examples.

Key words: *Theory of evidence, belief logic, consistent belief functions, simplicial complex, L_p norms, consistent transformation.*

1 Introduction

Belief functions (b.f.s) [35,15] 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, etcetera. Indeed, conflict and combinability play a

Email address: fabio.cuzzolin@brookes.ac.uk (Fabio Cuzzolin).
URL: http://perception.inrialpes.fr/people/Cuzzolin (Fabio Cuzzolin).

central role in the theory of evidence [42,37,33], and have been recently subject to novel analyses [27,22,28]. As a consequence, making decisions based on such objects can be misleading.

Consistent knowledge bases and belief functions. 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 our reasoning from [30]. 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 [21].
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 [32,2]. Even when a knowledge base is formally inconsistent, though, it may contain potentially useful information. Paris [30], 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 probability 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 b.f.s whose non-zero mass "focal elements" have non-empty intersection. We are therefore allowed to call them *consistent* belief functions (cs.b.f.s).

Consistent transformation. Analogously to consistent knowledge bases, consistent b.f.s are characterized by null internal conflict. It may be therefore be desirable to transform a generic belief function to 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 [41,4,14,5] and possibilistic [17,18,1] cases. A sensible approach, in particular, consists on studying the geometry [3,6] of the class of b.f.s of interest and projecting the original belief function onto the corresponding geometric locus.
Indeed, consistent transformations can be built by solving a minimization

problem of the form

$$cs[b] = \arg\min_{cs \in \mathcal{CS}} dist(b, cs) \tag{1}$$

where b is the original belief function, *dist* an appropriate distance measure between belief functions, and $\mathcal{CS}$ denotes the collection of all consistent b.f.s.

By plugging in different distance functions in (1) we get different consistent transformations. Indeed, Jousselme et al [24] have recently conducted a very nice survey of the distance or similarity measures so far introduced in belief calculus, come out with an interesting classification, and proposed a number of generalizations of known measures. Many of these measures could be in principle employed to define indifferently consistent transformations or conditional belief functions [8], or approximate belief functions by necessity or probability measures. Other similarity measures between belief functions have been proposed by Shi et al [36], Jiang et al [23], and others [25,16].
In this paper we focus on what happens when employing classical L_p norms in the approximation problem. The L_∞ norm, in particular, is closely related to consistent and consonant belief functions. The region of consistent b.f.s can be expressed as

$$\mathcal{CS} = \Big\{ b : \max_{x \in \Theta} pl_b(x) = 1 \Big\},$$

i.e., the set of b.f.s for which the L_∞ norm of the "plausibility distribution" or "contour function" $pl_b(x)$ is equal to 1. In addition, cs.b.f.s relate to possibility distributions, and possibility measures Pos are inherently associated with L_∞ as $Pos(A) = \max_{x \in A} Pos(x)$. In recent times, L_p norms have been successfully employed in different problems such as probability [5] and possibility [11] transformation/approximation, or conditioning [8,10].

Geometric approximation on a complex. As the author has proven [7], geometrically, consistent (as well as consonant [9]) belief functions live in a collection of simplices or "simplicial complex" $\mathcal{CS}$. Each maximal simplex $\mathcal{CS}^x$ of the consistent complex is associated with associated with an "ultrafilter" $\{A \supseteq \{x\}\}$, $x \in \Theta$ of focal elements. A partial solution has therefore to be found separately for each maximal simplex of the consistent complex: these partial solutions are later compared to determine the global approximation(s). In addition, geometric approximation can be performed in different Cartesian spaces. A belief function can be represented either by the vector $\vec{b} = [b(A), \emptyset \subsetneq A \subsetneq \Theta]'$ of its belief values, or the vector of its mass values $\vec{m}_b = [m_b(A), \emptyset \subsetneq A \subsetneq \Theta]'$. We call the set of vectors of the first kind *belief space* $\mathcal{B}$ [6,13], and the collection of vectors of the second kind *mass space* $\mathcal{M}$ [8].

Limitations. In some cases, *improper* partial solutions (in the sense that they potentially include negative mass assignments) can be generated by the L_p minimization process. This situation is not entirely new. For instance, outer consonant approximations [17] also include infinitely many improper solutions: nevertheless, only the subset of acceptable solutions is retained. As in the consonant case [11], the set of all (admissible and not) solutions is typically much simpler to describe geometrically, in terms of simplices or polytopes. Computing the set of *proper* approximations in all cases requires significant further effort, which for reasons of clarity and length we reserve for the near future.

Additionally, in this work only "normalized" belief functions, i.e., b.f.s whose mass of the empty set is nil, are considered. Unnormalized b.f.s, however, play an important role in the TBM [38] as the mass of the empty set is an indicator of conflicting evidence. The analysis of the unnormalized case is also left to future work for lack of sufficient space here.

Contribution. In this paper, we solve the L_p consistent transformation problem in full generality in both the mass and the belief space, and discuss the semantics of the results.

In the **mass space** representation, the partial L_1 consistent approximation focussed on a certain element x of the frame is simply obtained by reassigning all the mass outside the ultrafilter $\{A \supseteq \{x\}\}$ to Θ. Global approximations are, as expected, associated with cores containing the maximal plausibility element(s) of Θ. The L_∞ approximation generates a "rectangle" of partial approximations, with barycenter in the L_1 partial approximation. The corresponding global approximations span the components focussed on the element(s) x such that $\max_{B \not\supset x} m_b(B)$ is minimal. The L_2 partial approximation coincides with the L_1 one if mass vectors include $m_b(\Theta)$ as a component. Otherwise the L_2 partial approximation reassigns the mass $b(x^c)$ originally outside the desired ultrafilter to each element of $\{A \supseteq \{x\}\}$ on an equal basis. The related global approximation is of more difficult interpretation.

In the **belief space** representation, partial approximations determined by both L_1 and L_2 norms are unique and coincide, besides having a rather elegant interpretation in terms of classical inner approximations [17,1]. The L_1/L_2 consistent approximation onto each component $\mathcal{CS}^x$ of $\mathcal{CS}$ generates indeed the *consistent transformation focused on* x, i.e. a new belief function whose focal elements have the form $A' = A \cup \{x\}$, where A is a focal element of the original b.f. b. The associated global L_1/L_2 solutions do not lie in general on the component of the consistent complex related to the maximal plausibility element.

The L_∞ norm determines instead an entire polytope of solutions whose barycenter lies on the L_1/L_2 approximation, and which is natural to associate with the polytope of inner Bayesian approximations. Global optimal L_∞ approximations do focus on the maximal plausibility element(s), and their center of mass is the consistent transformation focused on the latter.

Paper outline. After briefly recalling a few basis notions of the theory of evidence, we prove that consistent belief functions are the counterparts of consistent knowledge bases in belief logic (Section 2). As we investigate the transformation problem in a geometric framework, we briefly recall in Section 3 the geometry of the simplicial complex of consistent belief functions, and explain how we need to solve the approximation/transformation problem separately for each maximal simplex of this complex. We then proceed to solve the L_1-, L_2- and L_∞-consistent approximation problems in full generality, in both the mass (Section 4) and the belief (Section 5) space representations. In Section 6 we compare and interpret the outcomes of L_p approximations in the two frameworks, with the help of the ternary example. We conclude and discuss the natural prosecution of this line of research in Section 7.

2 Semantics of consistent belief functions

We first recall some basis notions of the theory of evidence, its belief logic interpretation, and the role of consistent belief functions there as counterparts of consistent knowledge bases.

2.1 Belief functions and belief logic

A *basic probability assignment* (b.p.a.) on a finite set or *frame of discernment* Θ is a set function $m_b : 2^\Theta \to [0,1]$ on $2^\Theta \doteq \{A \subseteq \Theta\}$ s.t. $m_b(\emptyset) = 0$, $\sum_{A\subseteq\Theta} m_b(A) = 1$. Subsets of Θ associated with non-zero values of m_b are called *focal elements* (f.e.), and their intersection *core*:

$$\mathcal{C}_b \doteq \bigcap_{A\subseteq\Theta : m_b(A)\neq 0} A.$$

The *belief function* (b.f.) $b : 2^\Theta \to [0,1]$ 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.) $pl_b : 2^\Theta \to [0,1]$, $A \mapsto pl_b(A)$ where the plausibility value $pl_b(A)$ of an event A is given by $pl_b(A) \doteq 1 - b(A^c) = \sum_{B\cap A\neq\emptyset} m_b(B)$ and expresses the amount of evidence *not against* A.

Generalizations of classical logic in which propositions are assigned probability 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*, rather than probability values. This approach has been brought forward in particular by Saffiotti [34], Haenni [21], and others.
In propositional logic, propositions or formulas are either true or false, i.e., their truth value is either 0 or 1 [29]. 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 $A(\phi) \subseteq \Theta$ is then naturally assigned a degree of belief $b(A(\phi))$ between 0 and 1 [34,21], measuring the total amount of evidence supporting the proposition "ϕ is true".

2.2 Consistent belief functions generalize consistent knowledge bases

In classical logic, a set Φ of formulas or "knowledge base" is said to be *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 which form a consistent knowledge base. This is 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 [34] to a belief function, i.e., a set of propositions together with their non-zero belief values: $b = \{A \subseteq \Theta : b(A) \neq 0\}$. 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:

$$b \vdash B \Leftrightarrow A \subseteq B \quad \forall A : b(A) \neq 0. \tag{2}$$

An alternative definition requires the proposition B itself to receive non-zero support by the belief function b:

$$b \vdash B \Leftrightarrow b(B) \neq 0. \tag{3}$$

Whatever definition we choose for such implication relation, 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 (2), 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, no matter our 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 (2) or definition (3) 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 (2).*

Proof. We have seen above that a proposition A is implied (2) 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 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 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 (3)).*

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 $A \cap A^c = \emptyset$. □

2.3 Achieving consistency in belief logic

Belief functions are complex objects, in which different and sometimes contradictory bodies of evidence coexist, as they may result from the fusion of possibly conflicting expert opinions and/or imprecise/corrupted measurements. It is reasonable to conjecture that taking belief functions at face value may lead in some situations to incorrect/inappropriate decisions.
A variety of approaches have been proposed in the context of classical logics to solve the analogous problem with inconsistent knowledge bases: we mentioned some of them in the introduction [32,2]. Paris' approach [30] is particularly interesting as it tackles the problem by attributing to each proposition in the knowledge base a certain degree of belief, leading to something similar to a belief function.

