L_p consonant approximations of belief functions

Fabio Cuzzolin
fabio.cuzzolin@brookes.ac.uk

Oxford Brookes University, Oxford, UK

Abstract. In this paper we pose the problem of approximating an arbitrary belief function with a consonant one in a geometric framework. Given a belief function b, the consonant b.f. which minimizes an appropriate distance function from b can be sought. We consider here the classical L_1, L_2 and L_p norms. As consonant belief functions live in a collection of simplices in the belief space, partial approximations on each individual simplex have to be computed in order to find the overall approximation. Interpretations of the obtained approximations in terms of degrees of belief are proposed.

1 Introduction

The theory of evidence (ToE) [1] is a popular approach to uncertainty description. Probabilities are there replaced by *belief functions* (b.f.s), 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 $Nec(A) = 1 - Pos(A)$. Necessity measures have as counterparts in the theory of evidence *consonant* b.f.s, i.e., belief functions whose focal elements are nested [1]. The problem of approximating a belief function with a necessity measure is then equivalent to approximating a belief function with a consonant b.f. [3–6]. 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].

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 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.

In 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 indeed 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. Many of these measures could be in principle employed to define conditional belief functions, or approximate belief functions by necessity or probability measures.

Paper outline. In this paper we derive the expressions of all the consonant approximations of belief functions induced by minimizing L_p distances in the belief space. After providing the necessary background on consonant b.f.s and the approximation problem (Section 2), we compute the approximations induced by L_∞ (3.1), L_1 (3.2) and L_2 (3.3) 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 (Section 4).

2 Geometric consonant approximation

Consonant belief functions. We briefly recall some basic definitions. 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.

A geometric approach. 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 $\boldsymbol{b} = \sum_{\emptyset \subsetneq A \subsetneq \Theta} b(A) X_A$. If we denote by $\boldsymbol{b}_A \doteq b \in \mathcal{B}$ s.t. $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0$ $\forall B \subseteq \Theta$, $B \neq A$ the *categorical* [20] belief function assigning all the mass to a single subset $A \subseteq \Theta$, we can prove that [21, 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 categorical belief functions b_A: $\mathcal{B} = Cl(\boldsymbol{b}_A, \emptyset \subsetneq A \subseteq \Theta)$, where Cl denotes the convex closure operator: $Cl(\boldsymbol{b}_1, ..., \boldsymbol{b}_k) = \left\{ \boldsymbol{b} \in \mathcal{B} : \boldsymbol{b} = \alpha_1 \boldsymbol{b}_1 + \cdots + \alpha_k \boldsymbol{b}_k, \sum_i \alpha_i = 1,\ \alpha_i \geq 0\ \forall i \right\}$.

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

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

Its b.p.a. m_b is nothing but the set of simplicial coordinates of $\boldsymbol{b}$ 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[1]: $\mathcal{P} = Cl(\boldsymbol{b}_x, x \in \Theta)$.

The consonant complex. In this framework the geometry of consonant belief functions can be described by resorting to the notion of *simplicial complex* [22]. 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 $\mathcal{CO}$ of consonant belief functions in the belief space is a simplicial complex. More precisely, $\mathcal{CO}$ is the union of a collection of (maximal) simplices, each of them associated with a maximal chain $\mathcal{C} = \{A_1 \subset \cdots \subset A_n\}$, $|A_i| = i$ of subsets of Θ: $\mathcal{CO}_\mathcal{B} = \bigcup_{\mathcal{C} = A_1 \subseteq \cdots \subseteq A_n} Cl(\boldsymbol{b}_{A_1}, \cdots, \boldsymbol{b}_{A_n})$.

Binary example. In the case of a frame of discernment containing only two elements, $\Theta_2 = \{x, y\}$, each b.f. $b : 2^{\Theta_2} \rightarrow [0,1]$ is completely determined by its belief values $b(x) = m_b(x)$, $b(y) = m_b(y)$, as $b(\Theta) = 1$ and $b(\emptyset) = 0$ $\forall b$. We can then collect them in a vector of $\mathbb{R}^{N-2} = \mathbb{R}^2$ (since $N = 2^2 = 4$): $\boldsymbol{b} = [b(x) = m_b(x), b(y) = m_b(y)]' \in \mathbb{R}^2$. 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 $\mathcal{B}_2$ of all the possible belief functions on Θ_2 can be depicted as the triangle in the Cartesian plane of Figure 1, whose vertices are the points $\boldsymbol{b}_\Theta = [0,0]'$, $\boldsymbol{b}_x = [1,0]'$, $\boldsymbol{b}_y = [0,1]'$, 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 $Cl(\boldsymbol{b}_x, \boldsymbol{b}_y)$. On $\Theta_2 = \{x, y\}$ consonant belief

[1] 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 $m_b(\{x\}), pl_b(\{x\})$.

