L_p consonant approximation of belief functions in the mass space

Fabio Cuzzolin
Oxford Brookes University, Oxford, United Kingdom
Fabio.Cuzzolin@brookes.ac.uk

Abstract

In this paper we pose the problem of approximating an arbitrary belief function (b.f.) with a consonant one, in a geometric framework in which belief functions are represented by the vectors of their basic probabilities, or "mass space". Given such a vector $\vec{m}_b$, the consonant b.f. which minimizes an appropriate distance function from $\vec{m}_b$ 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 mass 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 basic probabilities are proposed, and the results compared with those of previous approaches, in particular outer consonant approximation.

Keywords. Consonant belief functions, (outer) consonant approximation, mass space, Lp norms.

1 Introduction

The theory of evidence (ToE) [22] 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 [10], 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 $Nec(A) = 1 - Pos(A^c)$.
Necessity measures have as counterparts in the theory of evidence *consonant* b.f.s, i.e., belief functions whose focal elements are nested [22]. The problem of approximating a belief function with a necessity measure is then equivalent to approximating a belief function with a consonant b.f. [1, 11, 15, 16]. 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. Several authors, such as Yager [25] and Romer [21] amongst others, have studied the connection between fuzzy numbers and Dempster-Shafer theory. Klir *et al* have published an excellent discussion [20] on the relations among fuzzy and belief measures and possibility theory. Heilpern [13] 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 [2], 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 [24].
Dubois and Prade [11], more specifically, have extensively worked on consonant approximations of belief functions. As belief functions are computationally expensive to work on (at least in a naive way), mapping them to necessity or possibility measures, which only depends on their values on singletons, can greatly reduce the complexity of making inferences or decisions under uncertainty. Dubois and Prade's work has been later considered in [15, 16]. In particular, the notion of "outer consonant approximation" has received considerable attention in the past. Indeed, belief functions admit the following order relation: $b \leq b' \Leftrightarrow b(A) \leq b'(A)\ \forall A \subseteq \Theta$, called "weak inclusion". It is then possible to introduce the notion of "outer consonant approximations" [11] 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 [1] to capacities. In [7] the author has indeed 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 himself introduced the notion of orthogonal projection $\pi[b]$ of a belief function onto the probability simplex [3], and studied consistent approximations of belief functions induced by classical L_p norms [8] in the space of belief functions [4]. In [6] he has shown that norm minimization can also be used to define families of geometric conditional b.f.s. Jousselme et al [17] have recently conducted a nice survey of the similarity measures between belief functions introduced so far. Other similarity measures between belief functions have been proposed by Shi et al [23], Jiang et al [14], and others [9, 14, 19]. Many of these measures could be in principle employed to define conditional belief functions, or approximate b.f.s by necessity 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 mass space (with respect to the counting measure on 2^Θ). After providing the necessary background on consonant b.f.s and the approximation problem (Section 2), we compute the approximations induced by L_1 (3.1), L_2 (3.2) and L_∞ (3.3) norms, respectively. Their interpretation in terms of mass re-assignment and their relation with outer consonant approximations are discussed in Section 4, and illustrated in the significant ternary case.

2 Consonant approximation

Consonant belief functions. We briefly recall here a few basis definitions. A *basic probability assignment* (b.p.a.) over a finite set *(frame of discernment [22])* Θ is a function $m_b : 2^\Theta \rightarrow [0,1]$ on its power set $2^\Theta = \{A \subseteq \Theta\}$ such that $m_b(\emptyset) = 0$ and $\sum_{A\subseteq\Theta} m_b(A) = 1$. 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: $b(A) = \sum_{B\subseteq A} m_b(B)$. The *plausibility function* (pl.f.) $pl_b : 2^\Theta \rightarrow [0,1]$, $A \mapsto pl_b(A)$, where $pl_b(A) \doteq 1-b(A^c) = 1-\sum_{B\subseteq A^c} m_b(B) = \sum_{B\cap A\neq\emptyset} m_b(B)$, expresses the amount of evidence *not against* A. A probability measure 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.