As consistent belief functions represent consistent knowledge bases in belief logic, such a mechanism can be be described there as an operator

$$cs : \mathcal{B} \rightarrow \mathcal{CS}, \quad b \mapsto cs[b] \tag{4}$$

where $\mathcal{B}$ and $\mathcal{CS}$ denote respectively the set of all belief functions, and that of all cs.b.f.s. Consistent transformations (4) can be built by posing a minimization problem of the form

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

where b is the belief function to approximate and $dist$ is some distance measure between b.f.s. It is natural to pose this problem in a geometric setup.

3 The L_p consistent approximation problem

3.1 Belief and mass vectors

As belief functions $b : 2^\Theta \rightarrow [0, 1]$, $b(A) = \sum_{B \subseteq A} m_b(B)$ are set functions defined on a the power set 2^Θ of a finite space Θ, they are obviously completely defined by the associate set of $2^{|\Theta|} - 2$ belief values, that we can collect in a vector $\vec{b} = [b(A), \emptyset \subsetneq A \subsetneq \Theta]'$ (since $b(\emptyset) = 0$, $b(\Theta) = 1$ for all b.f.s). They can therefore be represented as points of $\mathbb{R}^{N-2}$, $N = 2^{|\Theta|}$ [6]. The set $\mathcal{B}$ of points of $\mathbb{R}^{N-2}$ which correspond to belief functions is an N-dimensional triangle or *simplex* called *belief space* [6], namely: $\mathcal{B} = Cl(\vec{b}_A, \ \emptyset \subsetneq A \subseteq \Theta)$, where Cl denotes the convex closure operator

$$Cl(\vec{b}_1, ..., \vec{b}_k) = \left\{ \vec{b} \in \mathcal{B} : \vec{b} = \alpha_1 \vec{b}_1 + \cdots + \alpha_k \vec{b}_k, \sum_i \alpha_i = 1, \ \alpha_i \geq 0 \ \forall i \right\}$$

and $\vec{b}_A$ is the vector associated with the categorical [40] belief function b_A assigning all the mass to a single subset $A \subseteq \Theta$: $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0$ for all $B \neq A$. The vector $\vec{b} \in \mathcal{B}$ that corresponds to a belief function b has coordinates $m_b(A)$ in the simplex $\mathcal{B}$:

$$\vec{b} = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A) \vec{b}_A.$$

In the same way, each belief function is uniquely associated with the related set of mass values $\{m_b(A), \emptyset \subsetneq A \subseteq \Theta\}$ (Θ this time included). It can therefore be seen also as a point of $\mathbb{R}^{N-1}$, the vector $\vec{m}_b$ of its $N - 1$ mass components:

$$\vec{m}_b = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A) \vec{m}_A, \tag{5}$$

where $\vec{m}_A$ is the vector of mass values associated with the categorical b.f. b_A. The collection $\mathcal{M}$ of points which are valid basic probability assignments is also a simplex, which we can call *mass space*: $\mathcal{M} = Cl(\vec{m}_A, \emptyset \subsetneq A \subset \Theta)$.

3.2 Binary example

As an example let us consider a frame of discernment formed by just two elements, $\Theta_2 = \{x, y\}$. Each b.f. $b : 2^{\Theta_2} \to [0, 1]$ is completely determined by its mass/belief values $m_b(x) = b(x)$, $m_b(y) = b(y)$, as $m_b(\Theta) = 1 - m_b(x) - m_b(y)$ and $m_b(\emptyset) = 0$. We can therefore collect them in a vector of $\mathbb{R}^{N-2} = \mathbb{R}^2$ (since $N = 2^2 = 4$): $\vec{m}_b = \vec{b} = [m_b(x), m_b(y)]' = \vec{b} \in \mathbb{R}^2$. In this example mass space and belief space coincide.
Since $m_b(x) \geq 0$, $m_b(y) \geq 0$, and $m_b(x) + m_b(y) \leq 1$ we can easily infer that the set $\mathcal{B}_2 = \mathcal{M}_2$ of all the possible basic probability assignments (belief functions) on Θ_2 can be depicted as the triangle in the Cartesian plane of Figure 1, whose vertices are the points $\vec{b}_\Theta = \vec{m}_\Theta = [0, 0]'$, $\vec{b}_x = \vec{m}_x = [1, 0]'$, $\vec{b}_y = \vec{m}_y = [0, 1]'$, 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 with $m_{b_y}(y) = 1$. The region $\mathcal{P}_2$ of all Bayesian b.f.s on Θ_2 is the diagonal line segment $Cl(\vec{m}_x, \vec{m}_y) = Cl(\vec{b}_x, \vec{b}_y)$.
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: $\mathcal{CS}_2 = \mathcal{CS}^x \cup \mathcal{CS}^y = Cl(\vec{m}_\Theta, \vec{m}_x) \cup Cl(\vec{m}_\Theta, \vec{m}_y) = Cl(\vec{b}_\Theta, \vec{b}_x) \cup Cl(\vec{b}_\Theta, \vec{b}_y)$. It is easy to recognize that $\mathcal{CS}_2$ can be synthetically written in terms of the L_∞ (max) norm as

$$\mathcal{CS}_2 = \{b : \min\{b(x), b(y)\} = 0\} = \{b : \max\{pl_b(x), pl_b(y)\} = 1\}. \quad (6)$$

3.3 The consistent complex

In the general case the geometry of consistent belief functions can be described by resorting to the notion of *simplicial complex* [19]. 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 [7,12] the region $\mathcal{CS}$ of consistent belief functions in the belief space is a simplicial complex, the union

$$\mathcal{CS}_{\mathcal{B}} = \bigcup_{x \in \Theta} Cl(\vec{b}_A, A \ni x).$$

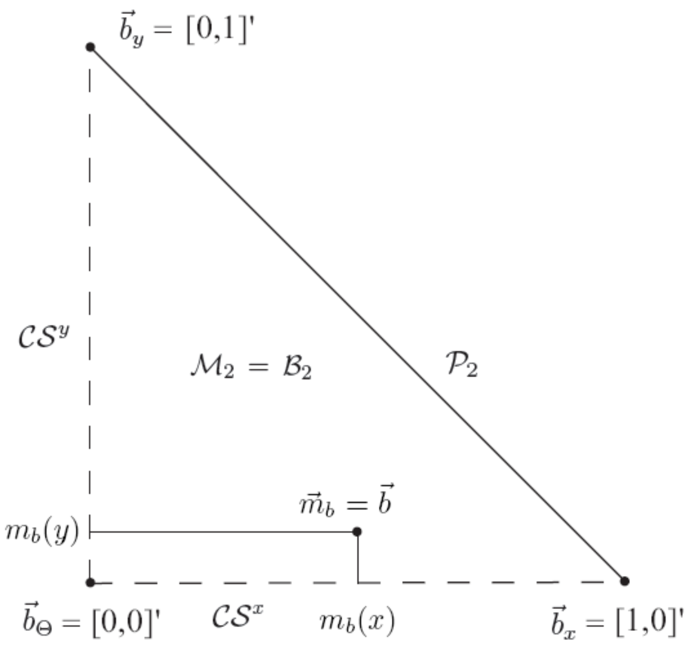

Fig. 1. Both the mass space $\mathcal{M}_2$ and the belief space $\mathcal{B}_2$ for a binary frame $\Theta = \{x, y\}$ coincide with the triangle of $\mathbb{R}^2$ whose vertices are the mass (belief) vectors associated with the categorical b.f.s focused on $\{x\}, \{y\}$ and Θ, respectively. Consistent b.f.s live in the union of the two segments $\mathcal{CS}^x = Cl(\vec{m}_\Theta, \vec{m}_x) = Cl(\vec{b}_\Theta, \vec{b}_x)$ and $\mathcal{CS}^y = Cl(\vec{m}_\Theta, \vec{m}_y) = Cl(\vec{b}_\Theta, \vec{b}_y)$.

of a number of (maximal) simplices, each associated with a "maximal ultrafilter" $\{A \supseteq \{x\}\}$, $x \in \Theta$ of subsets of Θ (those containing a given element x). It is not difficult to see that the same holds in the mass space, where the consistent complex is the union

$$\mathcal{CS}_\mathcal{M} = \bigcup_{x \in \Theta} Cl(\vec{m}_A, A \ni x)$$

of maximal simplices $Cl(\vec{m}_A, A \ni x)$ formed by the mass vectors associated with all the belief functions with core containing a particular element x of Θ.

3.4 Using L_p norms

The geometry of the binary case hints to a close relationship between consistent belief functions and L_p norms, in particular the L_∞ one (Equation (6)). It is easy to realize that this holds in general as, since the plausibility of all elements of 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 $\mathcal{CS} = \left\{ b : \max_{x \in \Theta} pl_b(x) = 1 \right\}$, i.e., the set of b.f.s for which the L_∞ norm of the "contour function" $pl_b(x)$ 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 (1) one of the classical L_p norms maybe be meaningful. For vectors $\vec{m}_b$, $\vec{m}_{b'} \in \mathcal{M}$ representing the b.p.a.s of two belief functions b, b', such norms read as:

$$
\begin{array}{lll}
\|\vec{m}_b - \vec{m}_{b'}\|_{L_1} & \doteq & \sum_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B) - m_{b'}(B)|; \\
\|\vec{m}_b - \vec{m}_{b'}\|_{L_2} & \doteq & \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta} (m_b(B) - m_{b'}(B))^2}; \\
\|\vec{m}_b - \vec{m}_{b'}\|_{L_\infty} & \doteq & \max_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B) - m_{b'}(B)|,
\end{array}
\tag{7}
$$

while the same norms in the belief space read as:

$$
\begin{array}{c}
\|\vec{b} - \vec{b'}\|_{L_1} \doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |b(B) - b'(B)|; \; \|\vec{b} - \vec{b'}\|_{L_2} \doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta} (b(B) - b'(B))^2}; \\
\|\vec{b} - \vec{b'}\|_{L_\infty} \doteq \max_{\emptyset\subsetneq B\subseteq\Theta} |b(B) - b'(B)|.
\end{array}
\tag{8}
$$

In recent times, L_p norms have been employed in different problems such as probability [5] and possibility [11] transformation/approximation, or conditioning [8,10]. In the probability transformation problem [41,4,39], $p[b] = \arg\min_{p\in\mathcal{P}} dist(b,p)$, the use of L_p norms leads indeed to quite interesting results. The L_2 approximation produces the so-called "orthogonal projection" of b onto $\mathcal{P}$ [5]. In addition, the set of L_1/L_∞ probabilistic approximations of b coincide with the set of probabilities consistent with b: $\{p : p(A) \geq b(A)\}$ (at least in the binary case).
Consonant approximations of belief functions obtained by minimizing L_p distances in the mass space have simple interpretations in terms of redistribution of the mass outside a desired chain $A_1 \subset \cdots \subset A_n$, $|A_i| = i$ of focal elements to a single element of the chain, or all in equal terms [11]. Conditional b.f.s can also be defined by minimizing L_p distances from a "conditioning simplex" $\mathcal{B}_A = Cl(\vec{b}_B, \emptyset \subsetneq B \subseteq A) \Big/ \mathcal{M}_A = Cl(\vec{m}_B, \emptyset \subsetneq B \subseteq A)$ in the belief/mass space determined by the conditioning event A [8]. In the mass space, the obtained conditional b.f.s have natural interpretation in terms of Lewis' "general imaging" [26,20] applied to belief functions.