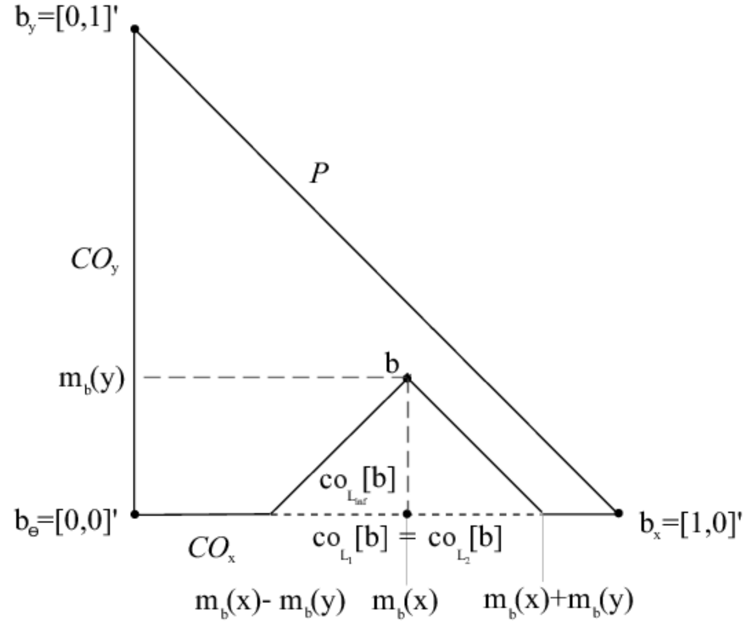

Fig. 1. The belief space $\mathcal{B}_2$ for a binary frame is a triangle in $\mathbb{R}^2$ whose vertices are the categorical belief functions focused on $\{x\}, \{y\}$ and Θ (b_x, b_y, b_Θ) respectively. Consonant belief functions are constrained to belong to the union of the two segments $\mathcal{CO}_x = Cl(\boldsymbol{b}_\Theta, \boldsymbol{b}_x)$ and $\mathcal{CO}_y = Cl(\boldsymbol{b}_\Theta, \boldsymbol{b}_y)$. The unique $L_1 = L_2$ consonant approximation and the set of L_∞ consonant approximations (dashed) on $\mathcal{CO}_x$ are also shown.

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): $\mathcal{CO}_2 = \mathcal{CO}_x \cup \mathcal{CO}_y = Cl(\boldsymbol{b}_\Theta, \boldsymbol{b}_x) \cup Cl(\boldsymbol{b}_\Theta, \boldsymbol{b}_y)$.

The consonant approximation problem. Given a belief function b, we call (metric) *consonant approximation of a belief function b induced by a distance function d* in $\mathcal{B}$ the b.f.(s) $b_d(.|A)$ which minimize(s) the distance $d(b, \mathcal{CO})$ between b and the consonant simplicial complex. We consider as distance functions the three major L_p norms: $d = L_1$; $d = L_2$; $d = L_\infty$. For vectors $\boldsymbol{b}$, $\boldsymbol{b'} \in \mathcal{B}$ representing two belief functions b, b', such norms read as:

$$\|\boldsymbol{b} - \boldsymbol{b'}\|_1 \doteq \sum_{\emptyset \subsetneq B \subseteq \Theta} |b(B) - b'(B)|; \quad \|\boldsymbol{b} - \boldsymbol{b'}\|_2 \doteq \sqrt{\sum_{\emptyset \subsetneq B \subseteq \Theta} (b(B) - b'(B))^2};$$
$$\|\boldsymbol{b} - \boldsymbol{b'}\|_\infty \doteq \max_{\emptyset \subsetneq B \subseteq \Theta} |b(B) - b'(B)|. \tag{2}$$