Mass vector representations. 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 $\mathbb{R}^{N-2}$ [4]. In the same way, each belief function is uniquely associated with the related set of mass values $\{m(A), \emptyset \subsetneq A \subseteq \Theta\}$ (Θ this time included). It can therefore be seen also as a point of $\mathbb{R}^{N-1}$, the vector $\vec{m}_b$ of its $N-1$ mass components:

$$\vec{m}_b = \sum_{\emptyset\subsetneq B\subseteq\Theta} m_b(B)\vec{m}_B, \qquad (1)$$

where $\vec{m}_B$ is the vector of mass values associated with the ("categorical") mass function $\vec{m}_B$ assigning all the mass to a single event B: $\vec{m}_B(B) = 1$, $\vec{m}_B(A) = 0 \ \forall A \neq B$. 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 $\vec{m}_b = \sum_{\emptyset\subsetneq B\subsetneq\Theta} m_b(B)\vec{m}_B$, in which this time the component Θ is neglected. We will consider both representations in the following. The collection $\mathcal{M}$ of points which are valid basic probability assignments is a *simplex*[1], which we call *mass space*. $\mathcal{M}$ is the convex closure[2] $\mathcal{M} = Cl(\vec{m}_A, \emptyset \subsetneq A \subseteq \Theta)$.

The consonant complex. In this framework the geometry of consonant belief functions can be described in terms of *simplicial complexes* [12], i.e., collections Σ of simplices of arbitrary dimensions such that: 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. Now, the region $\mathcal{CO}$ of consonant belief functions in the belief space is a simplicial complex [7]. Namely, $\mathcal{CO}$ 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. When the mass of some element of the maximal chain is zero, the simplicial coordinate of the associated b.f. is also zero. Analogously, the region of consonant belief functions in the mass space $\mathcal{M}$ will be the simplicial complex:

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

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 mass values $m_b(x)$, $m_b(y)$, as $m_b(\Theta) =$

[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*.

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

$1 - m_b(x) - m_b(y)$ 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. Since $m_b(x) \geq 0$, $m_b(y) \geq 0$,

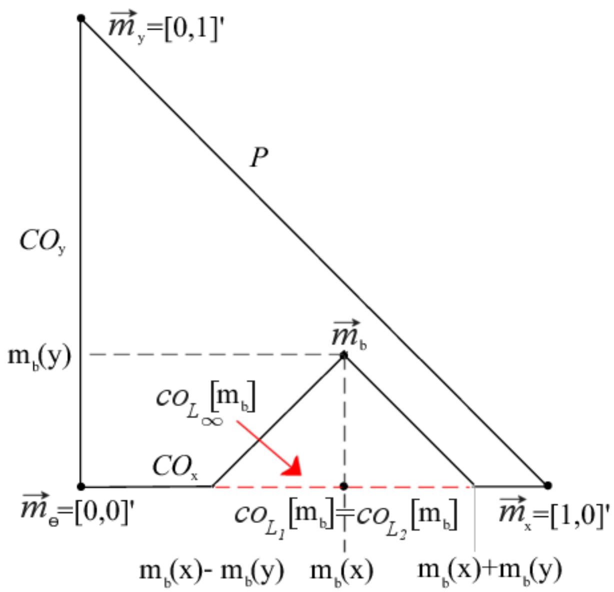

Figure 1: The belief space $\mathcal{M}_2$ for a binary frame is a triangle in $\mathbb{R}^2$ whose vertices are the mass vectors associated with the categorical belief functions focused on $\{x\}, \{y\}$ and Θ: $\vec{m}_x, \vec{m}_y, \vec{m}_\Theta$. Consonant b.f.s live in the union of the two segments $\mathcal{CO}_x = Cl(\vec{m}_\Theta, \vec{m}_x)$ 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.

and $m_b(x) + m_b(y) \leq 1$ we can easily infer that the set $\mathcal{M}_2$ of all the possible basic probability assignments on Θ_2 can be depicted as the triangle in the Cartesian plane of Figure 1, whose vertices are the points $\vec{m}_\Theta = [0,0]'$, $\vec{m}_x = [1,0]'$, $\vec{m}_y = [0,1]'$, 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 $Cl(\vec{m}_x, \vec{m}_y)$.

Consonant approximations in the binary case. 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): $\mathcal{CO}_2 = \mathcal{CO}_x \cup \mathcal{CO}_y = Cl(\vec{m}_\Theta, \vec{m}_x) \cup Cl(\vec{m}_\Theta, \vec{m}_y)$.
Figure 1 illustrates the L_p consonant approximations of a given $\vec{m}_b$ as well. We can notice that the L_1 and L_2 (partial) approximations coincide, and are located in the barycenter of the set of L_∞ approximations, which form instead a whole interval. Such L_1/L_2 approximations leave the mass of $\{x\}$ unchanged, and re-assign the mass of $\{y\}$ (which is not in the chain $\{\{x\}, \{x,y\}\}$) to Θ. Such features are retained in the general case (Section 4).

