L_p consonant approximations of belief functions

Fabio Cuzzolin

Abstract

In this paper we solve the problem of approximating a belief measure with a necessity measure or "consonant belief function" in a convex geometric framework. Consonant belief functions form a simplicial complex in both the space of all belief functions and the space of all mass vectors. As necessity measures are strictly related to the L_∞ norm it makes sense to look for approximations which minimize L_p norms. Partial approximations are fist sought in each simplicial component of the consonant complex, while global solutions are obtained from the set of such partial ones. The obtained L_p consonant approximations are discussed and their interpretation in terms of degrees of belief provided. Results are also compared to classical outer consonant approximations.

Index Terms

Theory of evidence, consonant belief functions, geometric approach, simplicial complex, (outer) consonant approximation, L_p norms.

I. Introduction: the consonant approximation problem

The theory of evidence (ToE) [1] is a popular approach to uncertainty description. Probabilities are there replaced by *belief functions* (b.f.s) $b : 2^\Theta \rightarrow [0,1]$, which assign values between 0 and 1 to subsets of the sample space Θ instead of single elements. Possibility theory [2], on its side, is based on *possibility measures*, i.e., functions $Pos : 2^\Theta \rightarrow [0,1]$ on Θ such that $Pos(\bigcup_i A_i) = \sup_i Pos(A_i)$ for any family $\{A_i | A_i \in 2^\Theta, i \in I\}$ where I is an arbitrary set index. Given a possibility measure Pos, the dual *necessity* measure is defined as $Nec(A) = 1 - Pos(A)$.

F. Cuzzolin is with the Department of Computing, Oxford Brookes University, Oxford, UK. Email: Fabio.Cuzzolin@brookes.ac.uk.

Necessity measures have as counterparts in the theory of evidence *consonant* b.f.s, i.e., belief functions whose non-zero mass events or "focal elements" are nested [1] and form a chain of subsets $A_1 \subset \cdots \subset A_m$, $A_i \subseteq \Theta$. The problem of approximating a belief function with a necessity measure is then equivalent to approximating a belief function with a consonant b.f. [3], [4], [5], [6]. As possibilities are completely determined by their values on the singletons ($Pos(x)$, $x \in \Theta$), they are less computationally expensive than b.f.s, making the approximation process interesting for many applications. Many authors, such as Yager [7] and Romer [8] amongst others, have studied the connection between fuzzy numbers and Dempster-Shafer theory. Klir *et al* have published an excellent discussion [9] on the relations among fuzzy and belief measures and possibility theory. Heilpern [10] has also presented the theoretical background of fuzzy numbers connected with the possibility and Dempster-Shafer theories, describing some types of representation of fuzzy numbers and studying the notions of distance and order between fuzzy numbers based on these representations. Caro and Nadjar [11], instead, have suggested a generalization of the Dempster-Shafer theory to a fuzzy valued measure. The links between transferable belief model and possibility theory have been briefly investigated by Ph. Smets in [12].

A. *Consonant approximation*

Dubois and Prade [3], more specifically, have extensively worked on consonant approximations of belief functions. Their work has been later considered in [4], [5]. In particular, the notion of "outer consonant approximation" has received considerable attention in the past. Indeed, belief functions admit the following order relation: $b \leq b' \equiv \forall A \subseteq \Theta b(A) \leq b'(A)$, called "weak inclusion". It is then possible to introduce the notion of "outer consonant approximations" [3] of a belief function b, i.e., those co.b.f.s $co : 2^\Theta \rightarrow [0, 1]$ such that $\forall A \subseteq \Theta \; co(A) \leq b(A)$. Dubois and Prade's work has been later extended by Baroni [6] to capacities. In [13] the author has provided a comprehensive description of the geometry of the set of outer consonant approximations.

B. *Geometric approach to approximation*

In more recent times the opportunity of seeking probability or consonant approximations / transformations of belief functions by minimizing appropriate distance functions has been explored. The author has himself introduced the notion of orthogonal projection $\pi[b]$ of a belief function onto the probability simplex [14], and studied consistent approximations of belief functions induced by classical L_p norms [15], [16] in the space of belief functions [17]. In [18] he has shown that norm minimization can also be used to define families of geometric conditional belief functions.

Jousselme et al [19] have recently conducted a very nice survey of the distance or similarity measures so far introduced between belief functions, come out with an interesting classification, and proposed a number of generalizations of known measures. Other similarity measures between belief functions have been proposed by Shi et al [20], Jiang et al [21], and others [22], [23], [21]. Many of these measures could be in principle employed to define conditional belief functions, or to approximate belief functions by necessity or probability measures.

C. Contribution

In this paper we derive the expressions of all the consonant approximations of belief functions induced by minimizing L_p distances in both the mass and the belief space.

As it turns out, approximations *in the mass space* do not take into account the contributions of subsets of Θ outside the desired chain of focal elements to the plausibility of the elements of the chain. A similar phenomenon has been observed in the case of geometric conditioning of belief functions by L_p minimization [18].

The problem can be posed in two different versions of the mass space, of dimensions $N-1$ and $N-2$, respectively.

In the first case, the set of L_1 approximations is the set of co.b.f.s whose mass values dominate that of the original belief function on the desired chain. This set is a simplex, whose vertices are obtained by re-assigning all the mass originally outside the desired chain to a single focal element of the chain itself. Its barycenter coincides with the L_2 partial approximation, which redistributes the mass outside the chain to all the elements of the chain on an equal basis. When the (partial) L_∞ approximation is unique, it coincides with the L_2 approximation and the barycenter of the L_1 approximations. When it is not unique, it is a simplex whose vertices assign to each element of the chain (but one) the maximal mass outside the chain, and whose barycenter is again the L_2 approximation.

In the second representation of $\mathcal{M}$, the L_∞ (partial) approximation is not unique, and it forms a generalized rectangle in the mass space $\mathcal{M}$, whose size is determined by the largest mass outside the desired maximal chain. The L_1 and L_2 partial approximations are uniquely determined, and coincide with the barycenter of the set of L_∞ partial approximations. Their semantic is straightforward: all the mass outside the chain is re-assigned to Θ, increasing the overall uncertainty of the belief state.

In the *belief space*, L_p approximations have rather more complex forms. Rather than appearing like suitable candidates for the role of possibilistic approximations, they tend to exhibit an interesting relationship with the notion of "derivative" of a belief function on a totally ordered chain of focal elements.

In particular, L_∞ minimization generates an entire convex set of (partial) approximations on each simplicial component. The barycenter of this set has indeed a potentially interesting interpretation in terms of a formally to specify notion of derivative of a belief function on a linearly ordered chain, while the global L_∞ approximations fall as expected on the component associated with the maximal plausibility singleton. The L_1 norm, on the other hand, does not seem to be suitable for the job. However, in the cases in which the analytical form of the set of L_1 approximations is obtainable in closed form, this seems to be connected to the difference of belief values on the chain. Finally, the L_2 partial approximation is unique and distinct from the above barycenter, while the related global L_2 approximation is rather elusive.

D. Paper outline

We first provide the necessary background on consonant b.f.s and consonant approximation (Section II), in particular the geometric representation of belief and mass vectors (II-B) and the geometric approach to the approximation problem (II-C). We then move to studying the consonant approximation problem by L_p minimization in the simple binary case (Section III). Using the intuition gathered in the binary case, we first approach the problem in the mass space (Section IV). We compute the approximations induced by L_1 (IV-B), L_2 (IV-C) and L_∞ (IV-D) norms, respectively, and discuss their interpretation in terms of mass re-assignment and their relation with outer consonant approximations in Section IV-E, illustrating them in the significant ternary case.

In the second part of the paper we analyze the L_p approximation problem in the belief space (Section V). Again, we compute the approximations induced by L_∞ (V-B), L_1 (V-C) and L_2 (V-D) norms, respectively. We propose an interpretation of the resulting approximations in terms of degrees of belief, and comment on their relationship with other known approximations with once again the help of the ternary case (Section V-E).

II. Geometry of consonant belief functions

A. Consonant belief functions as necessity measures

We briefly recall here some basis definitions of belief calculus.

A *basic probability assignment* (b.p.a.) over a finite set *(frame of discernment [1])* Θ is a function $m_b : 2^\Theta \rightarrow [0,1]$ on its power set $2^\Theta = \{A \subseteq \Theta\}$ such that $m_b(\emptyset) = 0$, $\sum_{A \subseteq \Theta} m_b(A) = 1$, and $m_b(A) \geq 0 \; \forall A \subseteq \Theta$. Subsets of Θ associated with non-zero values of m_b are called *focal elements*. 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 $pl_b(A) \doteq 1 - b(A^c) = 1 - \sum_{B \subseteq A^c} m_b(B) = \sum_{B \cap A \neq \emptyset} m_b(B)$, and expresses the amount of evidence *not against* A. In the theory of evidence a probability function is simply a special belief function assigning non-zero masses to singletons only (*Bayesian* b.f.): $m_b(A) = 0 \; |A| > 1$.

A belief function is said to be *consonant* if its focal elements are nested, and form a chain $A_1 \subset \cdots \subset A_m$. Consonant belief functions are characterized by the fact that $pl_b(A) = \max_{x \in A} pl_b(x)$ for all non-empty $A \subset \Theta$.

Consonant belief functions are the trait d'union between the theory of belief functions and possibility theory [2]. While the former relies on belief functions to represent uncertainty, the latter is centered on the notion of "possibility measure". A *possibility measure* on Θ is a function $Pos : 2^\Theta \rightarrow [0,1]$ such that $Pos(\emptyset) = 0$, $Pos(\Theta) = 1$ and $Pos(\bigcup_i A_i) = \sup_i Pos(A_i)$ for any family $\{A_i | A_i \in 2^\Theta, i \in I\}$ where I is an arbitrary set index. Each possibility measure is uniquely characterized by a *membership function* or *possibility distribution* $\pi : \Theta \rightarrow [0,1]$ s.t. $\pi(x) \doteq Pos(\{x\})$ via the formula $Pos(A) = \sup_{x \in A} \pi(x)$. $Nec(A) = 1 - Pos(A^c)$ is called *necessity measure*.

Now, given a belief function b, let us call *plausibility assignment* (pl.ass.) $\bar{pl}_b$ the restriction of the plausibility function to singletons $\bar{pl}_b(x) = pl_b(\{x\})$. From the above property of consonant b.f.s it follows that:

Proposition 1: The plausibility function pl_b associated with a belief function b on a domain Θ is a possibility measure iff b is consonant, in which case $\pi = \bar{pl}_b$. Equivalently, a b.f. b is a necessity measure iff b is consonant.
Consequently, approximating an arbitrary b.f. with a consonant belief function amounts to finding a possibility proxy for it.

B. Geometric representation of uncertainty measures

Belief space representation. Given a frame Θ, each belief function $b : 2^\Theta \rightarrow [0,1]$ is completely specified by its $N-2$ belief values $\{b(A), \emptyset \subsetneq A \subsetneq \Theta\}$, $N \doteq 2^n$ ($n \doteq |\Theta|$), (as $b(\emptyset) = 0$, $b(\Theta) = 1$ for all b.f.s) and can therefore be represented as a point of $\mathbb{R}^{N-2}$. Once introduced a set of coordinate axes $\{X_A, \emptyset \subsetneq A \subsetneq \Theta\}$ in $\mathbb{R}^{N-2}$, a belief function b can be represented by the vector $\vec{b} = \sum_{\emptyset \subsetneq A \subsetneq \Theta} b(A) X_A$. If we denote by $\vec{b}_A \doteq b \in \mathcal{B}$ s.t. $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0 \ \forall B \subseteq \Theta$, $B \neq A$ the *categorical* [24] belief function assigning all the mass to a single subset $A \subseteq \Theta$, we can prove that [25], [17] the set of points of $\mathbb{R}^{N-2}$ which correspond to a b.f. or "belief space" $\mathcal{B}$ coincides with the convex closure of all the vectors representing categorical belief functions: $\mathcal{B} = Cl(\vec{b}_A, \emptyset \subsetneq A \subseteq \Theta)$, where Cl denotes the convex closure operator:

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

The belief space $\mathcal{B}$ is a simplex[1] [17], and each vector $\vec{b} \in \mathcal{B}$ representing a belief function b can be written as a convex sum as:

$$\vec{b} = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A) \vec{b}_A. \tag{1}$$

The b.p.a. m_b of b is nothing but the set of simplicial coordinates of the vector $\vec{b}$ in $\mathcal{B}$. The set $\mathcal{P}$ of all Bayesian b.f.s is the simplex determined by all the categorical b.f.s associated with singletons[2]: $\mathcal{P} = Cl(\vec{b}_x, x \in \Theta)$.

Mass space representation. 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

[2]With a harmless abuse of notation we denote the categorical b.f. associated with a singleton x by b_x instead of $b_{\{x\}}$, and write $m_b(x), pl_b(x)$ instead of $m_b(\{x\}), pl_b(\{x\})$.

seen also as a point of $\mathbb{R}^{N-1}$, the vector $\vec{m}_b$ of its $N-1$ mass components:

$$\vec{m}_b = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B)\vec{m}_B, \tag{2}$$

where $\vec{m}_B$ is the vector of mass values associated with the ("categorical") mass functions $\vec{m}_A$ assigning all the mass to a single event A: $\vec{m}_A(A) = 1$, $\vec{m}_A(B) = 0 \; \forall B \neq A$. Note that in $\mathbb{R}^{N-1}$ $\vec{m}_\Theta = [0, ..., 0, 1]'$ and cannot be neglected.

However, since the mass of Θ is determined by all the other masses in virtue of the normalization constraint, we can also choose to represent mass vectors as vectors of $\mathbb{R}^{N-2}$ of the form

$$\vec{m}_b = \sum_{\emptyset \subsetneq B \subsetneq \Theta} m_b(B)\vec{m}_B \tag{3}$$

in which case the component Θ is neglected. This leads to two possible consonant approximations in the mass space. We will consider both representations in the following.

Whatever the chosen representation, the collection $\mathcal{M}$ of points which are valid basic probability assignments is also a simplex, which we can call *mass space*. $\mathcal{M}$ is the convex closure $\mathcal{M} = Cl(\vec{m}_A, \emptyset \subsetneq A \subset \Theta)$ (Θ included or not, depending on the representation).