3.5 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 space), solving the problem (1) involves finding a number of partial solutions in the belief/mass space

$$cs^x_{\mathcal{B},L_p}[b] = \arg\min_{\vec{cs}\in\mathcal{CS}^x_{\mathcal{B}}} \|\vec{b}-\vec{cs}\|_{L_p}, \quad cs^x_{\mathcal{M},L_p}[m_b] = \arg\min_{\vec{m}_{cs}\in\mathcal{CS}^x_{\mathcal{M}}} \|\vec{m}-\vec{m}_{cs}\|_{L_p}, \quad (9)$$

respectively (see Figure 2-left). Then, the distance of b from all such partial solutions has to be assessed in order to select a global approximation.

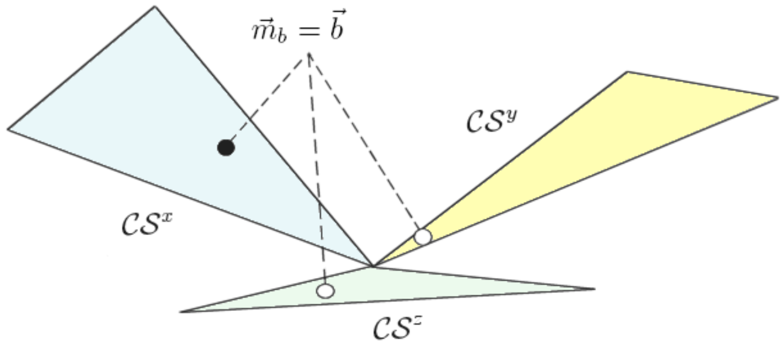

Fig. 2. Left: To minimize the distance of a point from a simplicial complex, we need to find all the partial solutions (9) on all the maximal simplices in the complex (empty circles), to later compare these partial solutions and select a global optimum (black circle).

4 Consistent approximation in $\mathcal{M}$

Let us therefore compute the analytical form of all L_p consistent approximations in the mass space. We start by describing the difference vector $\vec{m}_b - \vec{m}_{cs}$ between the original mass vector and its approximation. Using the notation

$$\vec{m}_{cs} = \sum_{B\supseteq\{x\},B\neq\Theta} m_{cs}(B)\vec{m}_B, \quad \vec{m}_b = \sum_{B\subsetneq\Theta} m_b(B)\vec{m}_B$$

(as in $\mathbb{R}^{N-2}$ $m_b(\Theta)$ is not included by normalization) the difference vector is

$$\vec{m}_b - \vec{m}_{cs} = \sum_{B\supseteq\{x\},B\neq\Theta} (m_b(B) - m_{cs}(B))\vec{m}_B + \sum_{B\not\supset\{x\}} m_b(B)\vec{m}_B \quad (10)$$

so that its classical L_p norms read as

$$\begin{aligned}
\|\vec{m}_b - \vec{m}_{cs}\|^{\mathcal{M}}_{L_1} &= \sum_{B\supseteq\{x\},B\neq\Theta} |m_b(B) - m_{cs}(B)| + \sum_{B\not\supset\{x\}} |m_b(B)|, \\
\|\vec{m}_b - \vec{m}_{cs}\|^{\mathcal{M}}_{L_2} &= \sqrt{\sum_{B\supseteq\{x\},B\neq\Theta} |m_b(B) - m_{cs}(B)|^2 + \sum_{B\not\supset\{x\}} |m_b(B)|^2}, \\
\|\vec{m}_b - \vec{m}_{cs}\|^{\mathcal{M}}_{L_\infty} &= \max\left\{ \max_{B\supseteq\{x\},B\neq\Theta} |m_b(B) - m_{cs}(B)|, \max_{B\not\supset\{x\}} |m_b(B)| \right\}.
\end{aligned} \quad (11)$$

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

$$\|\vec{m}_b - \vec{m}_{cs}\|_{L_1}^{\mathcal{M}} = \sum_{B \supseteq \{x\}, B \neq \Theta} |\beta(B)| + \sum_{B \not\supset \{x\}} |m_b(B)|, \tag{12}$$

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 partial L_1 consonant approximation $cs^x_{\mathcal{M},L_1}[m_b]$ in $\mathcal{M}$ with core containing x is the consonant b.f. whose mass distribution coincides with that of b on all the subsets containing x:*

$$m_{cs^x_{\mathcal{M},L_1}[m_b]}(B) = \begin{cases} m_b(B) & \forall B \supseteq \{x\}, B \neq \Theta \\ m_b(\Theta) + b(\{x\}^c) & B = \Theta. \end{cases} \tag{13}$$

The mass value for $B = \Theta$ comes from normalization, as follows:

$$m_{cs^x_{\mathcal{M},L_1}[m_b]}(\Theta) = 1 - \sum_{B \supseteq \{x\}, B \neq \Theta} m_{cs^x_{\mathcal{M},L_1}[m_b]}(B) = m_b(\Theta) + b(\{x\}^c).$$

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 [8], 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. (compare Equation (13)):

$$\begin{cases} m'(B) = m_b(B) + 1 - b(A), \\ m'(X) = m_b(X) & \forall \emptyset \subsetneq X \subsetneq A, X \neq B. \end{cases}$$

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.

4.1.1 Global approximation

The global L_1 consistent approximation in $\mathcal{M}$ coincides with the partial approximation (13) at minimal distance from the original mass vector $\vec{m}_b$. By (12) the partial approximation focussed on x has distance $b(\{x\}^c) = \sum_{B \not\supset \{x\}} m_b(B)$

from $\vec{m}_b$.
The global L_1 approximation(s) form therefore the union of the partial approximation(s) associated with the maximal plausibility singleton(s):

$$cs_{L_1,\mathcal{M}}[m_b] = \bigcup_{\arg\min_x b(x^c) = \arg\max_x pl_b(x)} cs^x_{\mathcal{M},L_1}[m_b]. \tag{14}$$

This is in accordance with our intuition, as it makes sense to focus on the singletons which are supported by the strongest evidence.

4.1.2 A running example

Consider as an example the belief function b with mass assignment

$$\begin{array}{c} m_b(x) = 0.2,\ m_b(y) = 0.1,\ m_b(z) = 0, \\ m_b(x,y) = 0.4,\ m_b(x,z) = 0,\ m_b(y,z) = 0.3,\ m_b(\Theta) = 0 \end{array} \tag{15}$$

on the ternary frame $\Theta = \{x, y, z\}$.
Suppose that we seek the L_1 approximation of b with core $\{x\}$. By Equation (13) we get the following consonant b.f.:

$$m_b'(x) = 0.2,\ m_b'(x,y) = 0.4,\ m_b'(x,z) = 0,\ m_b'(\Theta) = 0.4.$$

If, instead, we choose to focus on $\{y\}$ or $\{z\}$ we get, respectively (see Figure 3):

$$m_b'(y) = 0.1,\ m_b'(x,y) = 0.4,\ m_b'(y,z) = 0.3,\ m_b'(\Theta) = 0.2;$$
$$m_b'(z) = 0,\quad m_b'(x,z) = 0,\ m_b'(y,z) = 0.3,\ m_b'(\Theta) = 0.7.$$

Note that the partial approximation on $\mathcal{CS}^z$ has $\{y, z\}$ as actual core, and that

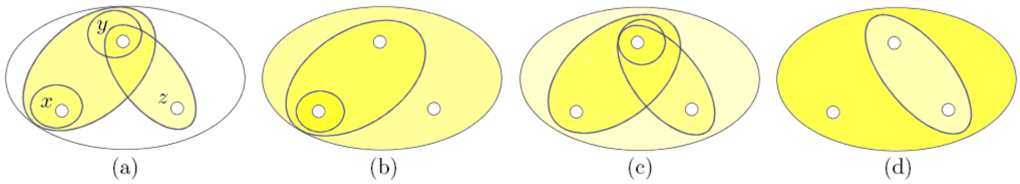

Fig. 3. (a) The belief function (15); (b) its partial L_1 consistent approximation in $\mathcal{M}$ focussed on $\{x\}$; (c) partial approximation on $\{y\}$; (d) partial approximation on $\{z\}$.

(b) and (d) are actually consonant (a special case of consistent b.f.), stressing the close relation between consonant and consistent approximation. Since $pl_b(x) = 0.6$, $pl_b(y) = 0.8$, $pl_b(z) = 0.3$, by (14) the global L_1 approximation of (15) is the partial approximation with core containing $\{y\}$. We can notice how the global approximation is visually much closer to the original b.f. than the partial ones, in terms of the structure of its focal elements.

4.2 L_∞ approximation

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

$$- \max_{C \not\supset \{x\}} m_b(C) \leq m_b(B) - m_{cs}(B) \leq \max_{C \not\supset \{x\}} m_b(C) \quad \forall\ B \supseteq \{x\}, B \neq \Theta.$$

Theorem 4 *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$ and an element $x \in \Theta$ of the frame, its partial L_∞ consistent approximations $cs^x_{\mathcal{M},L_\infty}[m_b]$ 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 $B \supset \{x\}$, $B \neq \Theta$*

$$m_b(B) - \max_{C \not\supset \{x\}} m_b(C) \leq m_{cs^x_{\mathcal{M},L_\infty}[m_b]}(B) \leq m_b(B) + \max_{C \not\supset \{x\}} m_b(C). \qquad (16)$$

Clearly this set of solutions can also include pseudo belief functions. Also, a comparison of Equations (16) and (13) shows that the barycenter of the partial L_∞ approximations coincides with the partial L_1 approximation, reassigning all the mass outside the ultrafilter to the whole frame.

4.2.1 Global approximation

Once again, the global L_∞ consistent approximation in $\mathcal{M}$ coincides with the partial approximation (16) at minimal distance from the original b.p.a. $\vec{m}_b$. The partial approximation focussed on x has distance $\max_{C \not\supset \{x\}} m_b(C)$ from $\vec{m}_b$. The global L_∞ approximation is therefore the (union of the) partial approximation(s) associated with the singleton(s) which minimize the maximal mass outside the ultrafilter:

$$cs_{\mathcal{M},L_\infty}[m_b] = \bigcup_{\arg\min_x \max_{C \not\supset \{x\}} m_b(C)} cs^x_{\mathcal{M},L_\infty}[m_b]. \qquad (17)$$

Global L_∞ solutions are not totally unrelated to their L_1 peers, as maximizing the plausibility of the core as in (14) involves minimizing the total mass outside the ultrafilter.