The consonant approximation problem. Given a belief function b with basic probability assignment m_b, we call (metric) *consonant approximation of a belief function b induced by a distance function d* in $\mathcal{M}$ the b.f.(s) $co_d[m_b]$ which minimize(s) the distance $d(\vec{m}_b, \mathcal{CO})$ between the mass vector $\vec{m}_b$ representing m_b and the consonant simplicial complex

$$co_d[m_b] = \arg \min_{\vec{m}_{co} \in \mathcal{CO}} d(\vec{m}_b, \vec{m}_{co}), \tag{2}$$

under the condition that such minima exist.

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 problem (2) one of the classical norms

$$\begin{array}{rcl} \|\vec{m}_b - \vec{m}_{b'}\|_{L_1} & = & \displaystyle\sum_{A \subseteq \Theta} |m_b(A) - m_{b'}(A)|, \\ \|\vec{m}_b - \vec{m}_{b'}\|_{L_2} & = & \sqrt{\sum_{A \subseteq \Theta} (m_b(A) - m_{b'}(A))^2}, \\ \|\vec{m}_b - \vec{m}_{b'}\|_{L_\infty} & = & \displaystyle\max_{A \subseteq \Theta} \Big\{ |m_b(A) - m_{b'}(A)| \Big\} \end{array} \tag{3}$$

would be meaningful. In the probabilistic case, in the belief space $\mathcal{B}$ ($p[b] = \arg\min_{p \in \mathcal{P}} dist(b, p)$), 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 $\mathcal{P}$ [3]. On the other, the set of L_1/L_∞ probabilistic approximations of b (in the belief space) coincides with the set of probabilities dominating b: $\{p : p(A) \geq b(A)\}$ (at least in the binary case).

Other norms. The L_p family of norms is important and useful also in classical probability theory. Clearly, however, a number of other norms can be introduced in the framework of belief functions and used to define consonant (or Bayesian) approximations. For instance, generalizations to belief functions of the classical Kullback-Leibler divergence of two probability distributions P, Q ($D_{KL}(P|Q) = \int_{-\infty}^{\infty} p(x) \log(\frac{p(x)}{q(x)}) dx$) or other measures based on information theory such as fidelity and entropy-based norms [18] can be studied. Many other similarity measures have indeed been proposed [9, 14, 19, 23]. The application of similarity measures more specific to belief functions or inspired by classical probability to the approximation problem

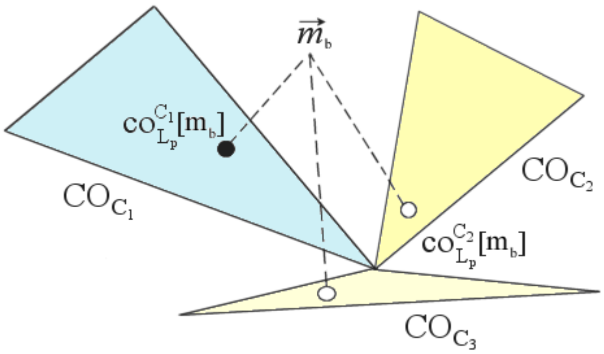

Figure 2: To minimize the distance of a point from a simplicial complex, we need to find all the partial solutions (4) on all the maximal simplices of the complex (empty circles), to later compare these partial solutions to select a global optimum (black circle).

is an enormous task, of which this paper can be seen as just a first step.

Distance of a point from a simplicial complex. As the consonant complex $\mathcal{CO}$ is a *collection* of simplices, solving the consonant approximation problem involves finding a number of partial solutions

$$co^{\mathcal{C}}_{L_p}[m_b] = \arg \min_{\vec{co} \in \mathcal{CO}_{\mathcal{C}}} \|\vec{m}_b - \vec{co}\|_{L_p} \qquad (4)$$

(see Figure 2), one for each maximal chain $\mathcal{C}$ of subsets of Θ. Then, the distance of $\vec{m}_b$ 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, $co_{L_1}[m_b] = co_{L_2}[m_b]$ coincide and are unique, lying on the barycenter of the set $co_{L_\infty}[m_b]$ 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 [8] approximations coincide, and there is no difference between belief and mass space [6] representation. In the rest of the paper we will explicitly compute the L_1, L_2, and L_∞ consonant approximations in the mass space and discuss the results.