As the consonant complex $\mathcal{CO}$ is a *collection* of linear spaces (better, simplices which generate a linear space), solving the consonant approximation problem involves finding a number of partial solutions

$$co^{\mathcal{C}}_{L_p}[b] = \arg \min_{\boldsymbol{co} \in \mathcal{CO}_{\mathcal{C}}} \|\boldsymbol{b} - \boldsymbol{co}\|_p \tag{3}$$

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

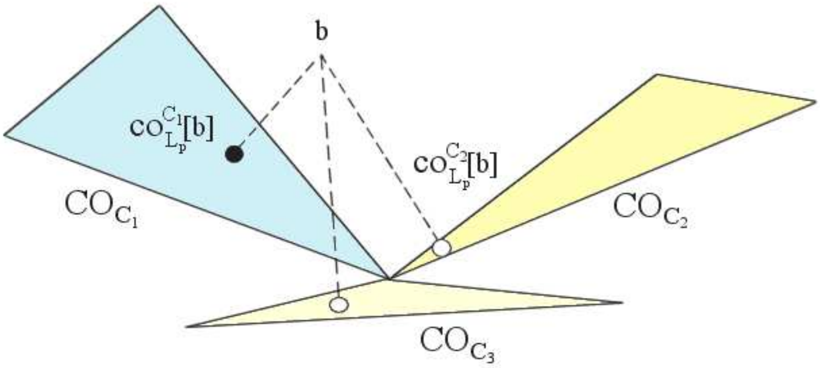

Fig. 2. To minimize the distance of a point from a simplicial complex, we need to find all the partial solutions (3) on all the maximal simplices in the complex (empty circles), to later compare these partial solutions to select a global optimum (black circle).

tance of b from all such partial solutions has to be assessed in order to select a global optimal approximation. Figure 1 shows the obtained (partial) L_p consonant approximations onto $\mathcal{CO}_x$ in the binary case. In such a toy example, $co_{L_1}[b] = co_{L_1}[b]$ coincide and are unique, lying on the barycenter of the set $co_{L_\infty}[b]$ of L_∞ approximations, which instead form a whole interval. Some of these 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. In the rest of the paper we will explicitly compute the L_1, L_2, and L_∞ consonant approximations in the belief space and discuss the results.

3 Consonant approximation in the belief space

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

$$
\begin{aligned}
\boldsymbol{b} - \boldsymbol{b}_{co} = & \sum_{A \not\supset A_1} b(A) X_A + \sum_{A \supseteq A_1} X_A \Big[b(A) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] = \sum_{A \not\supset A_1} b(A) X_A + \\
& + \sum_{A \supseteq A_1} X_A \Big[\sum_{\emptyset \subsetneq B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] = \sum_{A \not\supset A_1} b(A) X_A + \\
& + \sum_{A \supseteq A_1} X_A \Big[\sum_{B \subseteq A, B \in \mathcal{C}} (m_b(B) - m_{co}(B)) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big] \\
= & \sum_{A \not\supset A_1} b(A) X_A + \sum_{A \supseteq A_1} X_A \Big[\gamma(A) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big]
\end{aligned} \tag{4}
$$

after introducing the auxiliary variables $\gamma(A) = \sum_{B\subseteq A, B\in\mathcal{C}} \beta(B)$, $\beta(B) = m_b(A) - m_{co}(B)$. We can therefore write

$$\boldsymbol{b} - \boldsymbol{b}_{co} = \sum_{A \not\supset A_1} b(A) X_A + \sum_{i=1}^{n-1} \sum_{A \supset A_i, A \not\supset A_{i+1}} X_A \Big[\gamma(A_i) + \sum_{B\subseteq A, B\notin\mathcal{C}} m_b(B) \Big], \quad (5)$$

as all the terms in (4) 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 $1 - 1 = 0$.