4.2.2 Running example

For the b.f. (15) of our running example, the maximal mass outside the ultrafilter in x is

$$\max_{C \not\supset x} m_b(C) = \max\Big\{m_b(y), m_b(z), m_b(y,z)\Big\} = m_b(y,z) = 0.3,$$

so that the set of L_∞ consistent approximations with core containing $\{x\}$ is such that

$$\begin{cases} m_b(x) - m_b(y,z) \leq m'_b(x) \leq m_b(x) + m_b(y,z) \\ m_b(x,y) - m_b(y,z) \leq m'_b(x,y) \leq m_b(x,y) + m_b(y,z) \\ m_b(x,z) - m_b(y,z) \leq m'_b(x,z) \leq m_b(x,z) + m_b(y,z) \end{cases} \equiv \begin{cases} -0.1 \leq m'_b(x) \leq 0.5 \\ 0.1 \leq m'_b(x,y) \leq 0.7 \\ -0.3 \leq m'_b(x,z) \leq 0.3. \end{cases} \tag{18}$$

This set is clearly not entirely admissible. The same can be verified for partial approximations with cores containing $\{y\}$ and $\{z\}$, for which

$$\max_{C \not\supset y} m_b(C) = m_b(x) = 0.2, \quad \max_{C \not\supset z} m_b(C) = m_b(x,y) = 0.4.$$

Therefore, the global L_∞ approximations of (15) are the partial ones whose core contains y (as it was the case for the global L_1 approximation).

4.3 L_2 approximation

In order to find the L_2 consistent approximation(s) in $\mathcal{M}$ it is convenient to recall that the minimal L_2 distance between a point and a vector space is attained by the point of the vector space V such that the difference vector is orthogonal to all the generators $\vec{g}_i$ of V:

$$\arg\min_{\vec{q} \in V} \|\vec{p} - \vec{q}\|_2 = \hat{q} \in V : \langle \vec{p} - \hat{q}, \vec{g}_i \rangle = 0 \quad \forall i$$

whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$.

Instead of minimizing the L_2 norm of the difference vector $\|\vec{m}_b - \vec{m}_{cs}\|^{\mathcal{M}}_{L_2}$ we can just impose the orthogonality of the difference vector itself $\vec{m}_b - \vec{m}_{cs}$ and the subspace $\mathcal{CS}^x_{\mathcal{M}}$ associated with consistent mass functions focused on $\{x\}$. Clearly the generators of such linear space are the vectors in $\mathcal{M}$: $\vec{m}_B - \vec{m}_{\{x\}}$, for all $B \supsetneq \{x\}$.

Theorem 5 *Consider an arbitrary belief function $b : 2^\Theta \rightarrow [0,1]$ and an element $x \in \Theta$ of the frame. When using the $(N-2)$-dimensional representation of mass vectors (5), its unique L_2 partial consistent approximation in $\mathcal{M}$ with core containing x coincides with its partial L_1 approximation:*

$cs^x_{\mathcal{M},L_2}[m_b] = cs^x_{\mathcal{M},L_1}[m_b]$. *When representing belief functions as mass vectors* $\vec{m}_b$ *of* $\mathbb{R}^{N-1}$ *(*$B = \Theta$ *included)*

$$\vec{m}_b = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B) \vec{m}_B \tag{19}$$

the partial L_2 *approximation of* b *is obtained by equally redistributing to each element of the ultrafilter* $\{B \supseteq \{x\}\}$ *an equal fraction of the mass of focal elements originally not in it:*

$$m_{cs^x_{\mathcal{M},L_2}[m_b]}(B) = m_b(B) + \frac{b(\{x\}^c)}{2^{|\Theta|-1}} \quad \forall B \supseteq \{x\}. \tag{20}$$

The partial L_2 approximation in $\mathbb{R}^{N-1}$ redistributes the mass equally to all the elements of the ultrafilter.

4.3.1 Global approximation

The global L_2 consistent approximation in $\mathcal{M}$ is, as usual, given by the partial approximation (20) at minimal L_2 distance from $\vec{m}_b$. In the $N-2$ representation, by definition of L_2 norm in $\mathcal{M}$ (11), the partial approximation focussed on x has distance from $\vec{m}_b$

$$(b(x^c))^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 = \Big(\sum_{B \not\supset \{x\}} m_b(B) \Big)^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2,$$

which is minimized by the singleton(s) $\arg\min_x \sum_{B \not\supset \{x\}} (m_b(B))^2$. Therefore:

$$cs_{\mathcal{M},L_2}[m_b] = \bigcup_{\arg\min_x \sum_{B \not\supset \{x\}} (m_b(B))^2} cs^x_{\mathcal{M},L_2}[m_b].$$

In the $N-1$-dimensional representation, instead,

$$\begin{array}{l} \displaystyle\sum_{B \supseteq \{x\}, B \neq \Theta} \Big[m_b(B) - \Big(m_b(B) + \frac{b(x^c)}{2^{|\Theta|-1}} \Big) \Big]^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 = \\ \displaystyle\sum_{B \supseteq \{x\}, B \neq \Theta} \Big(\frac{b(x^c)}{2^{|\Theta|-1}} \Big)^2 + \sum_{B \not\supset \{x\}} (m_b(B))^2 = \frac{(\sum_{B \not\supset \{x\}} m_b(B))^2}{2^{|\Theta|-1}} + \sum_{B \not\supset \{x\}} (m_b(B))^2 \end{array} \tag{21}$$

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 consonant complex.

4.3.2 Running example

For the usual example b.f. (15), by Equation (20) the L_2 partial approximations with cores containing x, y and x respectively have mass assignments:

$$\begin{aligned} &m_b'(x) = 0.3, && m_b'(x,y) = 0.5, && m_b'(x,z) = 0.1, && m_b'(\Theta) = 0.1 \\ &m_b'(y) = 0.15, && m_b'(x,y) = 0.45, && m_b'(y,z) = 0.35, && m_b'(\Theta) = 0.05; \\ &m_b'(z) = 0.175, && m_b'(x,z) = 0.175, && m_b'(y,z) = 0.475, && m_b'(\Theta) = 0.175, \end{aligned}$$

since $b(x^c) = 1 - pl_b(x) = 0.4$, $b(x^c) = 1 - pl_b(x) = 0.4$, $b(x^c) = 1 - pl_b(x) = 0.4$ and $2^{|\Theta|-1} = 4$ (see Figure 4).

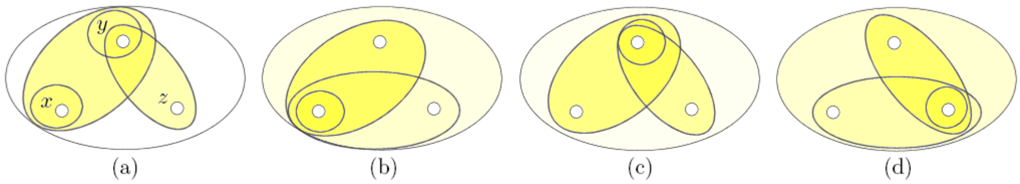

Fig. 4. (a) The belief function (15); (b) its partial L_2 consistent approximation in $\mathcal{M}$, focussed on $\{x\}$; (c) L_2 partial approximation on $\{y\}$; (d) L_2 partial approximation on $\{z\}$.

Note that, unlike the corresponding L_1 approximation, the partial L_2 approximation on $\{z\}$ has indeed the singleton as core. Also, as L_2 approximation redistributes mass evenly to all the focal elements in the desired ultrafilter, all partial approximations contain all possible f.e.s. As for the global solution

$$\begin{aligned} &(b(x^c))^2 = 0.16, \ \textstyle\sum_{B \not\supset x} (m_b(B))^2 = 0.1^2 + 0.3^2 = 0.01 + 0.09 = 0.1, \\ &(b(y^c))^2 = 0.04, \ \textstyle\sum_{B \not\supset y} (m_b(B))^2 = 0.2^2 = 0.04, \\ &(b(z^c))^2 = 0.49, \ \textstyle\sum_{B \not\supset z} (m_b(B))^2 = 0.2^2 + 0.1^2 + 0.4^2 = 0.04 + 0.01 + 0.16 = 0.21, \end{aligned}$$

and the L_2 norm (21) is minimized, once again, by the partial solution on $\{y\}$ (as visually confirmed by Figure 4).

4.4 Approximation in the mass space and general imaging in belief revision

As it is the case for geometric conditioning in the mass space [8], consistent approximations in the mass space can be interpreted as a generalization of Lewis' *imaging* approach to belief revision, originally formulated in the context of probabilities [26]. The idea behind imaging is that, upon observing that some state $x \in \Theta$ is impossible, you transfer the probability initially assigned to x completely towards the remaining state you deem the most similar to x [31]. Peter Gärdenfors [20] extended Lewis' idea by allowing a fraction λ_i of

the probability of such state x to be re-distributed to all remaining states x_i ($\sum_i \lambda_i = 1$).
In the case of partial consistent approximation of belief functions, the mass $m_b(A)$ of each focal element not in the desired ultrafilter $\{A \supseteq \{x\}\}$ should be re-assigned to the "closest" focal element in the latter. The partial L_1 consistent approximation in $\mathcal{M}$ amounts to admitting ignorance about the closeness of focal elements, and re-assigning all the mass outside the filter to the whole frame, therefore increasing the uncertainty of the state. If such ignorance is expressed by assigning instead equal weight $\lambda(A)$ to each $A \in \mathcal{C}$, the resulting partial consistent approximation is the unique partial L_2 approximation, the barycenter of the polytope of L_1 partial approximations.