Binary example. In the case of a frame of discernment containing only two elements, $\Theta_2 = \{x, y\}$, each b.f. $b : 2^{\Theta_2} \to [0, 1]$ is completely determined by its mass values $m_b(x)$, $m_b(y)$, 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 $N = 2^2 = 4$): $\vec{m}_b = [m_b(x), m_b(y)]' \in \mathbb{R}^2$. In this example we adopt therefore the $N-2$-dimensional version of the mass space, which indeed corresponds to the belief space representation. 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 $Cl(\vec{m}_x, \vec{m}_y) = Cl(\vec{b}_x, \vec{b}_y)$.

[1] An n-dimensional *simplex* is the convex closure $Cl(x_1, ..., x_{n+1})$ of $n+1$ (affinely independent) points $x_1, ..., x_{n+1}$ of the Euclidean space $\mathbb{R}^n$. An *affine combination* of k points $v_1, ..., v_k \in \mathbb{R}^m$ is a sum $\alpha_1 v_1 + \cdots + \alpha_k v_k$ such that $\sum_i \alpha_i = 1$. The affine subspace generated by the points $v_1, ..., v_k \in \mathbb{R}^m$ is the set $\{v \in \mathbb{R}^m : v = \alpha_1 v_1 + \cdots + \alpha_k v_k, \sum_i \alpha_i = 1\}$. If $v_1, ..., v_k$ generate an affine space of dimension k they are said to be *affinely independent*. The *faces* of an n-dimensional simplex are all the possible simplices generated by a subset of its vertices, i.e. $Cl(x_{j_1}, ..., x_{j_k})$ with $\{j_1, ..., j_k\} \subset \{1, ..., n+1\}$.

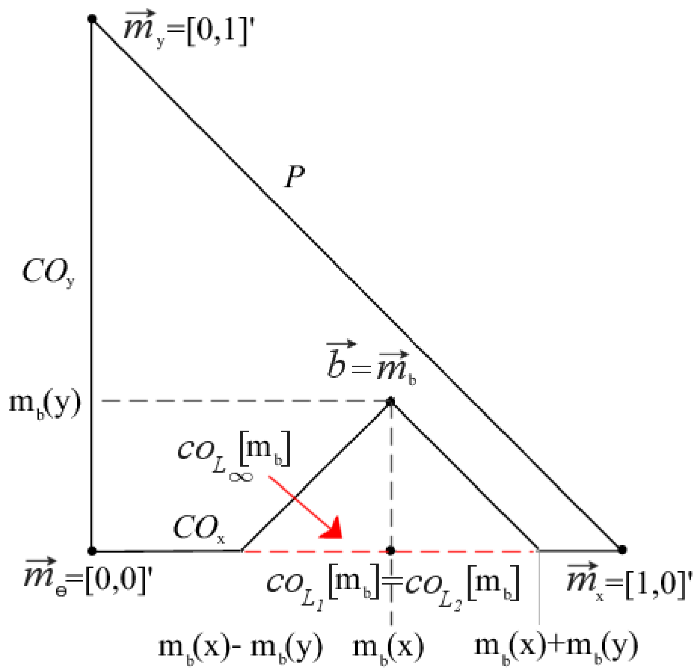

Fig. 1. The mass space $\mathcal{M}_2$ for a binary frame is a triangle in $\mathbb{R}^2$ whose vertices are the mass vectors associated with the categorical belief functions focused on $\{x\}, \{y\}$ and Θ: $\vec{m}_x, \vec{m}_y, \vec{m}_\Theta$. The belief space $\mathcal{B}_2$ coincides with $\mathcal{M}_2$ when $\Theta = \{x, y\}$. Consonant belief functions live in the union of the two segments $\mathcal{CO}_x = Cl(\vec{m}_\Theta, \vec{m}_x)$ and $\mathcal{CO}_y = Cl(\vec{m}_\Theta, \vec{m}_y)$. The unique $L_1 = L_2$ consonant approximation and the set of L_∞ consonant approximations (dashed) on $\mathcal{CO}_x$ are also shown.

On $\Theta_2 = \{x, y\}$ consonant belief functions can have as chain of focal elements either $\{\{x\}, \Theta_2\}$ or $\{\{y\}, \Theta_2\}$. Therefore the region $\mathcal{CO}_2$ of all the co.b.f.s on Θ_2 is the union of two segments (see Figure 1): $\mathcal{CO}_2 = \mathcal{CO}_x \cup \mathcal{CO}_y = Cl(\vec{m}_\Theta, \vec{m}_x) \cup Cl(\vec{m}_\Theta, \vec{m}_y)$.

C. The consonant approximation problem

The consonant complex. In this framework the geometry of consonant belief functions can be described by resorting to the notion of "simplicial complex" [26]. 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.

It can be proven that [13]: the region $\mathcal{CO}_\mathcal{B}$ of consonant belief functions in the belief space is a simplicial complex. More precisely, $\mathcal{CO}_\mathcal{B}$ is the union of a collection of (maximal) simplices, each of them associated with a maximal chain $\mathcal{C} = \{A_1 \subset \cdots \subset A_n\}$, $|A_i| = i$, $A_n = \Theta$ of subsets of Θ:

$$\mathcal{CO}_{\mathcal{B}} = \bigcup_{\mathcal{C}=A_1\subseteq\cdots\subseteq A_n} Cl(\vec{b}_{A_1}, \cdots, \vec{b}_{A_n}).$$

Analogously, the region $\mathcal{CO}_{\mathcal{M}}$ of consonant belief functions in the mass space $\mathcal{M}$ will be the simplicial complex:

$$\mathcal{CO}_{\mathcal{M}} = \bigcup_{\mathcal{C}=A_1\subseteq\cdots\subseteq A_n} Cl(\vec{m}_{A_1}, \cdots, \vec{m}_{A_n}).$$

Consonant approximation. Given a belief function b, we call *consonant approximation of a belief function* b *induced by a distance function* d in $\mathcal{M}$ ($\mathcal{B}$) the b.f.(s) $co_d[b/m_b]$ which minimize(s) the distance $d(\vec{m}_b, \mathcal{CO}_{\mathcal{M}})$ ($d(\vec{b}, \mathcal{CO}_{\mathcal{B}})$) between b and the consonant simplicial complex in $\mathcal{M}$ ($\mathcal{B}$):

$$co_d[m_b] = \arg\min_{\vec{m}_{co}\in\mathcal{CO}_{\mathcal{M}}} d(\vec{m}_b, \vec{m}_{co}) \Big/ co_d[b] = \arg\min_{\vec{co}\in\mathcal{CO}_{\mathcal{B}}} d(\vec{b}, \vec{co}). \tag{4}$$

Why use L_p norms. A close relation exists between consonant belief functions and L_p norms, in particular the L_∞ one. Consonant b.f.s are the counterparts of necessity measures in the theory of evidence, so that their plausibility functions are possibility measures. Possibility measures Pos, in turn, are inherently related to L_∞ as $Pos(A) = \max_{x\in A} Pos(x)$. It makes therefore sense to conjecture that a consonant transformation obtained by picking as distance function in the approximation problem (4) one of the classical L_p norms would 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:

$$\begin{aligned} \|\vec{m}_b - \vec{m}_{b'}\|_{L_1} &\doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B) - m_{b'}(B)|; \\ \|\vec{m}_b - \vec{m}_{b'}\|_{L_2} &\doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta} (m_b(B) - m_{b'}(B))^2}; \\ \|\vec{m}_b - \vec{m}_{b'}\|_{L_\infty} &\doteq \max_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B) - m_{b'}(B)|, \end{aligned} \tag{5}$$

while the same norms in the belief space read as:

$$\begin{array}{c} \|\vec{b} - \vec{b'}\|_{L_1} \doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |b(B) - b'(B)|; \quad \|\vec{b} - \vec{b'}\|_{L_2} \doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta} (b(B) - b'(B))^2}; \\ \|\vec{b} - \vec{b'}\|_{L_\infty} \doteq \max_{\emptyset\subsetneq B\subseteq\Theta} |b(B) - b'(B)|. \end{array} \tag{6}$$

In the probabilistic case, in the belief space ($p[b] = \arg\min_{p\in\mathcal{P}} dist(b,p)$), 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 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 consonant complex $\mathcal{CO}$ is a *collection* of linear spaces (better, simplices which generate a linear space) in both the belief and the mass space, solving the consonant approximation problem involves finding a number of partial solutions

$$co^{\mathcal{C}}_{L_p}[b] = \arg\min_{\vec{co}\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{B}}} \|\vec{b} - \vec{co}\|_{L_p} \Big/ co^{\mathcal{C}}_{L_p}[m_b] = \arg\min_{\vec{m}_{co}\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}} \|\vec{m} - \vec{m}_{co}\|_{L_p} \tag{7}$$

(see Figure 2), one for each maximal chain $\mathcal{C}$ of subsets of Θ. Then, the distance of b from

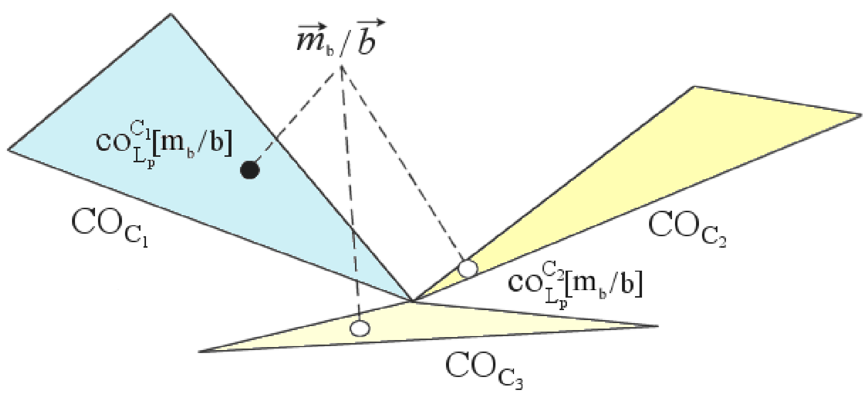

Fig. 2. To minimize the distance of a point from a simplicial complex, we need to find all the partial solutions (7) on all the maximal simplices in the complex (empty circles), to later compare these partial solutions to select a global optimum (black circle).

all such partial solutions has to be assessed in order to select a global optimal approximation. We start by analyzing the simple but interesting binary case (Figure 1). Some of its features are retained in the general case, others are not. Note also that, in the binary case, consonant and consistent [15] approximations coincide, and there is no difference between belief and mass space [18] representation.

III. Consonant approximation in the binary belief space

As we noticed above, in the binary case there is no distinction between approximation in the mass space or in the belief space, as those two representations coincide. Let us therefore consider the situation depicted in Figure 1, and find the probabilistic and consonant approximations of a b.f. $b \in \mathcal{B}_2$, using the classical norms (6).

In the Bayesian case

$$p_{L_2}[b] \doteq \arg\min_{p\in\mathcal{P}} \|\vec{b}-\vec{p}\|_{L_2} = \Big[m_b(x)+\frac{m_b(\Theta)}{2}, m_b(y)+\frac{m_b(\Theta)}{2}\Big]'.$$

This probability is called *orthogonal projection* $\pi[b]$ of b onto $\mathcal{P}$ [14], and coincides with the pignistic function $BetP[b]$ [27], [24], [28] only in the binary case. The L_∞ norm yields the same Bayesian approximation:

$$\begin{array}{rcl} p_{L_\infty}[b] & \doteq & \arg\min_{p\in\mathcal{P}} \|\vec{b}-\vec{p}\|_{L_\infty} = \arg\min_{p\in\mathcal{P}} \max\Big\{|b(x)-p(x)|, |b(y)-p(y)|\Big\} \\ & = & \arg\min_{p\in\mathcal{P}} \max\Big\{|m_b(x)-p(x)|, |m_b(y)-p(y)|\Big\} \\ & = & \Big[m_b(x)+\frac{m_b(\Theta)}{2}, m_b(y)+\frac{m_b(\Theta)}{2}\Big]' = p_{L_2}[b] \end{array}$$

while the optimization problem

$$\begin{array}{rcl} \arg\min_{p\in\mathcal{P}} \|\vec{b}-\vec{p}\|_{L_1} & = & \arg\min_{p\in\mathcal{P}}(|b(x)-p(x)|+|b(y)-p(y)|) \\ & = & \arg\min_{p\in\mathcal{P}}(|m_b(x)-p(x)|+|m_b(y)-p(y)|) \end{array}$$

has as solution the entire set of probabilities *compatible* with b [29], [30], i.e.

$$\{p\in\mathcal{P} : p(A)\geq b(A)\ \forall A\subseteq\Theta\}. \tag{8}$$

A. Consonant case

As depicted in Figure 2, in the consonant case we need to find a partial approximation on each component of the consonant complex, to later select a global approximation among those partial solutions. We get for L_2:

$$co_{L_2}[b] \doteq \arg\min_{co\in\mathcal{CO}} \|\vec{b}-\vec{co}\|_{L_2} = \begin{cases} \big[m_b(x), 0\big]' & m_b(x)\leq m_b(y) \\ \big[0, m_b(y)\big]' & m_b(x)\geq m_b(y) \end{cases} \tag{9}$$

(see Figure 1), while

$$\|\vec{b}-\vec{co}\|_{L_1} = |m_b(x)-m_{co}(x)| + |m_b(y)-m_{co}(y)| = |m_b(x)-m_{co}(x)| + m_b(y)$$

for $co \in \mathcal{CO}_x$. This is minimal for $m_{co}(x) = m_b(x)$ ($m_{co}(y) = 0$ by definition). Analogously for the component $\mathcal{CO}_y$

$$\arg\min_{co \in \mathcal{CO}_y} \|\vec{b} - \vec{co}\|_{L_1} = [0,\ m_b(y)]',$$

so that $co_{L_1}[b] = co_{L_2}[b]$ is again given by Equation (9).