3.1 L_∞ approximation

Partial approximation in each maximal simplex. Given the expression (5) for the difference vector of interest in the belief space, we can compute the explicit form of its L_∞ norm $\|b - co\|_\infty$ as

$$\begin{aligned} &\max\left\{ \max_i \max_{A\supseteq A_i, A\not\supset A_{i+1}} \Big|\gamma(A_i) + \sum_{B\subseteq A, B\notin\mathcal{C}} m_b(B)\Big|, \max_{A\not\supset A_1} \Big|\sum_{B\subseteq A} m_b(B)\Big| \right\} \\ &= \max\left\{ \max_i \max_{A\supseteq A_i, A\not\supset A_{i+1}} \Big|\gamma(A_i) + \sum_{B\subseteq A, B\notin\mathcal{C}} m_b(B)\Big|, b(A_1^c) \right\}, \end{aligned} \quad (6)$$

as $\max_{A\not\supset A_1} \Big|\sum_{B\subseteq A} m_b(B)\Big| = b(A_1^c)$. Now, (6) can be minimized separately for each $i = 1, ..., n-1$. Clearly, the minimum is attained when the variable elements in (6) are not greater than the constant element $b(A_1^c)$:

$$\max_{A\supseteq A_i, A\not\supset A_{i+1}} \Big|\gamma(A_i) + \sum_{B\subseteq A, B\notin\mathcal{C}} m_b(B)\Big| \leq b(A_1^c) \quad (7)$$

This is a function of the form $\max\{|x + x_1|, ..., |x + x_n|\}$ (see Figure 3). In (7)

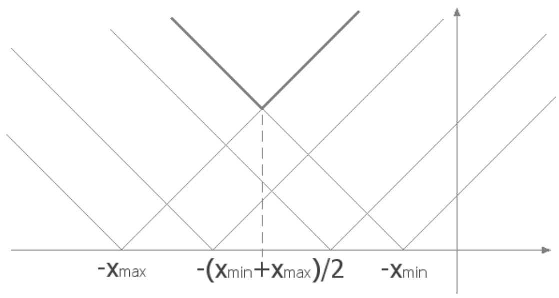

Fig. 3. The minimization of the L_∞ distance from the consonant subspace involves minimizing functions of the form $\max\{|x + x_1|, ..., |x + x_n|\}$ (in bold) depicted above.

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

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

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 3 again). In the case of (7), such minimum and maximum offset values are, respectively,

$$\begin{aligned} \gamma_{min} &= \sum_{B \subseteq A_i, B \notin \mathcal{C}} m_b(B) = b(A_i) - \sum_{j=1}^{i} m_b(A_j) \\ \gamma_{max} &= \sum_{B \subseteq (A_{i+1} \setminus A_i)^c, B \notin \mathcal{C}} m_b(B) = b(\{x_{i+1}\}^c) - \sum_{j=1}^{i} m_b(A_j) \end{aligned}$$

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.,

$$-\frac{\gamma_{min} + \gamma_{max}}{2} - b(A_1^c) \leq \gamma(A_i) \leq -\frac{\gamma_{min} + \gamma_{max}}{2} + b(A_1^c). \tag{8}$$

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

$$-\frac{\gamma_{min} + \gamma_{max}}{2} = -\frac{b(A_i) + b(\{x_{i+1}\}^c)}{2} + \sum_{j=1}^{i} m_b(A_j). \tag{9}$$

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

$$\begin{aligned} m_{co}(A_1) &= \frac{b(A_1) + b(\{x_2\}^c)}{2} \\ m_{co}(A_i) &= \frac{b(A_i) + b(\{x_{i+1}\}^c)}{2} - \frac{b(A_{i-1}) + b(\{x_i\}^c)}{2} \\ &= \frac{b(A_i) + b(A_i + A_{i+1}^c)}{2} - \frac{b(A_{i-1}) + b(A_{i-1} + A_i^c)}{2} \qquad i = 2, ..., n-1 \end{aligned} \tag{10}$$

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 1. *Given a belief function $b : 2^\Theta \to [0, 1]$, 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 (10).*

Interpretation. The partial L_∞ consonant approximation (10) 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 $\frac{b(A_i)+b(A_i+A^c_{i+1})}{2}$ 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 more extensive research report.