5 Consistent approximation in the belief space

5.1 L_1/L_2 approximations

We have seen that in the mass space (at least in its $N-2$ representation, Theorem 5) the L_1 and L_2 approximations coincide. This is true in the belief space in the general case as well. We will gather some intuition on the general solution by considering first the slightly more complex case of a ternary frame: $\Theta = \{x, y, z\}$. In this Section we will use the notation:

$$\vec{cs} = \sum_{B \supseteq \{x\}} m_{cs}(B)\vec{b}_B, \quad \vec{b} = \sum_{B \subsetneq \Theta} m_b(B)\vec{b}_B.$$

In the case of an arbitrary frame a cs.b.f. $cs \in \mathcal{CS}^x_{\mathcal{B}}$ is a solution of the L_2 approximation problem if, again, $\vec{b} - \vec{cs}$ is orthogonal to all generators $\{\vec{b}_B - \vec{b}_\Theta = \vec{b}_B, \{x\} \subseteq B \subsetneq \Theta\}$ of $\mathcal{CS}^x_{\mathcal{B}}$: $\langle \vec{b} - \vec{cs}, \vec{b}_B \rangle = 0 \ \forall B : \{x\} \subseteq B \subsetneq \Theta$. As $\vec{b} - \vec{cs} = \sum_{A \subsetneq \Theta} (m_b(A) - m_{cs}(A))\vec{b}_A$ we get

$$\left\{ \sum_{A \supseteq \{x\}} \beta(A)\langle \vec{b}_A, \vec{b}_B \rangle + \sum_{A \not\supset \{x\}} m_b(A)\langle \vec{b}_A, \vec{b}_B \rangle = 0 \qquad \forall B : \{x\} \subseteq B \subsetneq \Theta, \right. \quad (22)$$

where once again $\beta(A) = m_b(A) - m_{cs}(A)$. In the L_1 case, the minimization problem is (using again the notation $\vec{cs} = \sum_{B \supseteq \{x\}} m_{cs}(B)\vec{b}_B$)

$$\begin{aligned} &\arg\min_{m_{cs}(.)} \left\{ \sum_{B \supseteq \{x\}} \left| \sum_{A \subseteq B} m_b(A) - \sum_{A \subseteq B, A \supseteq \{x\}} m_{cs}(A) \right| \right\} \\ &= \arg\min_{\beta(.)} \left\{ \sum_{B \supseteq \{x\}} \left| \sum_{A \subseteq B, A \supseteq \{x\}} \beta(A) + \sum_{A \subseteq B, A \not\supset \{x\}} m_b(A) \right| \right\} \end{aligned}$$

which is clearly solved by setting all addenda to zero, obtaining

$$\left\{ \sum_{A\subseteq B, A\supseteq\{x\}} \beta(A) + \sum_{A\subseteq B, A\not\supset\{x\}} m_b(A) = 0 \quad \forall B : \{x\} \subseteq B \subsetneq \Theta. \right. \tag{23}$$

An interesting fact emerges when comparing the linear systems for L_1 and L_2 in the ternary case $\Theta = \{x, y, x\}$:

$$\begin{cases} 3\beta(x) + \beta(x,y) + \beta(x,z) + m_b(y) + m_b(z) = 0 \\ \beta(x) + \beta(x,y) + m_b(y) = 0 \\ \beta(x) + \beta(x,z) + m_b(z) = 0 \end{cases} \quad \begin{cases} \beta(x) = 0 \\ \beta(x) + \beta(x,y) + m_b(y) = 0 \\ \beta(x) + \beta(x,z) + m_b(z) = 0. \end{cases} \tag{24}$$

The solution is the same for both, due to the fact that the second linear system is obtained from the first one by a linear transformation of rows. We just need to replace the first equation e_1 in the first system with the difference: $e_1 \mapsto e_1 - e_2 - e_3$. This holds in the general case, too.

Lemma 1 $\sum_{C\supseteq B} \langle \vec{b}_C, \vec{b}_A \rangle (-1)^{|C\setminus B|} = 1$ *whenever* $A \subseteq B$, *0 otherwise.*

Corollary 1 *The linear system (22) can be reduced to the system (23) through the following linear transformation of rows:*

$$row_B \mapsto \sum_{C\supseteq B} row_C (-1)^{|C\setminus B|}. \tag{25}$$

Proof. If we apply the linear transformation (25) to the system (22) we get

$$\begin{aligned} &\sum_{C\supseteq B} \left[\sum_{A\supseteq\{x\}} \beta(A) \langle \vec{b}_A, \vec{b}_C \rangle + \sum_{A\not\supset\{x\}} m_b(A) \langle \vec{b}_A, \vec{b}_C \rangle \right] (-1)^{|C\setminus B|} \\ &= \sum_{A\supseteq\{x\}} \beta(A) \sum_{C\supseteq B} \langle \vec{b}_A, \vec{b}_C \rangle (-1)^{|C\setminus B|} + \sum_{A\not\supset\{x\}} m_b(A) \sum_{C\supseteq B} \langle \vec{b}_A, \vec{b}_C \rangle (-1)^{|C\setminus B|} \end{aligned}$$

$\forall B : \{x\} \subseteq B \subsetneq \Theta$. Therefore by Lemma 1 we get

$$\left\{ \sum_{A\supseteq\{x\}, A\subseteq B} \beta(A) + \sum_{A\not\supset\{x\}, A\subseteq B} m_b(A) = 0 \quad \forall B : \{x\} \subseteq B \subsetneq \Theta, \right.$$

i.e., the system of equations (23). □

5.1.1 *Form of the solution*

To obtain both the L_2 and the L_1 consistent approximations of b it then suffices to solve the system (23) associated with the L_1 norm.

Theorem 6 *The unique solution of the linear system (23) is:* $\beta(A) = -m_b(A \setminus \{x\})$ *for all* $A : \{x\} \subseteq A \subsetneq \Theta$.

We can prove it by simple substitution, as system (23) becomes:

$$
\begin{aligned}
& - \sum_{B \subseteq A, B \supseteq \{x\}} m_b(B \setminus \{x\}) + \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \\
& = - \sum_{C \subseteq A \setminus \{x\}} m_b(C) + \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) = 0.
\end{aligned}
$$

Therefore, according to what discussed in Section 3, the partial consistent approximations of b on the maximal component $\mathcal{CS}^x_{\mathcal{B}}$ of the consistent complex have b.p.a.

$$
m_{cs^x_{\mathcal{B},L_1}[b]}(A) = m_{cs^x_{\mathcal{B},L_2}[b]}(A) = m_b(A) - \beta(A) = m_b(A) + m_b(A \setminus \{x\})
$$

for all events A such that $\{x\} \subseteq A \subsetneq \Theta$. As for the mass of Θ, by normalization: $m_{cs^x_{\mathcal{B},L_1/L_2}[b]}(\Theta) =$

$$
\begin{aligned}
& = 1 - \sum_{\{x\} \subseteq A \subsetneq \Theta} m_{cs^x_{\mathcal{B},L_1/L_2}[b]}(A) = 1 - \sum_{\{x\} \subseteq A \subsetneq \Theta} \Big(m_b(A) + m_b(A \setminus \{x\}) \Big) \\
& = 1 - \sum_{\{x\} \subseteq A \subsetneq \Theta} m_b(A) - \sum_{\{x\} \subseteq A \subsetneq \Theta} m_b(A \setminus \{x\}) \\
& = 1 - \sum_{A \neq \Theta, \{x\}^c} m_b(A) = m_b(\{x\}^c) + m_b(\Theta)
\end{aligned}
$$

as all events $B \not\supset \{x\}$ can be written as $B = A \setminus \{x\}$ for $A = B \cup \{x\}$.

Corollary 2

$$
m_{cs^x_{\mathcal{B},L_1}[b]}(A) = m_{cs^x_{\mathcal{B},L_2}[b]}(A) = m_b(A) + m_b(A \setminus \{x\})
$$

$\forall x \in \Theta$, *and for all* A *s.t.* $\{x\} \subseteq A \subseteq \Theta$.

5.1.2 *Interpretation as focused consistent transformations*

The expression of the the basic probability assignment of the L_1/L_2 consistent approximations of b (Corollary 2) is simple and elegant. It also has a straightforward interpretation: to get a consistent belief function focused on a singleton x, the mass contribution of all events B such that $B \cup \{x\} = A$

coincide is assigned indeed to A. But there are just two such events: A itself, and $A \setminus \{x\}$.
The partial consistent approximation of a belief function on a frame $\Theta = \{x, y, z, w\}$ with core $\{x\}$ is illustrated in Figure 5. The b.f. with focal ele-

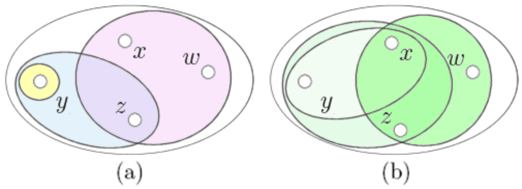

Fig. 5. A belief function (a) and its L_1/L_2 consistent approximation in $\mathcal{B}$ with core $\{x\}$ (b).

ments $\{y\}$, $\{y, z\}$, and $\{x, z, w\}$ is transformed by the map

$$\begin{aligned} \{y\} &\mapsto \{x\} \cup \{y\} = \{x, y\}, \\ \{y, z\} &\mapsto \{x\} \cup \{y, z\} = \{x, y, z\}, \\ \{x, z, w\} &\mapsto \{x\} \cup \{x, z, w\} = \{x, z, w\} \end{aligned}$$

into the consistent b.f. with focal elements $\{x, y\}$, $\{x, y, z\}$, and $\{x, z, w\}$ and the same b.p.a.
The partial solutions to the L_1/L_2 consistent approximation problem turn out to be related to the classical *inner consonant approximations* of a b.f. b, i.e., the set of consonant belief functions co such that

$$co(A) \geq b(A) \quad \forall A \subseteq \Theta$$

(or equivalently $pl_{co}(A) \leq pl_b(A)$ $\forall A$). Dubois and Prade [17] proved indeed that such an approximation exists iff b is consistent. However, when b is *not* consistent a "focused consistent transformation" can be applied to get a new belief function b' such that

$$m_{b'}(A \cup x_i) = m_b(A) \;\; \forall A \subseteq \Theta \tag{26}$$

and x_i is the element of Θ with highest plausibility.
Clearly then, the results of Theorem and Corollary say that the L_1/L_2 consistent approximation onto each component $\mathcal{CS}^x_{\mathcal{B}}$ of $\mathcal{CS}_{\mathcal{B}}$ generates the consistent transformation focused on x.

5.1.3 Global L_1 and L_2 approximations

Going back to the global approximation problem, to find the consistent approximation of b we need to work out which of the partial approximations $cs^x_{L_{1/2}}[b]$ has minimal distance from b.