The case of the L_∞ norm is more intriguing. For $co \in \mathcal{CO}_x$ the L_∞ distance between b and co is

$$\begin{aligned} \|\vec{b} - \vec{co}\|_{L_\infty} &= \max\Big\{|m_b(x) - m_{co}(x)|, |m_b(y) - m_{co}(y)|\Big\} \\ &= \max\Big\{|m_b(x) - m_{co}(x)|, m_b(y)\Big\}. \end{aligned}$$

Its minimum corresponds to all the consonant belief functions such that $|m_b(x) - m_{co}(x)| \leq m_b(y)$, i.e., $\arg\min_{co \in \mathcal{CO}_x} \|\vec{b} - \vec{co}\|_{L_\infty} =$

$$= \Big\{co \in \mathcal{CO}_x : \max\{0, m_b(x) - m_b(y)\} \leq co(x) \leq m_b(x) + m_b(y)\Big\},$$

and analogously for the $\mathcal{CO}_y$ component. Therefore we can write:

$$\arg\min_{co \in \mathcal{CO}_2} \|\vec{b} - \vec{co}\|_{L_\infty} = CO[b] \doteq$$
$$\doteq \begin{cases} \{co \in \mathcal{CO}_x : m_b(x) - m_b(y) \leq m_{co}(x) \leq m_b(x) + m_b(y)\}, & m_b(x) \geq m_b(y) \\ \{co \in \mathcal{CO}_y : m_b(y) - m_b(x) \leq m_{co}(y) \leq m_b(y) + m_b(x)\} & m_b(y) \geq m_b(x) \end{cases}$$

since when $m_b(x) \geq m_b(y)$ $\max\{0, m_b(x) - m_b(y)\} = m_b(x) - m_b(y)$, while when $m_b(y) \geq m_b(x)$ $\max\{0, m_b(y) - m_b(x)\} = m_b(y) - m_b(x)$.

It suffices to compare the expressions of the consonant approximations $co_{L_1}[b]$, $co_{L_2}[b]$, $co_{L_\infty}[b]$ of b to note that $co_{L_1}[b] = co_{L_2}[b]$ is the *center of mass* of the above set $CO[b]$ (see Figure 1 again): $co_{L_1}[b] = co_{L_2}[b] = \overline{CO}[b]$.

If we summarize the results we obtained,

$$\begin{array}{ll} c_{L_\infty}[b] = CO[b] & p_{L_1}[b] = P[b]; \\ c_{L_2}[b] = \overline{CO}[b] & p_{L_2}[b] = \overline{P}[b]; \\ c_{L_1}[b] = \overline{CO}[b] & p_{L_\infty}[b] = \overline{P}[b], \end{array}$$

we can recognize the dual role of the norms L_∞ and L_1 in the two problems (at least in the binary case). It is natural to call the set $CO[b]$ the collection of *consonant belief functions compatible with* b

B. Compatible consonant belief functions

We can try and give a characterization of compatible co.b.f.s in terms of an order relation similar to weak inclusion for $P[b]$ (8). Looking at Figure 3 we can note that, still in the binary

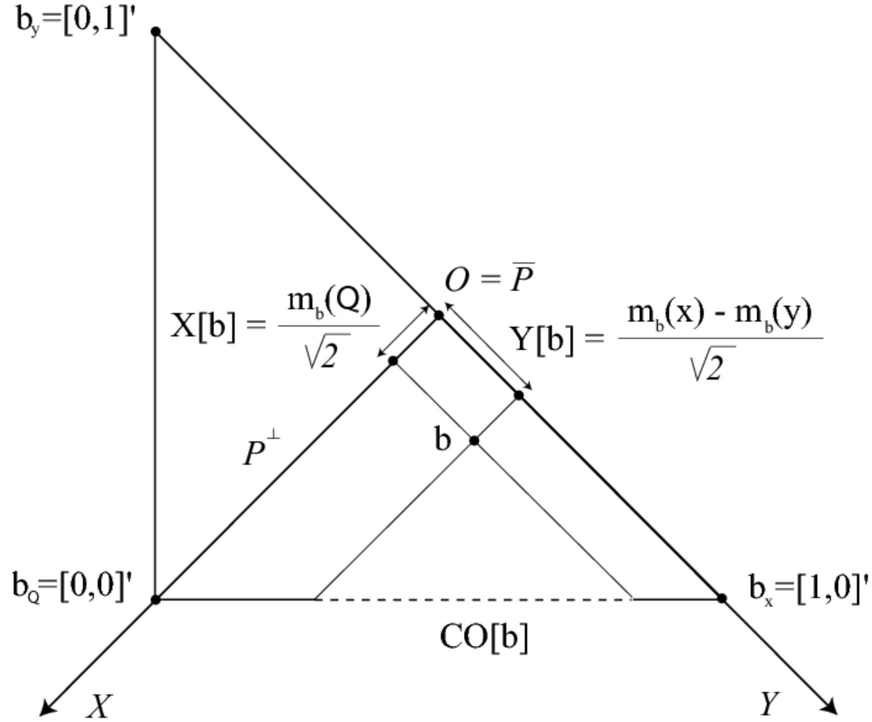

Fig. 3. Characterization of compatible consonant belief functions in terms of the reference frame (X, Y) formed by the probability line and the line $\mathcal{P}^{\perp}$ orthogonal to $\mathcal{P}$ in $\overline{\mathcal{P}} = [1/2, 1/2]'$.

case, $CO[b]$ is the set of co.b.f.s for which

$$X(c) \geq X(b), \qquad Y(c) \geq Y(b)$$

in the reference frame (X, Y) with origin $O = \overline{\mathcal{P}} = [1/2, 1/2]'$. The coordinates of a belief function b in this reference frame can be computed through simple trigonometric arguments, and are given by

$$X(b) = \frac{m_b(\Theta)}{\sqrt{2}}, \qquad Y(b) = \frac{m_b(x) - m_b(y)}{\sqrt{2}}.$$

More interestingly, it is easy to recognize from Figure 3 that $X(b)$ is nothing but the L_2 distance of b from the Bayesian region, while $Y(b)$ is the distance between b and the orthogonal complement of $\mathcal{P}$ in $\overline{\mathcal{P}}$:

$$X(b) = \|\vec{b} - \mathcal{P}\|_{L_2} \qquad Y(b) = \|\vec{b} - \mathcal{P}^{\perp}\|_{L_2}.$$

Furthermore, $\mathcal{P}^{\perp}$ (or better its segment $Cl(b_\Theta, \overline{\mathcal{P}})$ joining b_Θ and $\overline{\mathcal{P}}$) is the set of belief functions in which the mass is equally distributed among events of the same size ($\{x\}$ and $\{y\}$ in the binary case). This link between orthogonality and equidistribution is true in the general case too. In [31] we proved that:

Proposition 2: If a belief function b is such that

$$pl_b(x;k) \doteq \sum_{A \supset x, |A|=k} m_b(A) = const = pl_b(\cdot;k)$$

for all $k = 1, ..., |\Theta| - 1$ then $\vec{b}$ is orthogonal to the probabilistic subspace: $\vec{b} \perp \mathcal{P}$.
In conclusion we can claim that, at least in the binary case:

Theorem 1: The consonant belief functions compatible with $b \in \mathcal{B}_2$ are all the co.b.f.s which are at the same time less Bayesian and less equally distributed than b.
The possible existence of a set of compatible consonant belief functions in the general case is a very interesting topic of its own, which due to lack of space here we plan to analyze in a separate paper.

IV. Consonant approximation in the mass space

Let us compute the analytical form of all L_p consonant approximations in the mass space, in both its $\mathbb{R}^{N-1}$ and its $\mathbb{R}^{N-2}$ form (see Section II-B, Equations (2) and (3)). We start by describing the difference vector $\vec{m}_b - \vec{m}_{co}$ between the original mass vector and its approximation.

A. Difference vectors

If we choose the $N-1$-dimensional version of the mass space (see Equation (2)), the mass vector associated with an arbitrary consonant b.f. co with maximal chain of focal elements $\mathcal{C}$ reads as $\vec{m}_{co} = \sum_{A \in \mathcal{C}} m_{co}(A) \vec{m}_A$, so that:

$$\vec{m}_b - \vec{m}_{co} = \sum_{A \in \mathcal{C}} (m_b(A) - m_{co}(A)) \vec{m}_A + \sum_{A \notin \mathcal{C}} m_b(A) \vec{m}_A. \tag{10}$$

If we instead pick the $N-2$-dimensional version of the mass space (see Equation (3)), the mass vector associated with the same, arbitrary consonant b.f. co with maximal chain $\mathcal{C}$ reads as

$$\vec{m}_{co} = \sum_{A\in\mathcal{C}, A\neq\Theta} m_{co}(A)\vec{m}_A,$$

and the difference vector is:

$$\vec{m}_b - \vec{m}_{co} = \sum_{A\in\mathcal{C}, A\neq\Theta} (m_b(A) - m_{co}(A))\vec{m}_A + \sum_{A\not\in\mathcal{C}} m_b(A)\vec{m}_A. \quad (11)$$

B. L_1 approximation

1) $\mathbb{R}^{N-1}$ representation: Let us consider first the $\mathbb{R}^{N-1}$ representation of mass vectors. Given the difference vector (10) its L_1 norm is

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \sum_{A\in\mathcal{C}} |m_b(A) - m_{co}(A)| + \sum_{A\not\in\mathcal{C}} m_b(A) = \sum_{A\in\mathcal{C}} |\beta(A)| + \sum_{A\not\in\mathcal{C}} m_b(A)$$

expressed as a function of the variables $\{\beta(A), A\in\mathcal{C}, A\neq\Theta\}$, $\beta(A) \doteq m_b(A) - m_{co}(A)$. Here

$$\sum_{A\in\mathcal{C}} \beta(A) = \sum_{A\in\mathcal{C}} (m_b(A) - m_{co}(A)) = \sum_{A\in\mathcal{C}} m_b(A) - 1,$$

so that $\beta(\Theta) = \sum_{A\in\mathcal{C}} m_b(A) - 1 - \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A)$.

Therefore, the above norm reads as

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \Big|\sum_{A\in\mathcal{C}} m_b(A) - 1 - \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A)\Big| + \sum_{A\in\mathcal{C}, A\neq\Theta} |\beta(A)| + \sum_{A\not\in\mathcal{C}} m_b(A). \quad (12)$$

Partial approximation. The norm (12) is a function of the form

$$\sum_i |x_i| + \Big| - \sum_i x_i - k\Big|, \quad k \geq 0 \quad (13)$$

which has an entire simplex of minima, namely: $x_i \leq 0 \ \forall i$, $\sum_i x_i \geq -k$ (see [18] for a similar optimization problem in the geometric conditioning context). See Figure 4 for the case of two variables, x_1 and x_2 (corresponding to the case of a maximal chain of just three elements, i.e. $|\Theta| = n = 3$). The minima of the L_1 norm (12) are therefore given by the following system of constraints:

$$\begin{cases} \beta(A) \leq 0 & \forall A\in\mathcal{C}, A\neq\Theta, \\ \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A) \geq \sum_{A\in\mathcal{C}} m_b(A) - 1. \end{cases} \quad (14)$$

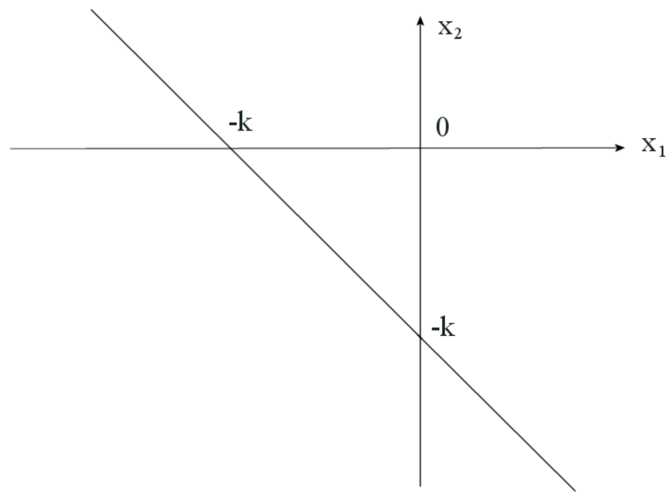

Fig. 4. The minima of a function of the form (13) with two variables x_1, x_2 form the triangle $x_1 \leq 0$, $x_2 \leq 0$, $x_1 + x_2 \geq -k$ depicted here.