3 Consonant approximation in $\mathcal{M}$

If we choose the $N-1$-dimensional version of the mass space (see Equation (1)), 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 the difference vector is

$$\vec{m}_b - \vec{m}_{co} = \sum_{A \in \mathcal{C}} (m_b(A) - m_{co}(A))\vec{m}_A + \sum_{A \not\in \mathcal{C}} m_b(A)\vec{m}_A. \qquad (5)$$

If we instead pick the $N-2$-dimensional version of the mass space, 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

$$\sum_{A \in \mathcal{C}, A \neq \Theta} (m_b(A) - m_{co}(A))\vec{m}_A + \sum_{A \not\in \mathcal{C}} m_b(A)\vec{m}_A. \qquad (6)$$

3.1 L_1 approximation

3.1.1 $\mathbb{R}^{N-1}$ representation

Consider first the $\mathbb{R}^{N-1}$ representation of mass vectors. Given the difference vector (5) its L_1 norm is $\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \sum_{A \in \mathcal{C}} |m_b(A) - m_{co}(A)| + \sum_{A \not\in \mathcal{C}} m_b(A) = \sum_{A \in \mathcal{C}} |\beta(A)| + \sum_{A \not\in \mathcal{C}} m_b(A)$, where $\beta(A) \doteq m_b(A) - m_{co}(A)$ and

$$\sum_{A \in \mathcal{C}} \beta(A) = \sum_{A \in \mathcal{C}} (m_b(A) - m_{co}(A)) = \sum_{A \in \mathcal{C}} m_b(A) - 1 \qquad (7)$$

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

The above norm reads therefore as, as a function of the variables $\{\beta(A), A \in \mathcal{C}, A \neq \Theta\}$,

$$\begin{aligned} \|\vec{m}_b - \vec{m}_{co}\|_{L_1} = & \Big| \sum_{A \in \mathcal{C}} m_b(A) - 1 - \sum_{A \in \mathcal{C}, A \neq \Theta} \beta(A) \Big| \\ & + \sum_{A \in \mathcal{C}, A \neq \Theta} |\beta(A)| + \sum_{A \not\in \mathcal{C}} m_b(A). \end{aligned} \qquad (8)$$

Partial approximation. This function has the form

$$\sum_i |x_i| + \Big| - \sum_i x_i - k \Big|, \quad k \geq 0 \qquad (9)$$

which has an entire simplex of minima, namely: $x_i \leq 0 \ \forall i$, $\sum_i x_i \geq -k$ (see [6] for a similar optimization problem in the geometric conditioning context). The minima of the L_1 norm (8) are therefore given by the following system of constraints:

$$\begin{cases} \beta(A) \leq 0 & \forall A \in \mathcal{C}, A \neq \Theta, \\ \sum_{A \in \mathcal{C}, A \neq \Theta} \beta(A) \geq \sum_{A \in \mathcal{C}} m_b(A) - 1. & \end{cases} \qquad (10)$$

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

$$\begin{cases} m_{co}(A) \geq m_b(A) & \forall A \in \mathcal{C}, A \neq \Theta, \\ \sum_{A \in \mathcal{C}, A \neq \Theta} (m_b(A) - m_{co}(A)) \geq \sum_{A \in \mathcal{C}} m_b(A) - 1 & \end{cases} \qquad (11)$$

where the last constraint reduces to

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

i.e., $m_{co}(\Theta) \geq m_b(\Theta)$. Therefore the solution is $m_{co}(A) \geq m_b(A)\ \forall A \in \mathcal{C}$.

Vertices and barycenter of the partial approximation. The vertices of the set of approximations which are the solutions of (10) are given by the vectors $\{\vec{\beta}_A, A \in \mathcal{C}\}$ 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 set of partial L_1 approximations are the vectors $\{\vec{m}_A^{L_1}, A \in \mathcal{C}\}$ such that

$$\vec{m}_A^{L_1}(B) = \begin{cases} m_b(B) + \sum_{A \notin \mathcal{C}} m_b(A) & B = A, \\ m_b(B) & B \neq A \end{cases} \quad (12)$$

whose barycenter is $co_{\overline{L_1}, N-1}[m_b](B) = m_b(B) + \frac{\sum_{A \notin \mathcal{C}} m_b(A)}{n}$.

Global approximation. To find the *global* L_1 approximation on the consonant complex, we need to find out which component is associated with the minimal L_1 distance. The partial approximations (11) onto $\mathcal{CO}^{\mathcal{C}}$ have L_1 distance from $\vec{m}_b$ given by:

$$\sum_{A \notin \mathcal{C}} m_b(A) = 1 - \sum_{A \in \mathcal{C}} m_b(A). \quad (13)$$

Therefore, the component of the consonant complex at minimal distance is that one associated with the chain that has maximal mass in the original b.f.