Theorem 7 *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$, its global L_1 consistent approximation in the belief space is its partial approximation associated with the singleton:*

$$\arg\min_x \Big\{ \sum_{A \subseteq \{x\}^c} b(A), x \in \Theta \Big\}. \tag{27}$$

In the binary case ($\Theta = \{x, y\}$) the optimal singleton cores (27) simplify as:

$$\arg\min_x \sum_{A \subseteq \{x\}^c} b(A) = \arg\min_x m_b(\{x\}^c) = \arg\max_x pl_b(x),$$

and the global approximation falls on the component of the consistent complex associated with the element of *maximal plausibility*. Unfortunately, in the case of an arbitrary frame Θ the element (27) is not necessarily the maximal plausibility element:

$$\arg\min \Big\{ \sum_{A \subseteq \{x\}^c} b(A), x \in \Theta \Big\} \neq \arg\max \{pl_b(x), x \in \Theta \Big\},$$

as a simple counterexample can prove. As for the L_2 case:

Theorem 8 *Given an arbitrary belief function $b : 2^\Theta \to [0,1]$, its global L_2 consistent approximation in the belief space is its partial approximation associated with the singleton:*

$$\arg\min_x \Big\{ \sum_{A \subseteq \{x\}^c} \Big(b(A)\Big)^2, x \in \Theta \Big\}. \tag{28}$$

Once again, in the binary case the optimal singleton cores (28) specialize as:

$$\arg\min_x \sum_{A \subseteq \{x\}^c} (b(A))^2 = \arg\min_x (m_b(\{x\}^c))^2 = \arg\max_x pl_b(x)$$

and the global approximation for L_2 also falls on the component of the consistent complex associated with the element of maximal plausibility, while this is not generally true for an arbitrary frame.

5.1.4 Running example

According to Equation (26), for the belief function (15) of our running example the partial L_1/L_2 consistent approximations in $\mathcal{B}$ with core containing $\{x\}$,

$\{y\}$ and $\{z\}$, respectively, have mass assignments (see Figure 6):

$$\begin{array}{l} m'_b(x) = 0.2,\ m'_b(x,y) = 0.5,\ m'_b(x,z) = 0, \quad m'_b(\Theta) = 0.3; \\ m'_b(y) = 0.1,\ m'_b(x,y) = 0.6,\ m'_b(y,z) = 0.3,\ m'_b(\Theta) = 0; \\ m'_b(z) = 0, \quad m'_b(x,z) = 0.2,\ m'_b(y,z) = 0.4,\ m'_b(\Theta) = 0.4, \end{array} \tag{29}$$

We can note that kind of approximation tends to penalize the smallest,

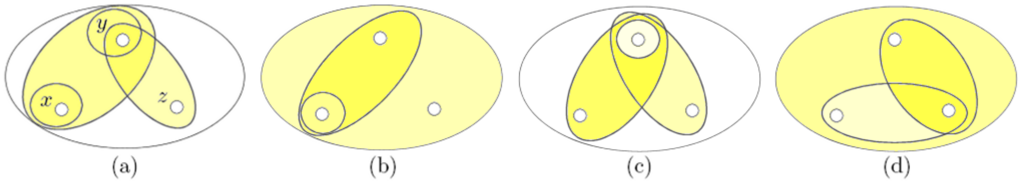

Fig. 6. (a) The belief function (15); (b) its partial L_1/L_2 consistent approximation in $\mathcal{B}$, focussed on $\{x\}$; (c) partial L_1/L_2 partial approximation on $\{y\}$; (d) partial L_1/L_2 partial approximation on $\{z\}$.

singleton element of the ultrafilter $\{B \supseteq x\}$: this is due to the fact that $m'_b(x) = m_b(x)$, while the masses of all the other elements of the filter are increased.
As for global approximations we have, for L_1:

$$\begin{array}{l} \sum_{A\subseteq\{x\}^c} b(A) = b(y) + b(z) + b(y,z) = 0.1 + 0 + 0.4 = 0.5, \\ \sum_{A\subseteq\{y\}^c} b(A) = b(x) + b(z) + b(x,z) = 0.2 + 0 + 0.2 = 0.4, \\ \sum_{A\subseteq\{z\}^c} b(A) = b(x) + b(y) + b(x,y) = 0.2 + 0.1 + 0.7 = 1, \end{array}$$

which is minimized by the partial solution in y. For L_2:

$$\begin{array}{l} \sum_{A\subseteq\{x\}^c} \big(b(A)\big)^2 = 0.01 + 0.16 = 0.17, \\ \sum_{A\subseteq\{y\}^c} \big(b(A)\big)^2 = 0.04 + 0.04 = 0.08, \\ \sum_{A\subseteq\{z\}^c} \big(b(A)\big)^2 = 0.04 + 0.01 + 0.49 = 0.54, \end{array}$$

which is also minimized by the same partial solution. This is not necessarily the case, in general.

5.2 L_∞ consistent approximation

While the partial L_1- and L_2-consistent approximations in the belief space are pointwise and coincide, the set of partial L_∞-approximations for each component $\mathcal{CS}^x_{\mathcal{B}}$ of the consistent complex form a polytope whose center of mass is exactly equal to $cs^x_{\mathcal{B},L_1}[b] = cs^x_{\mathcal{B},L_2}[b]$.

5.2.1 General solution

By applying the expression (8) of the L_∞ norm to the difference vector $\vec{b} - \vec{cs}$ in the belief space, where cs is a consistent b.f. with core containing x, we obtain $\max_{A \subsetneq \Theta} \left\{ \left| \sum_{B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \supseteq \{x\}} m_{cs}(B) \right| \right\}$. Therefore, the partial L_∞ consistent approximation with core containing x is:

$$cs^x_{\mathcal{B},L_\infty}[b] = \arg\min_{m_{cs}(.)} \max_{A \subsetneq \Theta} \left\{ \left| \sum_{B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \supseteq \{x\}} m_{cs}(B) \right| \right\}.$$

Again, $\max_{A \subsetneq \Theta}$ has as lower limit the value associated with the largest norm which does not depend on $m_{cs}(.)$, i.e.,

$$\max_{A \subsetneq \Theta} \left\{ \left| \sum_{B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \supseteq \{x\}} m_{cs}(B) \right| \right\} \geq b(\{x\}^c)$$

or equivalently $\max_{A \subsetneq \Theta} \left\{ \left| \sum_{B \subseteq A, B \supseteq \{x\}} \beta(B) + \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \right| \right\} \geq b(\{x\}^c)$.
In the above constraint only the expressions associated with $B \supseteq \{x\}$ contain variable terms. Therefore the desired optimal values of the variables $\{\beta(B)\}$ are such that:

$$\left\{ \left| \sum_{B \subseteq A, B \supseteq \{x\}} \beta(B) + \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \right| \leq b(\{x\}^c) \quad \forall A : \{x\} \subseteq A \subsetneq \Theta. \right. \tag{30}$$

After introducing the change of variables

$$\gamma(A) \doteq \sum_{B \subseteq A, B \supseteq \{x\}} \beta(B) \tag{31}$$

system (30) trivially reduces to $\left\{ \left| \gamma(A) + \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \right| \leq b(\{x\}^c) \right.$ for all A such that $\{x\} \subseteq A \subsetneq \Theta$, whose solution

$$-b(x^c) - \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \leq \gamma(A) \leq b(x^c) - \sum_{B \subseteq A, B \not\supset \{x\}} m_b(B) \tag{32}$$

defines a high-dimensional "rectangle" in the space of the solutions $\{\gamma(A), \{x\} \subseteq A \subsetneq \Theta\}$.
In the mass assignment $m_{cs}(.)$ of the desired approximations, as we can clearly see in the running ternary example, the solution set is polytope whose vertices do not appear to have straightforward interpretations. On the other hand, the barycenter of this polytope is easy to compute and interpret.

5.2.2 Barycenter of the L_∞ solution and global approximation

The center of mass of the set of solutions (32) to the L_∞ consistent approximation problem is clearly $\gamma(A) = -\sum_{B\subseteq A, B\not\supset\{x\}} m_b(B)$, $\{x\} \subseteq A \subsetneq \Theta$, which reads in the space of the variables $\{\beta(A), \{x\} \subseteq A \subsetneq \Theta\}$ as

$$\left\{ \sum_{B\subseteq A, B\supseteq\{x\}} \beta(B) = -\sum_{B\subseteq A, B\not\supset\{x\}} m_b(B), \qquad \{x\} \subseteq A \subsetneq \Theta. \right.$$

But this is exactly the linear system (23) which determines the L_1/L_2 consistent approximation $cs^x_{\mathcal{B},L_{1/2}}[b]$ of b onto $\mathcal{CS}^x_{\mathcal{B}}$.
Besides, the L_∞ distance between b and $\mathcal{CS}^x_{\mathcal{B}}$ is minimal for the element x which minimizes $\|\vec{b} - \vec{cs}^x_{\mathcal{B},L_\infty}\|_\infty = b(\{x\}^c)$. In conclusion,

Theorem 9 *Given a belief function $b : 2^\Theta \to [0,1]$, and an element of its frame $x \in \Theta$, its partial L_1/L_2 approximation onto any given component $\mathcal{CS}^x_{\mathcal{B}}$ of the consistent complex $\mathcal{CS}_{\mathcal{B}}$ in the belief space is also the geometric barycenter of the set its L_∞ consistent approximations on the same component. Its global L_∞ consistent approximations in $\mathcal{B}$ form the union of the partial L_∞ approximations associated with the maximal plausibility element(s) $x \in \Theta$.*

5.3 Running example

For the usual belief function (15) on the ternary frame, the set of partial L_∞ solutions (32) focussed on x becomes:

$$\begin{aligned}
b(x) - b(x^c) \leq {} & m_{cs}(x) && \leq b(x) + b(x^c), \\
b(x,y) - b(x^c) \leq {} & m_{cs}(x) + m_{cs}(x,y) && \leq b(x,y) + b(x^c), \\
b(x,z) - b(x^c) \leq {} & m_{cs}(x) + m_{cs}(x,z) && \leq b(x,z) + b(x^c),
\end{aligned}$$

whose $2^3 = 8$ vertices (represented as vectors $[m_{cs}(x), m_{cs}(x,y), m_{cs}(x,z), m_{cs}(\Theta)]'$)

$$\begin{aligned}
&\Big[b(x) - b(x^c),\ b(x,y) - b(x),\ b(x,z) - b(x), \\
&\qquad\qquad 1 + b(x) + b(x^c) - b(x,y) - b(x,z)\Big], \\
&\Big[b(x) - b(x^c),\ b(x,y) - b(x),\ b(x,z) - b(x) + 2b(x^c), \\
&\qquad\qquad 1 + b(x) - b(x,y) - b(x,z) - b(x^c)\Big],
\end{aligned}$$

$$\Big[b(x) - b(x^c),\ b(x,y) - b(x) + 2b(x^c),\ b(x,z) - b(x),\ 1 + b(x) - b(x,y) - b(x,z) - b(x^c)\Big],$$
$$\Big[b(x) - b(x^c),\ b(x,y) - b(x) + 2b(x^c),\ b(x,z) - b(x) + 2b(x^c),\ 1 + b(x) - b(x,y) - b(x,z) - 3b(x^c)\Big],$$
$$\Big[b(x) + b(x^c),\ b(x,y) - b(x) - 2b(x^c),\ b(x,z) - b(x) - 2b(x^c),\ 1 + b(x) - b(x,y) - b(x,z) + 3b(x^c)\Big],$$
$$\Big[b(x) + b(x^c),\ b(x,y) - b(x) - 2b(x^c),\ b(x,z) - b(x),\ 1 + b(x) - b(x,y) - b(x,z) + b(x^c)\Big],$$
$$\Big[b(x) + b(x^c),\ b(x,y) - b(x),\ b(x,z) - b(x) - 2b(x^c),\ 1 + b(x) - b(x,y) - b(x,z) + b(x^c)\Big],$$
$$\Big[b(x) + b(x^c),\ b(x,y) - b(x),\ b(x,z) - b(x),\ 1 + b(x) - b(x,y) - b(x,z) - b(x^c)\Big],$$

do not appear to be particularly meaningful. Their barycenter, instead,

$$\begin{aligned} &\Big[b(x), b(x,y) - b(x), b(x,z) - b(x), 1 + b(x) - b(x,y) - b(x,z)\Big] \\ = &\Big[m_b(x), m_b(x,y) + m_b(y), m_b(x,z) + m_b(z), m_b(y,z) + m_b(\Theta)\Big] \end{aligned} \tag{33}$$

clearly coincides with the focussed consistent approximation on x. The same holds for the barycenters of the L_∞ approximations focussed on y and z (see Equation (29)).