The solution in terms of the mass of the consonant approximation reads as:

$$\begin{cases} m_{co}(A) \geq m_b(A) & \forall A \in \mathcal{C}, A \neq \Theta, \\ \displaystyle\sum_{A \in \mathcal{C}, A \neq \Theta} (m_b(A) - m_{co}(A)) \geq \sum_{A \in \mathcal{C}} m_b(A) - 1, & \end{cases} \tag{15}$$

where the last constraint reduces to

$$\sum_{A \in \mathcal{C}, A \neq \Theta} (m_b(A) - m_{co}(A)) = \sum_{A \in \mathcal{C}, A \neq \Theta} m_b(A) - 1 + m_{co}(\Theta) \geq \sum_{A \in \mathcal{C}} m_b(A) - 1,$$

i.e., $m_{co}(\Theta) \geq m_b(\Theta)$. Therefore the solution is

$$m_{co}(A) \geq m_b(A) \quad \forall A \in \mathcal{C}.$$

Vertices and barycenter of the partial approximation. The vertices of the set of approximations which are the solutions of (14) are (compare Figure 4) given by the vectors of variables $\{\vec{\beta}_A, A \in \mathcal{C}\}$ such that

$$\vec{\beta}_A(B) = \begin{cases} -\sum_{A \notin \mathcal{C}} m_b(A) & B = A, \\ 0 & B \neq A \end{cases}$$

when $A \neq \Theta$, while $\vec{\beta}_\Theta = \vec{0}$. In terms of masses the vertices of the se of partial L_1 approximations are the vectors $\{\vec{m}^{L_1}_A[m_b], A \in \mathcal{C}\}$ such that

$$\vec{m}^{L_1}_A[m_b](B) = \begin{cases} m_b(B) + \sum_{A \notin \mathcal{C}} m_b(A) & B = A, \\ m_b(B) & B \neq A \end{cases} \tag{16}$$

whose barycenter is

$$co_{\overline{L_1},N-1}[b](B) = m_b(B) + \frac{\sum_{A\not\in\mathcal{C}} m_b(A)}{n}.$$

Global approximation. To find the *global* L_1 consonant approximation(s) on the consonant complex, we need to understand what component is associated with the minimal L_1 distance. All the partial approximations (15) onto $\mathcal{CO}^{\mathcal{C}}$ have L_1 distance from $\vec{m}_b$:

$$\sum_{A\not\in\mathcal{C}} m_b(A) = 1 - \sum_{A\in\mathcal{C}} m_b(A). \quad (17)$$

Therefore, the minimal distance component of the complex is that associated with the maximal chain $\mathcal{C}^*$ that has maximal mass, according to the belief function b to approximate:

$$\mathcal{C}^* = \arg\max_{\mathcal{C}} \sum_{A\in\mathcal{C}} m_b(A).$$

2) $\mathbb{R}^{N-2}$ representation: **Partial approximation.** Consider now the $\mathbb{R}^{N-2}$ representation of mass vectors. Given the difference vector (11) its L_1 norm is

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \sum_{A\in\mathcal{C},A\neq\Theta} |m_b(A) - m_{co}(A)| + \sum_{A\not\in\mathcal{C}} m_b(A)$$

which is obviously minimized by

$$m_{co}(A) = m_b(A) \quad \forall A\in\mathcal{C}, A\neq\Theta, \quad (18)$$

while the mass of Θ can be obtained by normalization:

$$m_{co}(\Theta) = 1 - \sum_{A\in\mathcal{C},A\neq\Theta} m_{co}(A) = m_b(\Theta) + 1 - \sum_{A\in\mathcal{C}} m_b(A).$$

Global approximation. As it is the case for $\mathbb{R}^{N-1}$, to find the global L_1 approximation(s) on the consonant complex in $\mathbb{R}^{N-2}$, we need to find the closest component. The (unique) partial approximation (18) onto $\mathcal{CO}^{\mathcal{C}}$ has L_1 distance from $\vec{m}_b$ given by (17), as in the previous case. Therefore, the minimal distance component of the consonant complex is once again that associated with the maximal chain $\mathcal{C}^*$ such that $\mathcal{C}^* = \arg\max_{\mathcal{C}} \sum_{A\in\mathcal{C}} m_b(A)$.
In conclusion:

Theorem 2: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, the global L_1 consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-1}$ is the set of partial approximations

$$co^{\mathcal{C}^*}_{L_1,\mathcal{M},N-1}[m_b] \;\; = \;\; \Big\{ m_{co}(A) \geq m_b(A) \; \forall A\in\mathcal{C} \Big\} = Cl(\vec{m}^{L_1}_A[m_b], A\in\mathcal{C}),$$

with vertices given by Equation (16), associated with the maximal chain of focal elements which maximizes the total original mass of the chain

$$\mathcal{C}^* = \arg\max_{\mathcal{C}} \sum_{A\in\mathcal{C}} m_b(A).$$

The global L_1 consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$ form the set of (unique) partial approximations $co^{\mathcal{C}^*}_{L_1,\mathcal{M},N-2}[m_b]$ such that

$$\begin{cases} m_{co}(A) = m_b(A) & \forall A \in \mathcal{C}^*, A \neq \Theta, \\ m_{co}(\Theta) = m_b(\Theta) + 1 - \sum_{A\in\mathcal{C}^*} m_b(A) \end{cases}$$

associated with the same chain(s) of focal elements.

Not only the two approximations are consistent (in the sense that they have the same chain of focal elements), but it is easy to see that the set of L_1 consonant approximations in $\mathbb{R}^{N-1}$ is convex and forms a polytope, one of whose vertices is indeed the L_1 approximation in $\mathbb{R}^{N-2}$.

C. L_2 approximation

In order to find the L_2 consonant approximation 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 such that the difference vector is orthogonal to all the generators $\vec{g}_i$ of the vector space:

$$\arg\min_{\vec{q}\in V} \|\vec{p} - \vec{q}\|_{L_2} = \hat{q} \in V : \langle \vec{p} - \hat{q}, \vec{g}_i \rangle = 0 \quad \forall i$$

whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$ (Figure 5).

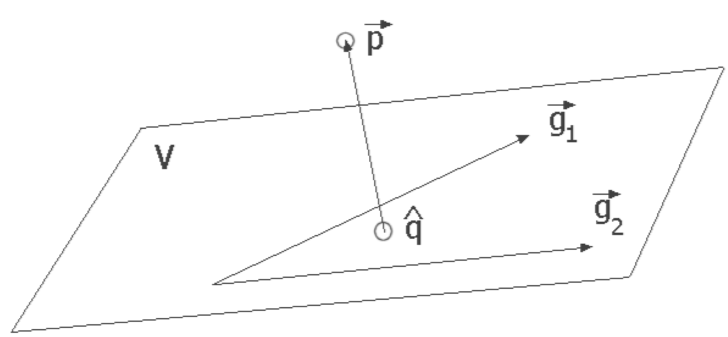

Fig. 5. The point $\vec{q}$ of a vector space V of minimal L_2 from a given point $\vec{p}$ external to it is such that the difference vector $\vec{p} - \vec{q}$ is orthogonal to all the generators $\vec{g}_i$ of V.

Instead of minimizing the L_2 norm of the difference vector $\|\vec{m}_b - \vec{m}_{co}\|_{L_2}$ we can just impose a condition of orthogonality between the difference vector itself $\vec{m}_b - \vec{m}_{co}$ and each component $\mathcal{CO}^{\mathcal{C}}$ of the consonant complex.

As the generators of $\mathcal{CO}^{\mathcal{C}}$ are the vectors in $\mathcal{M}$: $\{\vec{m}_A - \vec{m}_\Theta, A \in \mathcal{C}, A \neq \Theta\}$ we need to write:

$$\langle \vec{m}_b - \vec{m}_{co}, \vec{m}_A - \vec{m}_\Theta \rangle = 0 \quad \forall A \in \mathcal{C}, A \neq \Theta. \tag{19}$$

1) $\mathbb{R}^{N-1}$ representation: In the $N-1$ dimensional mass space the vector $\vec{m}_A - \vec{m}_\Theta$ is such that

$$\vec{m}_A - \vec{m}_\Theta(B) = \begin{cases} 1 & B = A \\ -1 & B = \Theta \\ 0 & B \neq A, \Theta. \end{cases}$$

Hence, the orthogonality condition becomes $\beta(A) - \beta(\Theta) = 0$ for all $A \in \mathcal{C}, A \neq \Theta$.

Partial approximation. To express $\beta(\Theta)$ in terms of the free variables we can use the following equality:

$$\sum_{A\in\mathcal{C}} \beta(A) = \beta(\Theta) + \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A) = \sum_{A\in\mathcal{C}} (m_b(A) - m_{co}(A)) = \sum_{A\in\mathcal{C}} m_b(A) - 1$$

so that $\beta(\Theta) = \sum_{A\in\mathcal{C}} m_b(A) - 1 - \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A)$ and the orthogonality condition becomes

$$\Big\{ 2\beta(A) + 1 - \sum_{B\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C}, B\neq A,\Theta} \beta(B) = 0$$

for all focal elements A in the maximal chain $\mathcal{C}$, $A \neq \Theta$. The solution is clearly $\beta(A) = \frac{\sum_{B\in\mathcal{C}} m_b(B) - 1}{n}$, so that the partial L_2 consonant approximation is s.t.

$$m_{co}(A) = m_b(A) + \frac{1 - \sum_{B\in\mathcal{C}} m_b(B)}{n} = m_b(A) + \frac{\sum_{B\notin\mathcal{C}} m_b(B)}{n}. \tag{20}$$

Global approximation. To find the global approximation, we need to compute the L_2 distance of b from the closest such partial solution. We have:

$$\begin{aligned} \|\vec{m}_b - \vec{m}_{co}\|^2_{L_2} &= \sum_{A\subseteq\Theta} (m_b(A) - m_{co}(A))^2 = \sum_{A\in\mathcal{C}} \Big(\frac{1 - \sum_{B\in\mathcal{C}} m_b(B)}{n} \Big)^2 + \\ &+ \sum_{A\notin\mathcal{C}} (m_b(A))^2 = \frac{(\sum_{B\notin\mathcal{C}} m_b(B))^2}{n} + \sum_{A\notin\mathcal{C}} (m_b(A))^2. \end{aligned}$$

which is minimized by the component $\mathcal{CO}^{\mathcal{C}}$ that minimizes $\sum_{A\notin\mathcal{C}} (m_b(A))^2$.

2) $\mathbb{R}^{N-2}$ representation: **Partial approximation.** In the case of the $\mathbb{R}^{N-2}$ representation, as $\vec{m}_\Theta = \vec{0}$, the orthogonality condition reads as:

$$\langle \vec{m}_b - \vec{m}_{co}, \vec{m}_A \rangle = \beta(A) = 0 \quad \forall A \in \mathcal{C}, A \neq \Theta$$

so that the L_2 partial consonant approximation of b is given by

$$\begin{cases} m_{co}(A) = m_b(A) & A \in \mathcal{C}, A \neq \Theta \\ m_{co}(\Theta) = m_b(\Theta) + \sum_{B \notin \mathcal{C}} m_b(B). \end{cases} \tag{21}$$

Global approximation. The optimal distance is, in this case,

$$\begin{aligned} \|\vec{m}_b - \vec{m}_{co}\|^2_{L_2} &= \sum_{A \subseteq \Theta} (m_b(A) - m_{co}(A))^2 = \sum_{A \in \mathcal{C}, A \neq \Theta} (m_b(A) - m_b(A))^2 + \\ &\quad + \sum_{A \notin \mathcal{C}} (m_b(A))^2 + \Big(m_b(\Theta) - m_b(\Theta) - \sum_{A \notin \mathcal{C}} m_b(A) \Big)^2 \\ &= \sum_{A \notin \mathcal{C}} (m_b(A))^2 + \Big(\sum_{A \notin \mathcal{C}} m_b(A) \Big)^2 \end{aligned}$$

which is once again minimized by the maximal chain $\mathcal{C}^* = \arg\min_{\mathcal{C}} \sum_{A \notin \mathcal{C}} (m_b(A))^2$.

Theorem 3: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, the global L_2 consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-1}$ form the collection of partial approximations

$$co^{\mathcal{C}^*}_{L_2, \mathcal{M}, N-1}[m_b] = \Big\{ m_{co}(A) = m_b(A) + \frac{1 - \sum_{B \in \mathcal{C}^*} m_b(B)}{n} \Big\}$$

associated with the maximal chain $\mathcal{C}^*$ of focal elements which minimizes the sum of square masses outside the chain:

$$\mathcal{C}^* = \arg\min_{\mathcal{C}} \sum_{A \notin \mathcal{C}} (m_b(A))^2.$$

The global L_2 consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$ are the union of the (unique) partial approximations $co^{\mathcal{C}^*}_{L_1, \mathcal{M}, N-2}[m_b] =$

$$= \Big\{ m_{co}(A) = m_b(A) \quad \forall A \in \mathcal{C}, A \neq \Theta, m_{co}(\Theta) = m_b(\Theta) + 1 - \sum_{A \in \mathcal{C}} m_b(A) \Big\}$$

associated with the same chain of focal elements, and coincides with the global L_1 consonant approximation in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$.

Indeed, in virtue of (21) and (18) all the partial L_1 and L_2 consonant approximations coincide in the mass space of dimension $N-2$.

D. L_∞ approximation

1) $\mathbb{R}^{N-1}$ representation: In the $N-1$ representation, the L_∞ norm of the difference vector is

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} = \max\Big\{ \max_{A\in\mathcal{C}} |\beta(A)|, \max_{B\not\in\mathcal{C}} m_b(B)\Big\}$$

while $\beta(\Theta) = \sum_{B\in\mathcal{C}} m_b(B) - 1 - \sum_{B\in\mathcal{C}, B\neq\Theta} \beta(B)$, so that

$$|\beta(\Theta)| = \Big| \sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B)\Big|$$

and the norm to minimize becomes $\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} =$

$$\max\Big\{ \max_{A\in\mathcal{C},A\neq\Theta} |\beta(A)|, \Big|\sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B)\Big|, \max_{B\not\in\mathcal{C}} m_b(B)\Big\}. \tag{22}$$

This is a function of the form

$$\max\Big\{|x_1|, |x_2|, |x_1+x_2+k_1|, k_2\Big\} \tag{23}$$

with $0 \leq k_2 \leq k_1 \leq 1$. Such a function has two possible behaviors in terms of its minimal region in the plane x_1, x_2.