Global solution To compute the global L_∞ approximation of the original belief function b, we need to detect 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 7). 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

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

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

Theorem 2. *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. if the maximal plausibility element of Θ:*

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

3.2 L_1 approximation

Recalling the expression (5) of the difference vector $\boldsymbol{b} - \boldsymbol{co}$ in the belief space, its L_1 norm reads as:

$$\|\boldsymbol{b} - \boldsymbol{co}\|_{L_1} = \sum_{i=1}^{n-1} \sum_{A \supseteq A_i, A \not\supset A_{i+1}} \left| \gamma(A_i) + \sum_{B \subseteq A, B \not\in \mathcal{C}} m_b(B) \right| + \sum_{A \not\supset A_1} |b(A)|. \quad (11)$$

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 that minimize such a function lie, let us consider the simple function (see Figure 4). It is easy

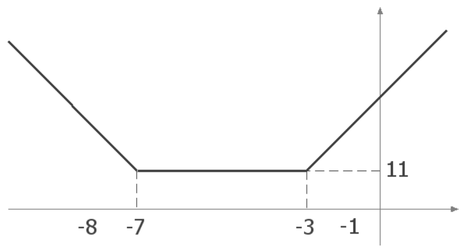

Fig. 4. The minimization of the L_1 distance from the consonant subspace involves minimizing functions such as the one depicted above, $|x+1|+|x+3|+|x+7|+|x+8|$, which is minimized for $3 \leq x \leq 7$.

to see that such a function is minimized by the interval of values comprised between the two innermost nodes. As a consequence:

Theorem 3. *Given a belief function* $b : 2^\Theta \rightarrow [0, 1]$*, 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 following intervals in the auxiliary variables* $\gamma(A_i)$*:*

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

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

3.3 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 $\boldsymbol{b} - \boldsymbol{co}$ with respect to any given simplicial component $\mathcal{CO}^{\mathcal{C}}$ of the complex $\mathcal{CO}$:

$$\langle \boldsymbol{b} - \boldsymbol{co}, \boldsymbol{b}_{A_j} - \boldsymbol{b}_{\Theta} \rangle = \langle \boldsymbol{b} - \boldsymbol{co}, \boldsymbol{b}_{A_j} \rangle = 0 \quad \forall A_j \in \mathcal{C}, j = 1, ..., n-1 \tag{12}$$

as $\boldsymbol{b}_\Theta = \boldsymbol{0}$ is the origin of the Cartesian space in $\mathcal{B}$, and $\boldsymbol{b}_{A_j} - \boldsymbol{b}_\Theta$ for $j = 1, ..., n-1$ are the generators of the component $\mathcal{CO}^{\mathcal{C}}$. Using once again the expression (5), the orthogonality conditions (12) generate the following linear system:

$$\left\{ \sum_{A \not\in \mathcal{C}} m_b(A) \langle \boldsymbol{b}_A, \boldsymbol{b}_{A_j} \rangle + \sum_{A \in \mathcal{C}, A \neq \Theta} \beta(A) \langle \boldsymbol{b}_A, \boldsymbol{b}_{A_j} \rangle = 0 \qquad j = 1, ..., n-1 \right. \tag{13}$$

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. In the ternary case the system reads as

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

This is a linear system whose matrix and its inverse are

$$\mathcal{A} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}, \mathcal{A}^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 3/2 \end{bmatrix}$$

so that its solution is

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

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

$$m_{co}(x) = m_b(x) + \frac{m_b(z) + m_b(x,z)}{2}, \; m_{co}(x,y) = m_b(y) + m_b(x,y) - \frac{m_b(z) + m_b(x,z)}{2}$$

and by normalization

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

By Equation (10) the barycenter of the set of partial L_∞ approximations is, as in this case $A_1 = \{x\}$, $A_2 = \{x,y\}$, $A_3 = \{x,y,z\}$:

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

which coincides with the L_2 partial approximation.

Partial solution in the general case. The solution of system (13) involves rather complicated combinatorial calculations. Nevertheless, the form of the general solution is rather simple.

Theorem 4. *Given a belief function $b : 2^\Theta \to [0,1]$, 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:*

$$
\begin{aligned}
m_{co}(A_1) &= m_b(A_1) + \sum_{A \not\in \mathcal{C}, A \not\supset x_1, x_2} m_b(A) 2^{-|A \setminus A_2|} + \sum_{A \not\in \mathcal{C}, A \supset x_1, A \not\supset x_2} m_b(A) 2^{-|A \setminus A_1|} \\
m_{co}(A_i) &= m_b(A_i) + \sum_{A \not\in \mathcal{C}, A \supset x_i, A \not\supset x_{i+1}} m_b(A) 2^{-|A \setminus A_i|} + \\
&\quad - \sum_{A \not\in \mathcal{C}, A \not\supset x_i, A \supset x_{i+1}} m_b(A) 2^{-|A \setminus A_{i-1}|} \qquad \forall i = 2, \cdots, n-1 \\
m_{co}(A_n) &= m_{co}(\Theta) = 1 - b(A_{n-1}).
\end{aligned} \tag{14}
$$

The proof can be done by substitution [23].
By comparison with Equation (10) 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.