6 Approximations in the belief versus the mass space

It is interesting to compare the results obtained in both belief (Section 5) and mass (Section 4) spaces.

6.1 Mass- versus belief- consistent approximations

Summarizing, in the mass space:

- the partial L_1 consistent approximation focussed on a certain element x of the frame is obtained by reassigning all the mass $b(x^c)$ outside the filter to Θ;
- the global approximation is associated, as expected, with cores containing the maximal plausibility element(s) of Θ;

- the L_∞ approximation generates a "rectangle" of partial approximations, with barycenter in the L_1 partial approximation;
- the corresponding global approximations spans the component(s) focussed on the element(s) x such that $\max_{B \not\supset x} m_b(B)$ is minimal;
- the L_2 partial approximation coincides with the L_1 one in the $N-2$ representation;
- in the $N-1$ representation the L_2 partial approximation reassigns the mass outside the desired filter ($b(x^c)$) to each element of the filter focussed on x on equal basis;
- global approximations in the L_2 case are of more difficult interpretation.

In the belief space:

- partial L_1/L_2 approximations coincide on each component of the consistent complex;
- such partial L_1/L_2 approximation turns out to be the consistent transformation [17] focused on the considered element of the frame: for all events A such that $\{x\} \subseteq A \subseteq \Theta$

$$m_{cs^x_{\mathcal{B},L_1}[b]}(A) = m_{cs^x_{\mathcal{B},L_2}[b]}(A) = m_b(A) + m_b(A \setminus \{x\});$$

- the corresponding global solutions have not in general as core the maximal plausibility element, and may lie in general on different components of $\mathcal{CS}$; the L_1 global consistent approximation is associated with the singleton(s) $x \in \Theta$ such that: $\hat{x} = \arg\min_x \sum_{A \subseteq \{x\}^c} b(A)$, while the L_2 global approximation is associated with $\hat{x} = \arg\min_x \sum_{A \subseteq \{x\}^c} (b(A))^2$ which do not appear to have simple epistemic interpretations;
- the set of partial L_∞ solutions form a polytope on each component of the consistent complex, whose center of mass lies on the partial L_1/L_2 approximation;
- the global L_∞ solutions fall on the component(s) associated with the maximal plausibility element(s), and their center of mass, when such element is unique, is the consistent transformation focused on the maximal plausibility singleton [17].

By comparing the behavior of the two classes of approximations we can notice a general pattern. Approximations in both mass and belief space reassign the mass outside the filter focussed on x, in different ways. However, mass consistent approximations reassign this mass $b(x^c)$ either to no focal element in the filter (i.e., to Θ) or to all on an equal basis. They do not distinguish between focal elements w.r.t. their set-theoretic relationships with subsets $B \not\supset x$ outside the filter. In opposition, approximations in the belief space do distinguish them according to the focussed consistent transformation principle.

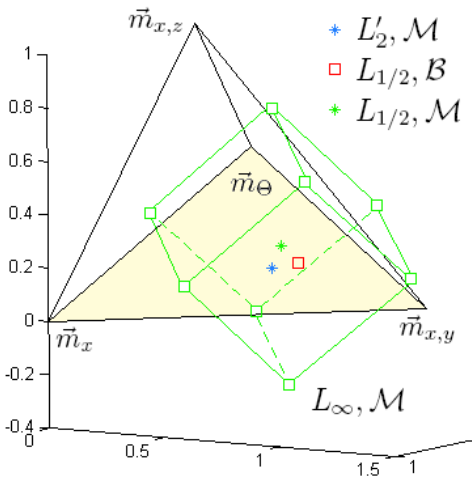

Fig. 7. The simplex (solid black tetrahedron) $Cl(\vec{m}_x, \vec{m}_{x,y}, \vec{m}_{x,z}, \vec{m}_\Theta)$ of consistent belief functions focussed on x, and the related L_p partial consistent approximations of a generic b.f. on $\Theta = \{x, y, z\}$.

6.2 Comparison on a ternary example

To conclude, let us graphically illustrate the different approximations for the belief function of the running ternary example. Figure 7 illustrates the different partial consistent approximations in the simplex $Cl(\vec{m}_x, \vec{m}_{x,y}, \vec{m}_{x,z}, \vec{m}_\Theta)$ of consistent belief functions focussed on x in a ternary frame, for the belief function (15). This is the solid black tetrahedron of Figure 7.

The set of partial L_∞ approximations in $\mathcal{M}$ (18) is represented by the green cube in the Figure. 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 approximation in $\mathcal{M}$ does also coincide, as expected, with the L_1 approximation.
There seems to be a strong case for the latter approximation, whose natural interpretation in terms of mass assignment is the following: all the mass outside the desired ultrafilter $\{B \supseteq x\}$ is reassigned to Θ, increasing the overall uncertainty of the belief state.
The $L_2, \mathcal{M}$ partial approximation in the $N-1$ representation is distinct from the previous ones, but still falls inside the polytope of L_∞ partial approximations (green cube) and is admissible, as it falls in the interior of the simplicial component $Cl(\vec{m}_x, \vec{m}_{x,y}, \vec{m}_{x,z}, \vec{m}_\Theta)$. It itself possesses a strong interpretation, as the overall mass not in the desired ultrafilter focused on x is equally redistributed to all elements of the filter.
Finally, the unique L_1/L_2 partial approximation in $\mathcal{B}$ of Equation (29) (which is also the barycenter (33) of the partial L_∞ approximations in $\mathcal{B}$) is shown as a red square. Just as the L_1/L_2 approximation in $\mathcal{M}$ (green star), it attributes zero mass to $\{x, z\}$, which fails to be supported by any focal element

of the original belief function. As a result, they both fall on the border of the tetrahedron of admissible consistent b.f.s focussed on x (the face highlighted in yellow).

7 Conclusions

Belief functions represent coherent knowledge bases in the theory of evidence. As consistency is not preserved by most operators used to update or elicit evidence, the use of a consistent transformation in conjunction with those combinations rules can be desirable. In this paper we solved the instance of the consistent approximation problem obtained by measuring distances between uncertainty measures by classical L_p norms, in both belief and mass spaces. This makes sense as cs.b.f.s live in a simplicial complex defined in terms of the L_∞ norms, where they correspond to possibility distributions. The obtained approximations typically have straightforward interpretations in terms of degrees of belief or mass redistribution. The interpretation of the polytope of all L_∞ solutions is a bit more complex, and worth to be fully investigated in the near future, in the light of the interesting analogy with the polytope of consistent probabilities.

Appendix

Proof of Theorem 5. The desired orthogonality condition reads as $\langle \vec{m}_b - \vec{m}_{cs}, \vec{m}_B - \vec{m}_{\{x\}} \rangle = 0$ where $\vec{m}_b - \vec{m}_{cs}$ is given by Equation (10), while $\vec{m}_B - \vec{m}_{\{x\}}(C) = 1$ if $C = B$, $= -1$ if $C = \{x\}$, 0 elsewhere. Therefore, using once again the variables $\{\beta(B)\}$, the condition simplifies as follows:

$$\langle \vec{m}_b - \vec{m}_{cs}, \vec{m}_B - \vec{m}_{\{x\}} \rangle = \begin{cases} \beta(B) - \beta(\{x\}) = 0 \ \forall B \supsetneq \{x\}, B \neq \Theta; \\ -\beta(x) = 0 \qquad\qquad B = \Theta. \end{cases} \tag{34}$$

When using the $(N-1)$-dimensional representation (19) of mass vectors, the orthogonality condition reads instead as:

$$\langle \vec{m}_b - \vec{m}_{cs}, \vec{m}_B - \vec{m}_{\{x\}} \rangle = \beta(B) - \beta(\{x\}) = 0 \quad \forall\, B \supsetneq \{x\}. \tag{35}$$

In the $N-2$ representation, by (34) 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 (13).
In the $N-1$ representation, the orthogonality condition (35) reads as

$$m_{cs}(B) = m_{cs}(x) + m_b(B) - m_b(x) \quad \forall B \supsetneq \{x\}.$$

By normalizing it 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 (20).