Case 1. If $k_1 \leq 3k_2$ its contour function has the form rendered in Figure 6. The set of minimal points is given by $x_i \geq -k_2$, $x_1 + x_2 \leq k_2 - k_1$. In the more general case of an arbitrary number $m-1$ of variables $x_1, ..., x_{m-1}$ such that $x_i \geq -k_2$, $\sum_i x_i \leq k_2 - k_1$, the set of minimal points is a simplex with m vertices: each vertex v^i is such that $v^i(j) = -k_2 \ \forall j \neq i$; $v^i(i) = -k_1 + (m-1)k_2$ (obviously $v^m = [-k_2, \cdots, -k_2]$). For the norm (22), in the first case

$$\max_{B\not\in\mathcal{C}} m_b(B) \geq \frac{1}{n} \sum_{B\not\in\mathcal{C}} m_b(B) \tag{24}$$

the set of partial L_∞ approximations is given by the following system:

$$\begin{cases} \beta(A) \geq -\max_{B\not\in\mathcal{C}} m_b(B) & A\in\mathcal{C}, A\neq\Theta \\ \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B) \leq \max_{B\not\in\mathcal{C}} m_b(B) - \sum_{B\not\in\mathcal{C}} m_b(B) \end{cases}$$

This determines a simplex $Cl(\vec{m}^{L_\infty}_{\bar{A}}[m_b], \bar{A}\in\mathcal{C})$ with vertices

$$\begin{cases} \beta_{\bar{A}}(A) = -\max_{B\not\in\mathcal{C}} m_b(B) & A\in\mathcal{C}, A\neq\bar{A} \\ \beta_{\bar{A}}(\bar{A}) = -\sum_{B\not\in\mathcal{C}} m_b(B) + (n-1)\max_{B\not\in\mathcal{C}} m_b(B) \end{cases}$$

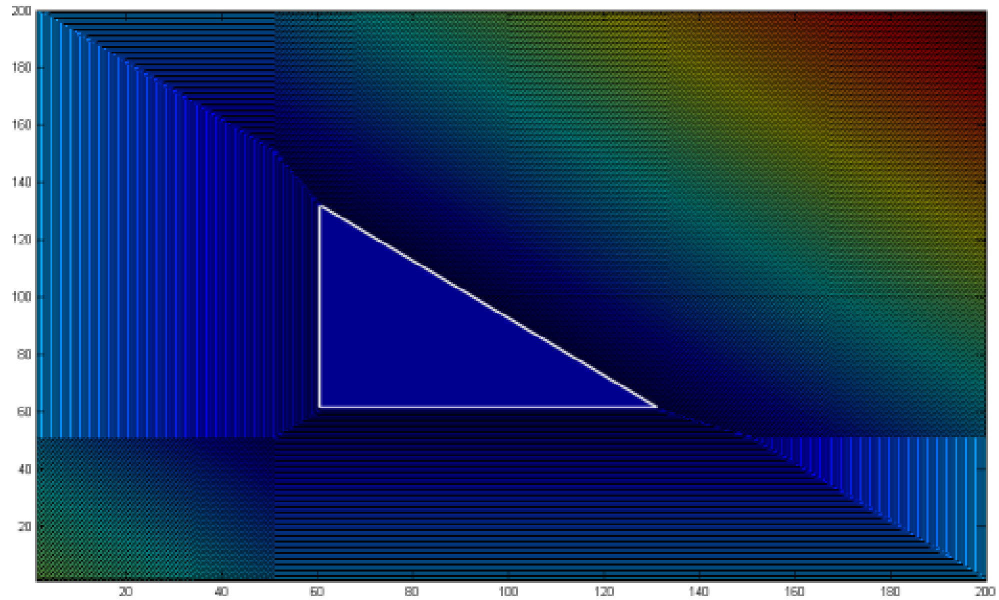

Fig. 6. Contour function (level sets) and minimal points (white triangle) of a function of the form (23), when $k_1 \leq 3k_2$. In the example $k_2 = 0.4$ and $k_1 = 0.5$.

or, in terms of their basic probability assignments,

$$\begin{cases} \vec{m}^{L_\infty}_{\bar{A}}[m_b](A) = m_b(A) + \max_{B \notin \mathcal{C}} m_b(B) & A \in \mathcal{C}, A \neq \bar{A} \\ \vec{m}^{L_\infty}_{\bar{A}}[m_b](\bar{A}) = m_b(\bar{A}) + \sum_{B \notin \mathcal{C}} m_b(B) - (n-1) \max_{B \notin \mathcal{C}} m_b(B). \end{cases} \tag{25}$$

Note that such quantity is not guaranteed to be positive. The barycenter of the above simplex can be computed as follows:

$$m^{\overline{L_\infty}}(A) = \frac{\sum_{\bar{A} \in \mathcal{C}} \vec{m}^{L_\infty}_{\bar{A}}[m_b](A)}{n} = \frac{n m_b(A) + \sum_{B \notin \mathcal{C}} m_b(B)}{n} = m_b(A) + \frac{\sum_{B \notin \mathcal{C}} m_b(B)}{n},$$

i.e., the L_2 partial approximation (20). The corresponding minimal L_∞ norm of the difference vector is, according to (22), equal to $\max_{B \notin \mathcal{C}} m_b(B)$.

Case 2. In the second case $k_1 > 3k_2$, i.e., for our problem,

$$\max_{B \notin \mathcal{C}} m_b(B) < \frac{1}{n} \sum_{B \notin \mathcal{C}} m_b(B), \tag{26}$$

the contour function of (23) is as in Figure 7. There is a single minimal point, located in $[-1/3k_1, -1/3k_1]$. For an arbitrary number $m-1$ of variables the minimal point is located in

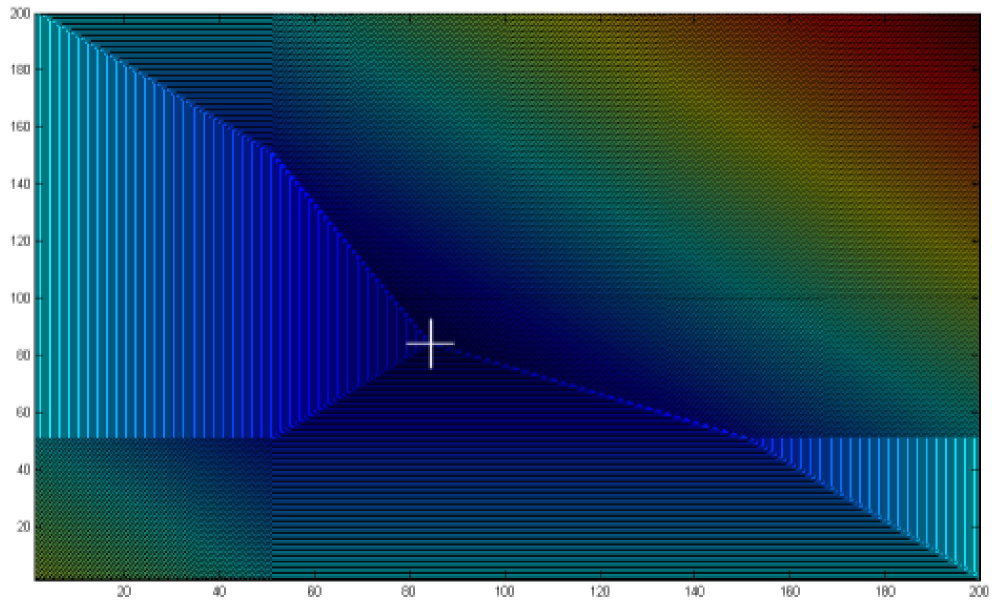

Fig. 7. Contour function (level sets) and minimal point (white cross) of a function of the form (23), when $k_1 \geq 3k_2$. In this example $k_2 = 0.1$ and $k_1 = 0.5$.

$[(-1/m)k_1, \cdots, (-1/m)k_1]'$, i.e., for system (22),

$$\beta(A) = -\frac{1}{n}\sum_{B \notin \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta$$

or, in terms of basic probability assignments,

$$m_{co_{L_\infty}[m_b]}(A) = m_b(A) + \frac{1}{n}\sum_{B \notin \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}.$$

The mass of Θ is obtained by normalization.

The corresponding minimal L_∞ norm of the difference vector is $\frac{1}{n}\sum_{B \notin \mathcal{C}} m_b(B)$.

2) $\mathbb{R}^{N-2}$ representation: In the $N-2$ representation, the L_∞ norm of the difference vector is

$$\begin{array}{rcl} \|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} & = & \max_{\emptyset \subsetneq A \subsetneq \Theta} |m_b(A) - m_{co}(A)| \\ & = & \max\left\{ \max_{A \in \mathcal{C}, A \neq \Theta} |\beta(A)|, \max_{B \notin \mathcal{C}} m_b(B) \right\} \end{array} \tag{27}$$

which is minimized by

$$|\beta(A)| \leq \max_{B \notin \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta \tag{28}$$

i.e., in the original mass coordinates,

$$m_b(A) - \max_{B \not\in \mathcal{C}} m_b(B) \leq m_{co}(A) \leq m_b(A) + \max_{B \not\in \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta. \tag{29}$$

According to (27) the corresponding minimal L_∞ norm is: $\max_{B \not\in \mathcal{C}} m_b(B)$.

Vertices and barycenter. Clearly, the vertices of the set (28) are all the vectors of β variables such that $\beta(A) = +/- \max_{B \not\in \mathcal{C}} m_b(B)$ for all $A \in \mathcal{C}$, $A \neq \Theta$. Its barycenter is clearly given by $\beta(A) = 0$ for all $A \in \mathcal{C}$, $A \neq \Theta$, i.e.:

$$m_{co}(B) = \begin{cases} m_b(B) & B \in \mathcal{C}, B \neq \Theta \\ m_b(B) + \sum_{B \not\in \mathcal{C}} m_b(B) & B = \Theta. \end{cases} \tag{30}$$

Summarizing:

Theorem 4: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, the partial L_∞ consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-1}$ can form either a simplex

$$co^{\mathcal{C}^*}_{L_\infty, \mathcal{M}, N-1}[m_b] = Cl(\vec{m}^{L_\infty}_{\bar{A}}[m_b], \bar{A} \in \mathcal{C})$$

with vertices (25) when

$$\max_{B \not\in \mathcal{C}} m_b(B) \geq \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B),$$

or a reduce to a single belief function when the opposite is true, the barycenter of the above simplex, located on the partial L_2 approximation (20).

In both cases, the global L_∞ consonant approximation is associated with the maximal chain(s) of focal elements:

$$\mathcal{C}^* = \arg\min_{\mathcal{C}} \max_{B \not\in \mathcal{C}} m_b(B). \tag{31}$$

The partial L_∞ consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$ form the set $co^{\mathcal{C}}_{L_\infty, \mathcal{M}, N-2}[m_b]$ given by Equation (29). Its barycenter reassigns all the mass originally outside the desired chain $\mathcal{C}$ to Θ, leaving the masses of the other elements of the chain untouched. The related global approximations of b are associated with the same optimal chain(s) as in (31).

E. Semantics of consonant approximations in $\mathcal{M}$

Let us interpret the results we obtained in terms of the basic probability assignments of the various consonant approximations, and compare those results with the outer consonant approximations [3] whose geometry has been described in [13].

1) Summary of L_p consonant approximation in $\mathcal{M}$: $N-1$ **representation.** Let us summarize all the results obtained so far. In the $\mathbb{R}^{N-1}$ mass representation the partial L_p approximations are:

$$\begin{array}{lllll} co^{\mathcal{C}}_{L_1,N-1}[m_b] & = & Cl(\vec{m}^{L_1}_{A}[m_b], A \in \mathcal{C}) & : & m_{co}(A) \geq m_b(A) \; \forall A \in \mathcal{C}; \\ co^{\mathcal{C}}_{\overline{L_1},N-1}[m_b] & = & co^{\mathcal{C}}_{L_2,N-1}[m_b] & : & m_{co}(A) = m_b(A) + \frac{\sum_{B \not\in \mathcal{C}} m_b(B)}{n}. \end{array} \tag{32}$$

Concerning the L_∞ approximation, if (24) holds

$$co^{\mathcal{C}}_{L_\infty,N-1}[m_b] = Cl(\vec{m}^{L_\infty}_{\bar{A}}, \bar{A} \in \mathcal{C}), \quad co^{\mathcal{C}}_{\overline{L_\infty},N-1}[m_b] : m_{co}(A) = \frac{1}{n} \; \forall A \in \mathcal{C}$$

while if (26) holds: $co^{\mathcal{C}}_{L_\infty,N-1}[m_b] = co^{\mathcal{C}}_{L_2,N-1}[m_b]$.

We can observe the following facts:

1) the set of L_1 partial approximation is the set of inner consonant approximations of b according to the order relation $b \geq b'$ iff $m_b(A) \geq m_{b'}(A)$;
2) this set is a simplex, whose vertices are obtained by re-assigning all the mass outside the desired chain to a single focal element of the chain itself (see (16));
3) its barycenter coincides with the L_2 partial approximation;
4) such approximation redistributes the mass of focal elements outside the chain on an equal basis to all the elements of the chain;
5) when the partial L_∞ approximation is unique, it coincides with the L_2 approximation and the barycenter of the L_1 approximations;
6) when it is not unique, is a simplex whose vertices assign to each element of the chain but one the maximal mass outside the chain, with barycenter still in the L_2 approximation.

In particular, points (2) and (4) (and (5)) remind us of the behavior of geometric conditional belief functions in the mass space [18]. There:

Proposition 3: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ and an arbitrary non-empty focal element $\emptyset \subsetneq A \subseteq \Theta$, the unique L_2 conditional belief functions $b_{L_2,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is the b.f. whose b.p.a. redistributes the mass $1-b(A)$ to each focal element $B \subseteq A$ in an equal way.

The set of L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is a simplex whose vertices re-assign the mass $1-b(A)$ of focal elements not in the conditioning event A to a specific subset of A.

It is tempting to speculate that this be a consistent behavior of L_1 and L_2 minimization in the $\mathbb{R}^{N-1}$ representation of the mass space.

$N-2$ **representation.** In the $\mathbb{R}^{N-2}$ mass representation the partial L_p approximations are:

$$\begin{array}{rcl} co^{\mathcal{C}}_{L_\infty,N-2}[m_b] & : & |m_{co}(A) - m_b(A)| \leq \max_{B\not\in\mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta; \\ co^{\mathcal{C}}_{\overline{L_\infty,N-2}}[m_b] & = & co^{\mathcal{C}}_{L_1,N-2}[m_b] = co^{\mathcal{C}}_{L_2,N-2}[m_b] : \left\{ \begin{array}{l} m_{co}(A) = m_b(A), \ A \in \mathcal{C}, \neq \Theta \\ m_{co}(\Theta) = m_b(\Theta) + \sum_{B\not\in\mathcal{C}} m_b(B). \end{array} \right. \end{array} \tag{33}$$

We can notice a number of facts in this case too:

- the L_∞ (partial) approximation is not unique, and it falls entirely inside the simplex of admissible consonant b.f. only if each focal element in the desired chain has mass greater then all focal elements outside the chain: $m_b(A) \leq \max_{B\not\in\mathcal{C}} m_b(B)$;
- it forms a generalized rectangle in the mass space $\mathcal{M}$, whose size is determined by the largest mass outside the desired maximal chain;
- the L_1 and L_2 partial approximations are uniquely determined, and coincide with the barycenter of the set of L_∞ partial approximations;
- their semantic is straightforward: all the mass outside the chain is re-assigned to Θ, increasing the overall uncertainty of the belief state.

Clearly, approximations in the mass space do not take into account the contributions of focal elements outside the chain to the plausibility of elements of the chain. A similar phenomenon has been observed in the case of geometric conditioning of belief functions by L_p minimization [18].