4 Critical discussion and conclusions

In this paper we posed the consonant approximation problem in geometric terms, by computing the approximations obtained by minimizing a-priori sensible L_p distances between the original b.f. and the consonant complex in the belief space $\mathcal{B}$. We observed that: 1. the L_1 norm is not really suitable for the job; 2. L_∞ minimization generates an entire convex set of (partial) approximations on each simplicial component; 3. 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; 4. the global L_∞ approximations fall as expected on the component associated with the maximal plausibility singleton; 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 [23, 16]. This is material for further analysis.

References

1. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
2. Dubois, D., Prade, H.: Possibility theory. Plenum Press, New York (1988)
3. Dubois, D., Prade, H.: Consonant approximations of belief functions. International Journal of Approximate Reasoning **4** (1990) 419–449
4. Joslyn, C., Klir, G.: Minimal information loss possibilistic approximations of random sets (1992)
5. Joslyn, C.: Possibilistic normalization of inconsistent random intervals. Advances in Systems Science and Applications (1997) 44–51
6. Baroni, P.: Extending consonant approximations to capacities. In: IPMU. (2004) 1127–1134
7. Yager, R.R.: Class of fuzzy measures generated from a Dempster-Shafer belief structure. International Journal of Intelligent Systems **14** (1999) 1239–1247
8. Roemer, C., Kandel, A.: Applicability analysis of fuzzy inference by means of generalized Dempster-Shafer theory. IEEE Transactions on Fuzzy Systems **3:4** (November 1995) 448–453
9. Klir, G.J., Zhenyuan, W., Harmanec, D.: Constructing fuzzy measures in expert systems. Fuzzy Sets and Systems **92** (1997) 251–264
10. Heilpern, S.: Representation and application of fuzzy numbers. Fuzzy Sets and Systems **91** (1997) 259–268
11. Caro, L., Nadjar, A.B.: Generalization of the Dempster-Shafer theory: a fuzzy-valued measure. IEEE Transactions on Fuzzy Systems **7** (1999) 255–270
12. Smets, P.: The transferable belief model and possibility theory. In Y., K., ed.: Proceedings of NAFIPS-90. (1990) 215–218
13. Cuzzolin, F.: The geometry of consonant belief functions: simplicial complexes of necessity measures. Fuzzy Sets and Systems **161**(10) (2010) 1459–1479
14. Cuzzolin, F.: Two new Bayesian approximations of belief functions based on convex geometry. IEEE Tr. SMC-B **37**(4) (2007) 993–1008
15. Cuzzolin, F.: Consistent approximation of belief functions. In: Proceedings of ISIPTA'09, Durham, UK. (June 2009)
16. Cuzzolin, F.: l_p consistent approximations of belief functions. IEEE Transactions on Fuzzy Systems (under review) (2008)
17. Cuzzolin, F.: A geometric approach to the theory of evidence. IEEE Transactions on Systems, Man and Cybernetics part C **38**(4) (2008) 522–534
18. Cuzzolin, F.: Geometric conditioning of belief functions. In: Proceedings of BELIEF'10, Brest, France. (2010)
19. In: Proceedings of BELIEF'10, Brest, France. (2010)
20. Smets, P., Kennes, R.: The transferable belief model. Artificial Intelligence **66** (1994) 191–234
21. Cuzzolin, F., Frezza, R.: Geometric analysis of belief space and conditional subspaces. In: Proceedings of the 2^{nd} International Symposium on Imprecise Probabilities and their Applications (ISIPTA2001), Cornell University, Ithaca, NY. (26-29 June 2001)
22. Dubrovin, B., Novikov, S., Fomenko, A.: Sovremennaja geometrija. Metody i prilozenija. Nauka, Moscow (1986)
23. Cuzzolin, F.: l_p consonant approximations of belief functions. IEEE Tr.Fuzzy Systems (to submit) (2011)