Proof of Lemma 1. We first note that, by definition of categorical belief function b_A, $\langle \vec{b}_B, \vec{b}_C \rangle = \sum_{D\supseteq B,C; D\neq\Theta} 1 = \sum_{E\subsetneq(B\cup C)^c} 1 = 2^{|(B\cup C)^c|} - 1$. Hence:

$$\begin{aligned} &\sum_{B\subseteq A} \langle \vec{b}_B, \vec{b}_C \rangle (-1)^{|B\setminus A|} = \sum_{B\subseteq A} (2^{|(B\cup C)^c|} - 1)(-1)^{|B\setminus A|} \\ &= \sum_{B\subseteq A} 2^{|(B\cup C)^c|}(-1)^{|B\setminus A|} - \sum_{B\subseteq A} (-1)^{|B\setminus A|} = \sum_{B\subseteq A} 2^{|(B\cup C)^c|}(-1)^{|B\setminus A|} \end{aligned} \tag{36}$$

as $\sum_{B\subseteq A} (-1)^{|B\setminus A|} = \sum_{k=0}^{|B\setminus A|} 1^{|A^c|-k}(-1)^k = 0$ by Newton's binomial: $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n$. Now, as both $B \supseteq A$ and $C \supseteq A$ the set B can be decomposed into the disjoint sum $B = A + B' + B''$, where $\emptyset \subseteq B' \subseteq C \setminus A$, $\emptyset \subseteq B'' \subseteq (C \cup A)^c$. Therefore (36) becomes: $\sum_{\emptyset\subseteq B'\subseteq C\setminus A} \sum_{\emptyset\subseteq B''\subseteq(C\cup A)^c} 2^{|(A\cup C)|^c - |B''|}(-1)^{|B'|+|B''|} =$
$= \sum_{\emptyset\subseteq B'\subseteq C\setminus A} (-1)^{|B'|} \sum_{\emptyset\subseteq B''\subseteq(C\cup A)^c} (-1)^{|B''|} 2^{|(A\cup C)|^c - |B''|}$, where
$\sum_{\emptyset\subseteq B''\subseteq(C\cup A)^c} (-1)^{|B''|} 2^{|(A\cup C)|^c - |B''|} = [2 + (-1)]^{|(A\cup C)|^c} = 1^{|(A\cup C)|^c} = 1$, again by Newton's binomial. The quantity (36) is therefore equal to $\sum_{\emptyset\subseteq B'\subseteq C\setminus A} (-1)^{|B'|}$, which is nil for $C \setminus A \neq \emptyset$, equal to 1 when $C \subseteq A$.

Proof of Theorem 7. The L_1 distance between the partial approximation and $\vec{b}$ can be easily computed as: $\|\vec{b} - \vec{cs}^{x}_{\mathcal{B},L_1}[b]\|_{L_1} =$

$$\begin{aligned} &= \sum_{A\subseteq\Theta} |b(A) - cs^x_{\mathcal{B},L_1}[b](A)| \\ &= \sum_{A\not\supseteq\{x\}} |b(A) - 0| + \sum_{A\supseteq\{x\}} \Big| b(A) - \sum_{B\subseteq A, B\supseteq\{x\}} m_{cs}(B) \Big| \\ &= \sum_{A\not\supseteq\{x\}} b(A) + \sum_{A\supseteq\{x\}} \Big| \sum_{B\subseteq A} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} (m_b(B) + m_b(B\setminus\{x\})) \Big| \\ &= \sum_{A\not\supseteq\{x\}} b(A) + \sum_{A\supseteq\{x\}} \Big| \sum_{B\subseteq A, B\not\supseteq\{x\}} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} m_b(B\setminus\{x\}) \Big| \\ &= \sum_{A\not\supseteq\{x\}} b(A) + \sum_{A\supseteq\{x\}} \Big| \sum_{C\subseteq A\setminus\{x\}} m_b(C) - \sum_{C\subseteq A\setminus\{x\}} m_b(C) \Big| \\ &= \sum_{A\not\supseteq\{x\}} b(A) = \sum_{A\subseteq\{x\}^c} b(A). \end{aligned}$$

Proof of Theorem 8. The L_2 distance between the partial approximation and $\vec{b}$ can be easily computed as: $\|\vec{b} - \vec{cs}^x_{\mathcal{B},L_2}[b]\|^2 =$

$$
\begin{aligned}
&= \sum_{A\subseteq\Theta} (b(A) - cs^x_{\mathcal{B},L_2}[b](A))^2 = \sum_{A\subseteq\Theta} \bigg(\sum_{B\subseteq A} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} m_{cs}(B) \bigg)^2 \\
&= \sum_{A\subseteq\Theta} \bigg(\sum_{B\subseteq A} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} m_b(B\setminus\{x\}) \bigg)^2 \\
&= \sum_{A\not\supset\{x\}} (b(A))^2 + \sum_{A\supseteq\{x\}} \bigg(\sum_{B\subseteq A, B\not\supset\{x\}} m_b(B) - \sum_{B\subseteq A, B\supseteq\{x\}} m_b(B\setminus\{x\}) \bigg)^2 \\
&= \sum_{A\not\supset\{x\}} (b(A))^2 + \sum_{A\supseteq\{x\}} \bigg(\sum_{C\subseteq A\setminus\{x\}} m_b(C) - \sum_{C\subseteq A\setminus\{x\}} m_b(C) \bigg)^2
\end{aligned}
$$

so that $\|\vec{b} - \vec{cs}^x_{\mathcal{B},L_2}[b]\|^2 = \sum_{A\not\supset\{x\}} (b(A))^2 = \sum_{A\subseteq\{x\}^c} (b(A))^2$.

References

[1] P. Baroni, *Extending consonant approximations to capacities*, Proc. of IPMU, 2004, pp. 1127–1134.

[2] D. Batens, C. Mortensen, and G. Priest, *Frontiers of paraconsistent logic*, Studies in logic and computation, vol. 8, Research Studies Press, 2000.

[3] P. Black, *An examination of belief functions and other monotone capacities*, PhD dissertation, Department of Statistics, Carnegie Mellon University.

[4] B.R. Cobb and P.P. Shenoy, *A comparison of Bayesian and belief function reasoning*, Information Systems Frontiers **5** (2003), no. 4, 345–358.

[5] F. Cuzzolin, *Two new Bayesian approximations of belief functions based on convex geometry*, IEEE Transactions on Systems, Man, and Cybernetics part B **37** (2007), no. 4, 993–1008.

[6] ———, *A geometric approach to the theory of evidence*, IEEE Transactions on Systems, Man, and Cybernetics part C **38** (2008), no. 4, 522–534.

[7] ———, *An interpretation of consistent belief functions in terms of simplicial complexes*, Proc. of ISAIM'08, Fort Lauderdale, Florida, 2008.

[8] ———, *Geometric conditioning of belief functions*, Proc. of BELIEF'10, Brest, France, 2010.

[9] ———, *The geometry of consonant belief functions: simplicial complexes of necessity measures*, Fuzzy Sets and Systems **161** (2010), no. 10, 1459–1479.

[10] ———, *Geometric conditional belief functions in the belief space*, Proc. of ISIPTA'11, Innsbruck, Austria, 2011.

[11] ———, *Lp consonant approximations of belief functions in the mass space*, Proceedings of ISIPTA'11, Innsbruck, Austria, 2011.

[12] ———, *Consistent approximation of belief functions*, Proceedings of ISIPTA'09, Durham, UK, 2009.

[13] F. Cuzzolin and R. Frezza, *Geometric analysis of belief space and conditional subspaces*, Proc. of ISIPTA, Ithaca, NY, 2001.

[14] M. Daniel, *On transformations of belief functions to probabilities*, International Journal of Intelligent Systems, DETAILS??

[15] A.P. Dempster, *Upper and lower probabilities induced by a multivariate mapping*, Annals of Mathematical Statistics **38** (1967), 325–339.

[16] J. Diaz, M. Rifqi, and B. Bouchon-Meunier, *A similarity measure between basic belief assignments*, Proc. of FUSION, 2006.

[17] D. Dubois and H. Prade, *Consonant approximations of belief functions*, International Journal of Approximate Reasoning **4** (1990), 419–449.

[18] D. Dubois, H. Prade, and S. Sandri, *On possibility/probability transformations*, JOURNAL?? 1993.

[19] B.A. Dubrovin, S.P. Novikov, and A.T. Fomenko, *Sovremennaja geometrija. metody i prilozenija*, Nauka, Moscow, 1986.

[20] P. Gardenfors, *Knowledge in flux: Modeling the dynamics of epistemic states*, MIT Press, Cambridge, MA, 1988.

[21] R. Haenni, *Towards a unifying theory of logical and probabilistic reasoning*, Proc. of ISIPTA, 2005.

[22] A. Hunter and W. Liu, *Fusion rules for merging uncertain information*, Information Fusion **7** (2006), no. 1, 97–134.

[23] W. Jiang, A. Zhang, and Q. Yang, *A new method to determine evidence discounting coefficient*, Lecture Notes in Computer Science, vol. 5226/2008, 2008, pp. 882–887.

[24] A.-L. Jousselme and P. Maupin, *On some properties of distances in evidence theory*, Proc. of BELIEF'10, Brest, France, 2010.

[25] V. Khatibi and G.A. Montazer, *A new evidential distance measure based on belief intervals*, Scientia Iranica - Transactions D: Computer Science and Engineering and Electrical Engineering **17** (2010), no. 2, 119–132.

[26] D.K. Lewis, *Probabilities of conditionals and conditional probabilities*, Philosophical Review **85** (1976), 297–315.

[27] W. Liu, *Analyzing the degree of conflict among belief functions*, Artificial Intelligence **170** (2006), no. 11, 909–924.

[28] K–C. Lo, *Agreement and stochastic independence of belief functions*, Mathematical Social Sciences **51** (2006), no. 1, 1–22.

[29] B. Mates, *Elementary logic*, Oxford University Press, 1972.

[30] J.B. Paris, D. Picado-Muino, and M. Rosefield, *Information from inconsistent knowledge: A probability logic approach*, Advances in Soft Computing, vol. 46, Springer-Verlag, 2008.

[31] A. Perea, *A model of minimal probabilistic belief revision*, Theory and Decision **67** (2009), no. 2, 163–222.

[32] G. Priest, R. Routley, and J. Norman, *Paraconsistent logic: Essays on the inconsistent*, Philosophia Verlag, 1989.

[33] A. Ramer and G.J. Klir, *Measures of discord in the Dempster-Shafer theory.*, Information Sciences **67** (1993), no. 1-2, 35–50.

[34] A. Saffiotti, *A belief-function logic*, Universitè Libre de Bruxelles, MIT Press, pp. 642–647.

[35] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[36] C. Shi, Y. Cheng, Q. Pan, and Y. Lu, *A new method to determine evidence distance*, Proc. of CiSE, 2010, pp. 1–4.

[37] Ph. Smets, *The degree of belief in a fuzzy event*, Information Sciences **25** (1981), 1–19.

[38] ———, *The nature of the unnormalized beliefs encountered in the transferable belief model*, Proc. of UAI, San Mateo, CA, 1992, pp. 292–29.

[39] ———, *Belief functions versus probability functions*, Uncertainty and Intelligent Systems (Saitta L. Bouchon B. and Yager R., eds.), Springer Verlag, Berlin, 1988, pp. 17–24.

[40] ———, *Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem*, International Journal of Approximate Reasoning **9** (1993), 1–35.

[41] F. Voorbraak, *A computationally efficient approximation of Dempster-Shafer theory*, International Journal on Man-Machine Studies **30** (1989), 525–536.

[42] R.R. Yager, *On the Dempster-Shafer framework and new combination rules*, Information Sciences **41** (1987), 93–138.