2) Relation with outer consonant approximations: Let us first recall the main results on the geometry of outer consonant approximations [13].

Proposition 4: For each simplicial component $\mathcal{CO}_\mathcal{C}$ of the consonant space associated with any maximal chain of focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ the set of outer consonant approximation of any b.f. b is the convex closure $O_\mathcal{C}[b] = Cl(o^{\vec{B}}[b], \forall \vec{B})$ of the co.b.f.s with basic probabilities

$$m_{o^{\vec{B}}[b]}(B_i) = \sum_{A\subseteq\Theta:\vec{B}(A)=A_i} m_b(A), \tag{34}$$

each associated with an "assignment function" $\vec{B} : 2^\Theta \rightarrow \mathcal{C}$, $A \mapsto \vec{B}(A) \supseteq A$ which maps each event A to one of the elements of the chain containing it.

The points (34) are not guaranteed to be proper vertices of the polytope $O_\mathcal{C}[b]$, as some of them can be obtained as a convex combination of the others. The outer approximation produced by the permutation ρ of singletons which generates the desired chain $A_i = \{x_{\rho(1)}, \cdots, x_{\rho(i)}\}$, i.e.

$$m_{co^\rho}(A_i) = \sum_{B \subseteq A_i, B \not\subset A_{i-1}} m_b(B), \tag{35}$$

is an actual vertex of $O_\mathcal{C}[b]$, and corresponds to the *maximal* outer consonant approximation with maximal chain $\mathcal{C}$.

It can seen in Figure 1 that, in the binary case, such maximal outer approximation coincides with the (partial) $L_1 = L_2 = \overline{L}_\infty$ approximation in the $N-2$ representation. It looks unclear what the relationship should be in the general case. By Equation (15)) the partial L_1 approximations in $\mathbb{R}^{N-1}$ are such that $m_{co}(A) \geq m_b(A)$ for all $A \in \mathcal{C}$: *they are the opposite of outer consonant approximations, using the natural order relation between basic probabilities* (rather than belief values).

F. Ternary example

It can be useful to compare the different approximations in the toy case of a ternary frame, $\Theta = \{x, y, z\}$, to look for further insights. Let us assume that we want the consonant approximation to have maximal chain $\mathcal{C} = \{\{x\}, \{x, y\}, \Theta\}$.

Figure 8 illustrates the different partial consonant approximations in the simplex of consonant belief functions with focal element $\{\{x\}, \{x, y\}, \Theta\}$, for a belief function with masses

$$m_b(x) = 0.2, \quad m_b(y) = 0.3, \quad m_b(x, z) = 0.5 \tag{36}$$

We notice that the different simplices of L_p consonant approximations are distinct, with the $L_{1,N-1}$ one (red simplex) falling entire in the consonant simplex $Cl(\vec{m}_x, \vec{m}_{x,y}, \vec{m}_\Theta)$, while most of $L_{\infty,N-2}$ (green quadrangle) does not. It is interesting to note, though, they are not unrelated to each other: indeed, the $L_1/L_2/\overline{L_\infty}$ consonant approximation in $\mathbb{R}^{N-2}$ (green little square) is a vertex of the simplex of L_1 approximation in $N-1$.

Even though the case for the unique $L_{1,N-2}/L_{2,N-2}/\overline{L_{\infty,N-2}}$ and $\overline{L_{1,N-1}}$ approximations seems

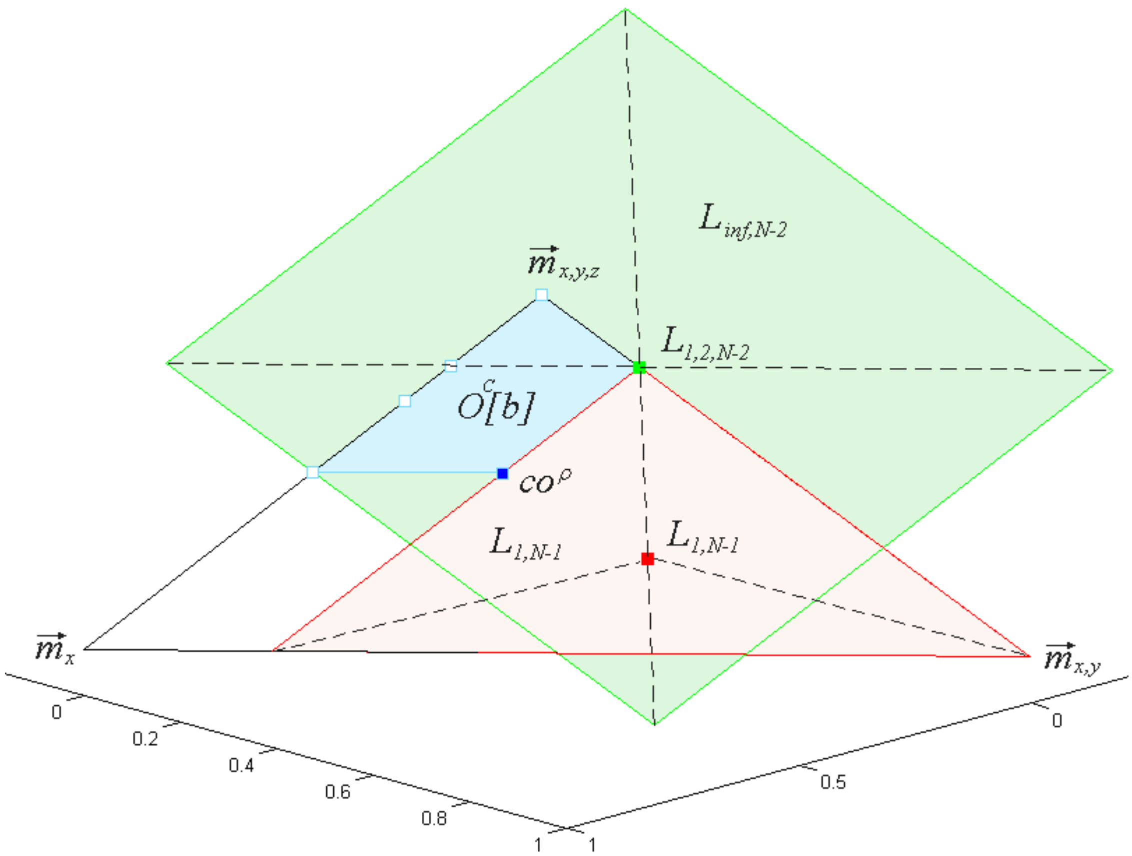

Fig. 8. The simplex (solid black triangle) of consonant belief functions with maximal chain $\{\{x\},\{x,y\},\Theta\}$, and the L_p partial consonant approximations in $\mathcal{M}$ of the belief function with basic probabilities (36) on $\Theta = \{x,y,z\}$. The related set $O^{\mathcal{C}}[b]$ of outer consonant approximations (37) is also shown.

compelling, it will be worth exploring in the near future the behavior of the intersection of the set of approximations not entirely admissible with the consonant complex.

According to the formulae at page 8 of [32], the set of outer consonant approximations of (36) with chain $\{\{x\},\{x,y\},\Theta\}$ is the closure of the points:

$$\begin{array}{lll} \vec{m}_{B_1,B_2} & = & [m_b(x), m_b(y), 1-m_b(x)-m_b(y)]', \\ \vec{m}_{B_3,B_4} & = & [m_b(x), 0, 1-m_b(x)]', \\ \vec{m}_{B_5,B_6} & = & [0, m_b(x)+m_b(y), 1-m_b(x)-m_b(y)]', \\ \vec{m}_{B_7,B_8} & = & [0, m_b(x), 1-m_b(x)]', \\ \vec{m}_{B_9,B_{10}} & = & [0, m_b(y), 1-m_b(y)]', \\ \vec{m}_{B_{11},B_{12}} & = & [0,0,1]', \end{array} \tag{37}$$

These points are plotted as light blue squares in Figure 8. We can notice again a number of interesting facts:

- the set $O^{\mathcal{C}}[b]$ of outer consonant approximations with chain $\mathcal{C}$ is a subset of (the admissible part of) the set of $L_{\infty,N-2}$ partial approximations; actually, the barycenter of the latter is a vertex of $O^{\mathcal{C}}[b]$;
- on the contrary, outer approximations and $L_{1,N-1}$ approximations are mutually exclusive, as it can be inferred by Equation (15);
- the maximal outer approximation co^{ρ} is on the border between the two, the line on which $m_{co}(x,y) = m_b(x,y)$.

Several other intriguing facts can be noticed there: they surely deserve further analysis.

V. Consonant approximation in the belief space

Consonant approximations in the mass space have quite natural semantics. As we see in this Section, L_p approximations in the belief space are of more complex interpretation. We will come back to this later.

A. Difference vector

In the belief space the original b.f. and the desired consonant approximation are written as $\vec{b} = \sum_{\emptyset \subsetneq A \subsetneq \Theta} b(A) X_A$, $\vec{b}_{co} = \sum_{A \supseteq A_1} \Big(\sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big) X_A$.
Their difference vector is

$$\begin{array}{rcl} \vec{b} - \vec{b}_{co} & = & \displaystyle\sum_{A \not\supset A_1} b(A) X_A + \sum_{A \supseteq A_1} X_A \Big[b(A) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] = \sum_{A \not\supset A_1} b(A) X_A + \\ & & \displaystyle + \sum_{A \supseteq A_1} X_A \Big[\sum_{\emptyset \subsetneq B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] = \sum_{A \not\supset A_1} b(A) X_A + \\ & & \displaystyle + \sum_{A \supseteq A_1} X_A \Big[\sum_{B \subseteq A, B \in \mathcal{C}} (m_b(B) - m_{co}(B)) + \sum_{B \subseteq A, B \not\in \mathcal{C}} m_b(B) \Big] \\ & = & \displaystyle\sum_{A \not\supset A_1} b(A) X_A + \sum_{A \supseteq A_1} X_A \Big[\gamma(A) + \sum_{B \subseteq A, B \not\in \mathcal{C}} m_b(B) \Big] \end{array} \tag{38}$$

once introduced the auxiliary variables

$$\gamma(A) = \sum_{B \subseteq A, B \in \mathcal{C}} \beta(B), \quad \beta(B) = m_b(A) - m_{co}(B).$$

We can therefore write

$$\vec{b} - \vec{b}_{co} = \sum_{A \not\supset A_1} b(A) X_A + \sum_{i=1}^{n-1} \sum_{A \supset A_i, A \not\supset A_{i+1}} X_A \Big[\gamma(A_i) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big], \tag{39}$$

as all the terms in (38) associated with subsets $A \supset A_i, A \not\supset A_{i+1}$ depend on the same auxiliary variable $\gamma(A_i)$, while the difference in the component X_Θ is trivially $1 - 1 = 0$.

B. L_∞ approximation

1) Partial approximation in each maximal simplex: **Problem formulation.** Given the expression (39) for the difference vector of interest in the belief space, we can compute the explicit form of its L_∞ norm as $\|\vec{b} - \vec{co}\|_\infty =$

$$\begin{array}{l} = \max \left\{ \max\limits_i \max\limits_{A \supseteq A_i, A \not\supset A_{i+1}} \Big| \gamma(A_i) + \sum\limits_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big|, \max\limits_{A \not\supset A_1} \Big| \sum\limits_{B \subseteq A} m_b(B) \Big| \right\} \\ = \max \left\{ \max\limits_i \max\limits_{A \supseteq A_i, A \not\supset A_{i+1}} \Big| \gamma(A_i) + \sum\limits_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big|, b(A_1^c) \right\}, \end{array} \tag{40}$$

as $\max_{A \not\supset A_1} \Big| \sum_{B \subseteq A} m_b(B) \Big| = b(A_1^c)$. Now, (40) can be minimized separately for each $i = 1, ..., n-1$. Clearly, the minimum is attained when the variable elements in (40) are not greater than the constant element $b(A_1^c)$:

$$\max_{A \supseteq A_i, A \not\supset A_{i+1}} \Big| \gamma(A_i) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big| \leq b(A_1^c). \tag{41}$$

The left hand side of (41) is a function of the form $\max\{|x + x_1|, ..., |x + x_n|\}$ (see Figure 9). In (41) there are $|\{A \supseteq A_i, A \not\supset A_{i+1}\}| = 2^{|A_{i+1}^c|}$ components of the form $x + x_i$ (an even number). In the binary case $\Theta = \{x, y\}$ the difference vector reads as

$$\vec{b} - \vec{b}_{co} = \sum_{A \supseteq \{x\}, A \not\supset \{x,y\}} X_A \Big[\gamma(A) - \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big] = [\gamma(x) - 0] \cdot X_x,$$

whose L_∞ norm is obviously minimized by $\gamma(x) = 0$. The partial consonant approximation of b in $\mathcal{CO}_x$ has therefore b.p.a. $m_{co}(x) = m_b(x)$, $m_{co}(\Theta) = 1 - m_b(x)$, as confirmed by Figure 1.

L_∞ **partial approximation in the variables** γ**.** In the general case, functions of the above form are minimized by $x = -\frac{x_{min} + x_{max}}{2}$ (see Figure 9 again). In the case of (41), such minimum

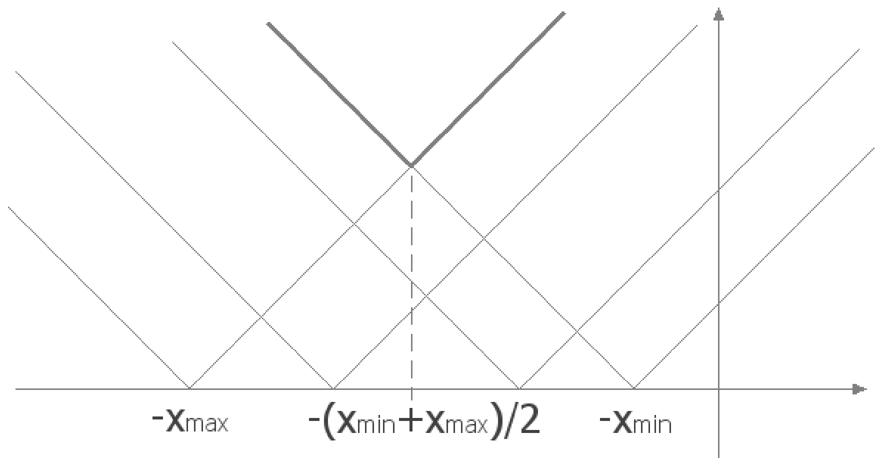

Fig. 9. The minimization of the L_∞ distance from the consonant subspace involves minimizing functions of the form $\max\{|x+x_1|, ..., |x+x_n|\}$ (in bold) depicted above.

and maximum offset values are, respectively,

$$\begin{aligned}\gamma_{min} &= \sum_{B\subseteq A_i, B\not\in\mathcal{C}} m_b(B) = b(A_i) - \sum_{j=1}^{i} m_b(A_j)\\ \gamma_{max} &= \sum_{B\subseteq (A_{i+1}\setminus A_i)^c, B\not\in\mathcal{C}} m_b(B) = b(\{x_{i+1}\}^c) - \sum_{j=1}^{i} m_b(A_j)\end{aligned}$$

once defined $\{x_{i+1}\} = A_{i+1}\setminus A_i$. As for each value of γ, $|\gamma(A_i)-\gamma|$ is dominated by either $|\gamma(A_i)-\gamma_{min}|$ or $|\gamma(A_i)-\gamma_{max}|$, the norm of the difference vector is minimized by the values of $\gamma(A_i)$ such that

$$\max\{|\gamma(A_i)-\gamma_{min}|, |\gamma(A_i)-\gamma_{max}|\} \leq b(A_1^c)$$

for all $i = 1, ..., n-1$, i.e.,

$$-\frac{\gamma_{min}+\gamma_{max}}{2} - b(A_1^c) \leq \gamma(A_i) \leq -\frac{\gamma_{min}+\gamma_{max}}{2} + b(A_1^c). \quad (42)$$

Barycenter of the L_∞ partial approximation. The barycenter of the set of L_∞ partial approximations lies therefore in

$$-\frac{\gamma_{min}+\gamma_{max}}{2} = -\frac{b(A_i)+b(\{x_{i+1}\}^c)}{2} + \sum_{j=1}^{i} m_b(A_j). \quad (43)$$

In terms of basic belief assignment, simple maths lead to the following formulae:

$$\begin{array}{rcl} m_{co}(A_1) & = & \dfrac{b(A_1)+b(\{x_2\}^c)}{2}; \\ m_{co}(A_i) & = & \dfrac{b(A_i)+b(\{x_{i+1}\}^c)}{2} - \dfrac{b(A_{i-1})+b(\{x_i\}^c)}{2} \\ & = & \dfrac{b(A_i)+b(A_i+A_{i+1}^c)}{2} - \dfrac{b(A_{i-1})+b(A_{i-1}+A_i^c)}{2} \qquad i=2,...,n-1, \end{array} \tag{44}$$

while $m_{co}(A_n) = m_{co}(\Theta) = 1 - \sum_{i=2}^{n-1} \Big[\frac{b(A_i)+b(\{x_{i+1}\}^c)}{2} - \frac{b(A_{i-1})+b(\{x_i\}^c)}{2} \Big] - \frac{b(A_1)+b(\{x_2\}^c)}{2} = 1 - b(A_{n-1})$.

Theorem 5: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, the barycenter of the set of partial L_∞ consonant approximations $co^{\mathcal{C}}_{L_\infty}[b]$ of b (in the belief space) onto a simplicial component $\mathcal{CO}_{\mathcal{C}}$ has b.p.a. given by Equation (44).

Semantics of L_∞ approximation in $\mathcal{B}$. The partial L_∞ consonant approximation (44) has a rather interesting form: it is the difference of two positive vectors, on of which is the "shifted" version of the other. This vector

$$\Big[\frac{b(A_i)+b(A_i+A_{i+1}^c)}{2}, i=1,...,n-1 \Big]'$$

measures the average between the belief value of the given element of the chain and the belief value you would obtain if x_{i+1} was the last element of Θ. Clearly, such a barycenter is not guaranteed to be a proper belief function, as its b.b.a. can be negative. However, an interpretation in terms of degrees of belief is possible when we note that the mass of the focal element A_i under the barycenter of $co^{\mathcal{C}}_{L_\infty}[b]$ can be also written as

$$m_{co}(A_i) = \frac{b(A_i)-b(A_{i-1})}{2} + \frac{pl_b(\{x_i\}) - pl_b(\{x_{i+1}\})}{2}.$$

This is proportional to the backward difference $b(A_i) - b(A_{i-1})$ between the belief values of two consecutive focal elements in the desired chain (which is always positive) plus the forward difference $pl_b(\{x_i\}) - pl_b(\{x_{i+1}\})$ between the plausibility values of two consecutive singletons (which can be positive or negative). The first addendum in itself is a sort of "derivative" of the original belief function on the desired chain. The second addendum is some sort of "derivative" of the plausibility distribution $pl_b(x)$ or contour function w.r.t. the desired order between singletons. This fact deserves to be further investigated in a subsequent research report.

Global solution. To compute the global L_∞ approximation of the original belief function b, we need to detect as usual the partial solution whose L_∞ distance from b is the smallest. Clearly, such (partial) optimal distance is, for each given component $\mathcal{CO}_\mathcal{C}$ of the consonant complex with $\mathcal{C} = \{A_1 \subset \cdots \subset A_n = \Theta\}$, equal to $b(A_1^c)$ (see Equation 41). Therefore the global L_∞ consonant approximation of b in the belief space $\mathcal{B}$ is the partial solution associated with the chain of focal elements such that

$$\arg\min_{\mathcal{C}} b(A_1^c) = \arg\min_{\mathcal{C}} 1 - pl_b(A_1) = \arg\max_{\mathcal{C}} pl_b(A_1).$$

Consequently, the partial solutions for all the chains which have the same singleton $A_1 = \{x\}$ as the smallest focal element are equally optimal.

Theorem 6: Given a belief function $b : 2^\Theta \to [0,1]$, the set of global L_∞ consonant approximations of b in the belief space is the collection of partial approximations $co^{\mathcal{C}}_{L_\infty}[b]$ onto the simplicial components $\mathcal{CO}_\mathcal{C}$ associated with chains of focal elements whose smallest f.e. is the maximal plausibility element of Θ:

$$co_{L_\infty}[b] = \bigcup_{\mathcal{C}:A_1 = \arg\max_x pl_b(x)} co^{\mathcal{C}}_{L_\infty}[b].$$

C. L_1 approximation

Recalling the expression (39) of the difference vector $\vec{b} - \vec{co}$ in the belief space, the latter's L_1 norm reads as:

$$\|\vec{b} - \vec{co}\|_{L_1} = \sum_{i=1}^{n-1} \sum_{A \supseteq A_i, A \not\supseteq A_{i+1}} \Big| \gamma(A_i) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big| + \sum_{A \not\supseteq A_1} |b(A)|. \quad (45)$$

Once again, it can be decomposed into a number of summations which depend on a single auxiliary variable $\gamma(A_i)$. Such components are of the form $|x + x_1| + ... + |x + x_n|$, with an even number of "nodes" x_i.

Partial solution in terms of nodes. To understand where the values which minimize such a function lie, let us consider the simple function of Figure 10. It is easy to see that these functions are minimized by the interval of values comprised between their two innermost nodes. As a consequence:

Theorem 7: Given a belief function $b : 2^\Theta \to [0,1]$, the set of partial L_1 consonant approximations $co^{\mathcal{C}}_{L_1}[b]$ of b (in the belief space) onto the simplicial component $\mathcal{CO}_\mathcal{C}$ is given by the

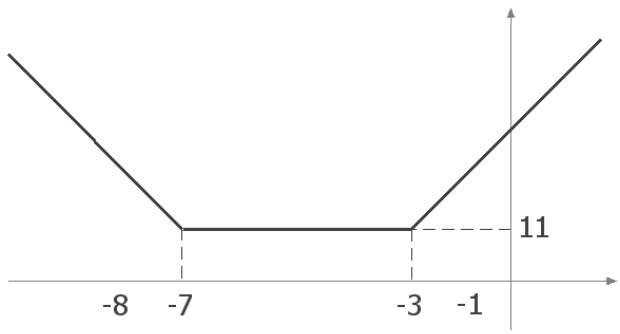

Fig. 10. The minimization of the L_1 distance from the consonant subspace involves minimizing functions such as the one depicted above, $|x+1|+|x+3|+|x+7|+|x+8|$, which is minimized for $3 \leq x \leq 7$.

following intervals in the auxiliary variables $\gamma(A_i)$:

$$\gamma^i_{int1} \leq \gamma(A_i) \leq \gamma^i_{int2}, \quad i = 1, ..., n-1$$

where γ^i_{int1}, γ^i_{int2} are the two innermost values in the list

$$\Big\{ \sum_{B\subseteq A, B\not\in \mathcal{C}} m_b(B), A \supseteq A_i, A \not\supset A_{i+1} \Big\}.$$

As the innermost values of the above list cannot identified analytically but in the simplest special cases, we need to conclude that the L_1 norm is not suitable for consonant approximation in the belief space.

Partial solution in special cases. Nevertheless, the partial L_1 approximations can be computed in some of the most simple cases. For instance, the *binary case* in the L_1 problem is a bit of an exception, as the norm reads as

$$\sum_{A\supseteq\{x\}, A\not\supset\Theta} \Big| \gamma(A) + \sum_{B\subseteq A, B\not\in\mathcal{C}} m_b(B) \Big| = |\gamma(x)|.$$

In this case there is a single innermost node, and the minimal is attained for $\gamma(x) = 0$, i.e., $m_{co}(x) = m_b(x)$ (see Figure 1). There is a single L_1 approximation, while there is in general a whole polytope of such L_1 consonant approximations.

In the *ternary case* a computation of the set of L_1 approximations in closed form is also possible. The norm of the difference vector is, for the following chain of focal elements $\mathcal{C} = \{\{x\}, \{x, y\}, \Theta\}$

$$\Big|\gamma(x) + \sum_{B\subseteq\{x\}, B\not\in\mathcal{C}} m_b(B)\Big| + \Big|\gamma(x) + \sum_{B\subseteq\{x,z\}, B\not\in\mathcal{C}} m_b(B)\Big| + \Big|\gamma(x,y) + \sum_{B\subseteq\{x,y\}, B\not\in\mathcal{C}} m_b(B)\Big|$$

which reduces to $|\gamma(x)| + |\gamma(x) + (m_b(z) + m_b(x,z))| + |\gamma(x,y) + m_b(y)|$, which in turn is minimized for

$$-(m_b(z) + m_b(x,z)) \leq \gamma(x) \leq 0, \quad \gamma(x,y) = -m_b(y).$$

In terms of the basic probability assignment m_{co} of the sought consonant approximation co the solution reads as follows:

$$b(x) \leq m_{co}(x) \leq b(x,z), \quad m_{co}(\{x,y\}) = b(x,y) - m_{co}(x).$$

In other words, the set of L_1 partial consonant approximations with chain of focal elements $\mathcal{C} = \{\{x\},\{x,y\},\Theta\}$ in a ternary frame is a segment $Cl(\vec{co}_1, \vec{co}_2)$, with extrema:

$$\begin{array}{lll} \vec{m}_{co_1} & = & [b(x), b(x,y) - b(x), 1 - b(x,y)]', \\ \vec{m}_{co_2} & = & [b(x,z), b(x,y) - b(x,z), 1 - b(x,y)]'. \end{array} \tag{46}$$

Comments. We can observe a couple of interesting facts in the L_1 partial solution in the ternary case (46).

1) even in this special case the set of L_1 partial approximations is not entirely admissible, as $b(x,y) - b(x)$ and $b(x,y) - b(x,z)$ can be negative;
2) some sort of structure seems to emerge, as the components of their extremal points have a form $b(A_i) - b(A_{i-1})$ for $i = 1,...,n$.

This recalls the properties of L_∞ approximations in $\mathcal{B}$ (Section V-B) and their possible interpretation in terms of the derivative of a belief function on an ordered chain of focal elements. This deserves further investigation in the near future.

D. L_2 approximation

Partial approximation in each maximal simplex. To find the partial consonant approximation at minimal L_2 distance we need to impose the orthogonality of the difference vector $\vec{b} - \vec{co}$ with respect to any given simplicial component $\mathcal{CO}^{\mathcal{C}}$ of the complex $\mathcal{CO}$:

$$\langle \vec{b} - \vec{co}, \vec{b}_{A_j} - \vec{b}_\Theta \rangle = \langle \vec{b} - \vec{co}, \vec{b}_{A_j} \rangle = 0 \quad \forall A_j \in \mathcal{C}, j = 1,...,n-1 \tag{47}$$

as $\vec{b}_\Theta = \vec{0}$ is the origin of the Cartesian space in $\mathcal{B}$, and $\vec{b}_{A_j} - \vec{b}_\Theta$ for $j = 1,...,n-1$ are the generators of the component $\mathcal{CO}^{\mathcal{C}}$. Using once again the expression (39), the orthogonality conditions (47) generate the following linear system:

$$\left\{ \sum_{A \notin \mathcal{C}} m_b(A) \langle \vec{b}_A, \vec{b}_{A_j} \rangle + \sum_{A \in \mathcal{C}, A \neq \Theta} \beta(A) \langle \vec{b}_A, \vec{b}_{A_j} \rangle = 0 \quad j = 1,...,n-1 \right. \tag{48}$$

where again $\beta(A) = m_b(A) - m_{co}(A)$. This is a linear system in $n-1$ unknowns $\{\beta(A_i) = m_b(A_i) - m_{co}(A_i), i = 1, ..., n-1\}$ and $n-1$ equations. If the matrix determining the system is non-singular, then the latter has a unique solution.

Partial solution in the ternary case. When $\Theta = \{x, y, z\}$ (48) reads as

$$\begin{cases} 3\beta(x) + \beta(x,y) & = & -(m_b(y) + m_b(x) + m_b(x,z)) \\ \beta(x) + \beta(x,y) & = & -m_b(y). \end{cases}$$

This is a linear system whose solution is

$$\beta(x) = -\frac{m_b(z) + m_b(x,z)}{2}, \quad \beta(x,y) = -m_b(y) + \frac{m_b(z) + m_b(x,z)}{2}$$

or, in the b.b.a. m_{co} of the sought partial approximation:

$$\begin{array}{lll} m_{co}(x) & = & m_b(x) + \frac{m_b(z)+m_b(x,z)}{2}, \\ m_{co}(x,y) & = & m_b(y) + m_b(x,y) - \frac{m_b(z)+m_b(x,z)}{2}. \end{array} \tag{49}$$

By normalization $m_{co}(\Theta) = 1 - m_{co}(x) - m_{co}(x,y) = 1 - b(x,y)$.

By Equation (44) the barycenter of the set of partial L_∞ approximations is, as in this case $A_1 = \{x\}$, $A_2 = \{x,y\}$, $A_3 = \{x,y,z\}$:

$$\begin{array}{lll} m(x) & = & \frac{b(A_1) + b(\{x_2\}^c)}{2} = \frac{b(x) + b(\{y\}^c)}{2} = \frac{b(x) + b(x,z)}{2} \\ & = & m_b(x) + \frac{m_b(z) + m_b(x,z)}{2} \\ m(x,y) & = & \frac{b(A_2) + b(\{x_3\}^c)}{2} - \frac{b(A_1) + b(\{x_2\}^c)}{2} = \frac{b(x,y) + b(x,y)}{2} + \\ & & -\frac{b(x) + b(\{y\}^c)}{2} = m_b(y) + m_b(x,y) - \frac{m_b(z) + m_b(x,z)}{2} \\ m(\Theta) & = & 1 - b(A_2) = 1 - b(x,y) \end{array}$$

which coincides with the L_2 partial approximation.

Partial solution in the general case. The solution of system (48) involves rather complicated combinatorial calculations. Nevertheless, the form of the general solution is rather simple.

Theorem 8: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, the unique partial L_2 consonant approximations $co^{\mathcal{C}}_{L_2}[b]$ of b (in the belief space) onto the simplicial component $\mathcal{CO}_{\mathcal{C}}$ is associated with the following basic probability assignment:

$$\begin{array}{l} m_{co}(A_1) = m_b(A_1) + \sum_{A \not\in \mathcal{C}, A \not\supset x_1, x_2} m_b(A) 2^{-|A \setminus A_2|} + \sum_{A \not\in \mathcal{C}, A \supset x_1, A \not\supset x_2} m_b(A) 2^{-|A \setminus A_1|} \\ m_{co}(A_i) = m_b(A_i) + \sum_{A \not\in \mathcal{C}, A \supset x_i, A \not\supset x_{i+1}} m_b(A) 2^{-|A \setminus A_i|} + \\ \qquad - \sum_{A \not\in \mathcal{C}, A \not\supset x_i, A \supset x_{i+1}} m_b(A) 2^{-|A \setminus A_{i-1}|} \qquad \forall i = 2, \cdots, n-1 \\ m_{co}(A_n) = m_{co}(\Theta) = 1 - b(A_{n-1}). \end{array} \tag{50}$$

The proof is rather combinatorial, and can be done by substitution.

By comparison with Equation (44) we can infer that the coincidence of the L_2 partial approximation and the barycenter of the set of L_∞ approximations is an artifact of the, otherwise instructive, ternary case.

Global solution. The computation of the global L_2 approximation is not simple. We plan to solve this issue in the near future.

E. Critical discussion

Summary of the results in the belief space. As we did for the L_p consonant approximations in the mass space, it is worth summarizing the main outcomes of our analysis of L_p consonant approximation in $\mathcal{B}$:

1) L_∞ minimization generates an entire convex set of (partial) approximations on each simplicial component;
2) the barycenter of this set has a potentially interesting interpretation in terms of a formally to specify notion of derivative of a belief function on a linearly ordered chain;
3) the global L_∞ approximations fall as expected on the component associated with the maximal plausibility singleton;
4) the L_1 norm does not seem to be suitable for the job, even though in the cases in which the analytical form of the set of L_1 approximations is obtainable in closed form, this seems to be connected to the difference of belief values on the chain;
5) the L_2 partial approximation is unique and distinct from the above barycenter, while the global L_2 approximation is rather elusive.

It is interesting to compare these results with those obtained in the consistent approximation problem, also in the belief space [15]. The partial L_1/L_2 consistent approximations of b focused on a given element x coincide, and have b.p.a. $m_{cs^x_{L_1}}(A) = m_{cs^x_{L_2}}(A) = m_b(A) + m_b(A \setminus \{x\})$. They coincide with Dubois and Prade's "focused consistent transformations" [3]. On the other hand, global approximations do not have a strong meaning in terms of degrees of belief. Much is still there to be understood on the use of geometric norms for consonant and consistent approximations: other results seem to point to more natural results for approximations in the mass space [16].

Ternary example. We can use the formulae for the ternary case (49), (46) to visualize the outcomes of L_p consonant approximation in the belief space when $\Theta = \{x, y, z\}$, and compare them with approximations in the mass on the same example of Section IV-F. The results can be seen in Figure 11.

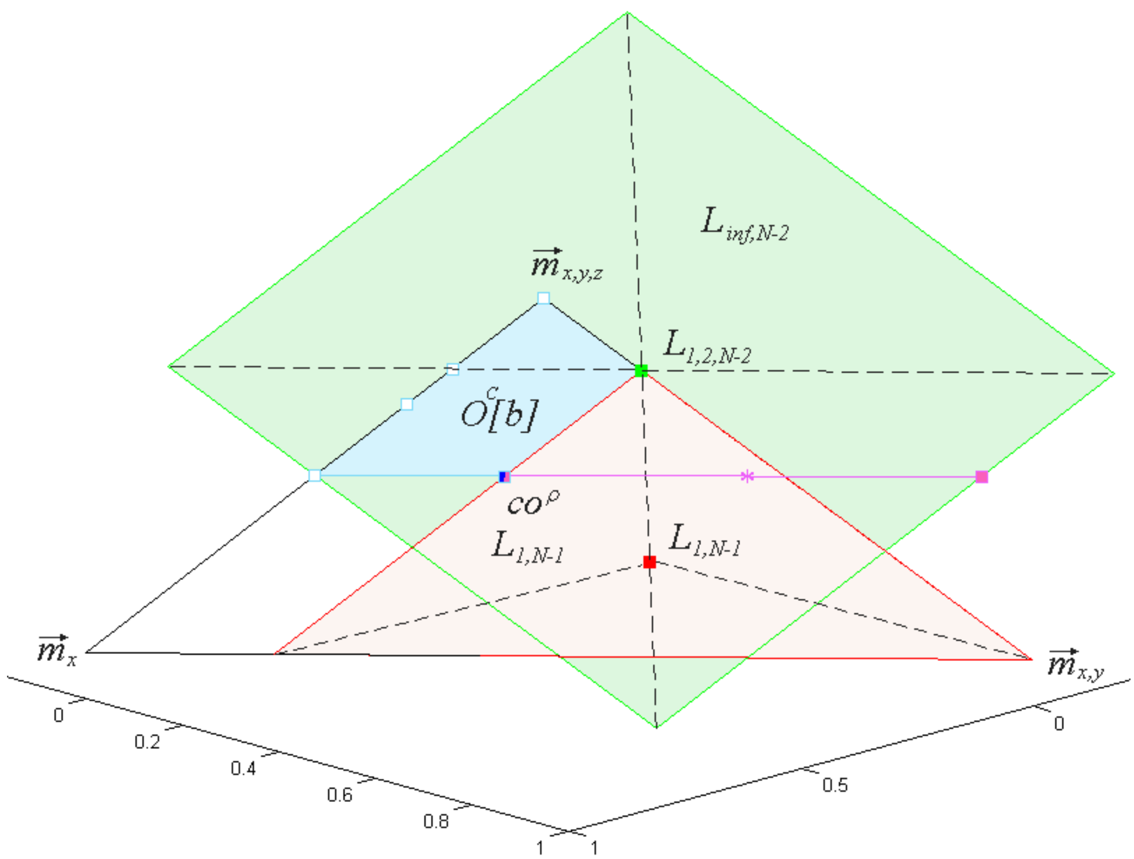

Fig. 11. Comparison between L_p partial consonant approximations in the mass and belief spaces for the belief function with basic probabilities (36) on $\Theta = \{x, y, z\}$.

Once again we can notice several interesting things about approximations in the mass and the belief space. The set of L_1 (partial) approximations in $\mathcal{B}$ (magenta segment), in particular, seems to be very much related to outer consonant approximations, as a continuation of one of their sides. The maximal outer approximation co^ρ is indeed one of the vertices of the segment of L_1 approximations, which is also a subset of the simplex of $L_{\infty,N-2}$ approximations in $\mathcal{M}$. It is reasonable to conjecture that the L_2 approximation in $\mathcal{B}$ (magenta star), while distinct from the barycenter of the L_∞ one in the general case, could be related to the barycenter of the L_1 set. The general form of the set of L_∞ approximations remains elusive.

VI. Conclusions

In this paper we computed all the consonant approximations of a belief function induced by minimizing its L_p distances to the consonant complex, in both the mass space of basic probability vectors and the belief space of belief values. Interpretations for such approximations in the mass space are rather natural in terms of mass redistribution. We compared them with each other and related them with classical outer consonant approximations, with the help of an example. In the belief space results are rather more complex, and seem to be related to the difference of belief values along the desired chain of focal elements.

References

[1] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[2] D. Dubois and H. Prade, *Possibility theory*. New York: Plenum Press, 1988.

[3] ——, "Consonant approximations of belief functions," *International Journal of Approximate Reasoning*, vol. 4, pp. 419–449, 1990.

[4] C. Joslyn and G. Klir, "Minimal information loss possibilistic approximations of random sets," in *Proc. 1992 FUZZ-IEEE Conference*, J. Bezdek, Ed., 1992, pp. 1081–1088.

[5] C. Joslyn, "Possibilistic normalization of inconsistent random intervals," 1997, pp. 44–51.

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

[7] R. R. Yager, "Class of fuzzy measures generated from a Dempster-Shafer belief structure," *International Journal of Intelligent Systems*, vol. 14, pp. 1239–1247, 1999.

[8] C. Roemer and A. Kandel, "Applicability analysis of fuzzy inference by means of generalized Dempster-Shafer theory," *IEEE Transactions on Fuzzy Systems*, vol. 3:4, pp. 448–453, November 1995.

[9] G. J. Klir, W. Zhenyuan, and D. Harmanec, "Constructing fuzzy measures in expert systems," *Fuzzy Sets and Systems*, vol. 92, pp. 251–264, 1997.

[10] S. Heilpern, "Representation and application of fuzzy numbers," *Fuzzy Sets and Systems*, vol. 91, pp. 259–268, 1997.

[11] L. Caro and A. B. Nadjar, “Generalization of the Dempster-Shafer theory: a fuzzy-valued measure,” *IEEE Transactions on Fuzzy Systems*, vol. 7, pp. 255–270, 1999.

[12] P. Smets, “The transferable belief model and possibility theory,” in *Proceedings of NAFIPS-90*, K. Y., Ed., 1990, pp. 215–218.

[13] F. Cuzzolin, “The geometry of consonant belief functions: simplicial complexes of necessity measures,” *Fuzzy Sets and Systems*, vol. 161, no. 10, pp. 1459–1479, 2010.

[14] ——, “Two new Bayesian approximations of belief functions based on convex geometry,” *IEEE Tr. SMC-B*, vol. 37, no. 4, pp. 993–1008, 2007.

[15] ——, “Consistent approximation of belief functions,” in *Proceedings of ISIPTA’09, Durham, UK*, June 2009.

[16] ——, “Lp consistent approximations of belief functions,” *IEEE Transactions on Fuzzy Systems (under review)*, 2008.

[17] ——, “A geometric approach to the theory of evidence,” *IEEE Transactions on Systems, Man and Cybernetics part C*, vol. 38, no. 4, pp. 522–534, 2008.

[18] ——, “Geometric conditioning of belief functions,” in *Proceedings of BELIEF’10, Brest, France*, 2010.

[19] A.-L. Jousselme and P. Maupin, “On some properties of distances in evidence theory,” in *Proceedings of BELIEF’10, Brest, France*, 2010.

[20] C. Shi, Y. Cheng, Q. Pan, and Y. Lu, “A new method to determine evidence distance,” in *Proceedings of the 2010 International Conference on Computational Intelligence and Software Engineering (CiSE)*, 2010, pp. 1–4.

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

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

[23] J. Diaz, M. Rifqi, and B. Bouchon-Meunier, “A similarity measure between basic belief assignments,” in *Proceedings of FUSION’06*, 2006.

[24] P. Smets and R. Kennes, “The transferable belief model,” *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[25] F. Cuzzolin and R. Frezza, “Geometric analysis of belief space and conditional subspaces,” in *Proc. of ISIPTA’01, Ithaca, NY*, 26-29 June 2001.

[26] B. Dubrovin, S. Novikov, and A. Fomenko, *Sovremennaja geometrija. Metody i prilozenija*. Moscow: Nauka, 1986.

[27] P. Smets, “Belief functions versus probability functions,” in *Uncertainty and Intelligent Systems*, S. L. Bouchon B. and Y. R., Eds. Springer Verlag, Berlin, 1988, pp. 17–24.

[28] B. Cobb and P. Shenoy, “On transforming belief function models to probability models,” University of Kansas, School of Business, Working Paper No. 293, Tech. Rep., February 2003.

[29] H. Kyburg, “Bayesian and non-Bayesian evidential updating,” *Artificial Intelligence*, vol. 31:3, pp. 271–294, 1987.

[30] D. Dubois, H. Prade, and P. Smets, “New semantics for quantitative possibility theory.” in *ISIPTA*, 2001, pp. 152–161.

[31] F. Cuzzolin, “The geometry of relative plausibility and belief of singletons,” *submitted to the International Journal of Approximate Reasoning*, 2006.

[32] ——, “Complexes of outer consonant approximations,” in *Proceedings of ECSQARU’09, Verona, Italy*, 2009.