3.1.2 $\mathbb{R}^{N-2}$ representation

In the $\mathbb{R}^{N-2}$ representation of mass vectors, the L_1 norm of the difference vector (6) is

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \sum_{A \in \mathcal{C}, A \neq \Theta} |m_b(A) - m_{co}(A)| + \sum_{A \notin \mathcal{C}} m_b(A)$$

which is obviously minimized by

$$m_{co}(A) = m_b(A) \quad \forall A \in \mathcal{C}, A \neq \Theta. \quad (14)$$

Again, to find the global L_1 approximation on the consonant complex in $\mathbb{R}^{N-2}$, we need to find the closest simplicial component. As the partial approximation (14) onto $\mathcal{CO}^{\mathcal{C}}$ has L_1 distance from $\vec{m}_b$ given as before by (13), we have the following.

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

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

with vertices given by Equation (12), associated with the maximal chain of focal elements which maximizes the total original mass of the chain

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

Its global L_1 consonant approximations in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$ is the (unique) partial approximation $co^{\mathcal{C}^*}_{L_1, \mathcal{M}, N-2}[m_b]$ 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 of focal elements.

Not only the two approximations are consistent in the sense that they have the same chain of focal elements, but 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 $\mathbb{R}^{N-2}$.

3.2 L_2 approximation

In order to find the L_2 consonant approximation(s) 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 $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$. Hence, 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. In the two cases $\mathbb{R}^{N-1}$ and $\mathbb{R}^{N-2}$ we will therefore have two different difference vectors and two different orthogonality conditions. In the both cases we need to write:

$$\langle \vec{m}_b - \vec{m}_{co}, \vec{m}_A - \vec{m}_\Theta \rangle = 0 \quad \forall A \in \mathcal{C}, A \neq \Theta. \quad (15)$$

3.2.1 $\mathbb{R}^{N-1}$ representation

In the $N-1$ dimensional mass space, however, the vector $\vec{m}_A - \vec{m}_\Theta$ is such that $\vec{m}_A - \vec{m}_\Theta(B) = 1$ if $B = A$, $\vec{m}_A - \vec{m}_\Theta(B) = -1$ if $B = \Theta$, 0 otherwise. Hence, the orthogonality condition becomes

$$\beta(A) - \beta(\Theta) = 0 \quad \forall A \in \mathcal{C}, A \neq \Theta.$$

Partial approximation. By Equation (7) $\beta(\Theta) = \sum_{A \in \mathcal{C}} m_b(A) - 1 - \sum_{A \in \mathcal{C}, A \neq \Theta} \beta(A)$ 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 $\mathcal{C}$, $A \neq \Theta$. By substitution it can be proven that the solution is $\beta(A) = \frac{\sum_{B\in\mathcal{C}} m_b(B)-1}{n}$. The mass of the partial L_2 consonant approximation is therefore, $\forall A \in \mathcal{C}$:

$$m_{co}(A) = m_b(A) + \frac{1 - \sum_{B\in\mathcal{C}} m_b(B)}{n}. \qquad (16)$$

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:

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

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

3.2.2 $\mathbb{R}^{N-2}$ representation

In 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 approximation is given by

$$\begin{cases} m_{co}(A) = m_b(A) & A \in \mathcal{C}, A \neq \Theta \\ m_{co}(\Theta) = m_b(\Theta) + \sum_{B\notin\mathcal{C}} m_b(B). \end{cases} \qquad (17)$$

The optimal distance is, in this case, $\|\vec{m}_b - \vec{m}_{co}\|^2_{L_2} = \sum_{A\subseteq\Theta}(m_b(A) - m_{co}(A))^2 = \sum_{A\notin\mathcal{C}}(m_b(A))^2 + (\sum_{A\notin\mathcal{C}} m_b(A))^2$, 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 2. *Given a belief function $b : 2^\Theta \to [0, 1]$ with b.p.a. m_b, the global L_2 consonant approximations of b in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-1}$ is the set 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 of focal elements which minimizes the sum of square masses outside the chain: $\mathcal{C}^* = \arg\min_{\mathcal{C}} \sum_{A\notin\mathcal{C}}(m_b(A))^2$.
Its global L_2 consonant approximations in the mass space $\mathcal{M}$ of dimension $\mathbb{R}^{N-2}$ is the (unique) partial approximation $co^{\mathcal{C}^*}_{L_1,\mathcal{M},N-2}[m_b] =$

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

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 (17) and (14) all partial L_1 and L_2 consonant approximations coincide in the mass space of dimension $N - 2$.

3.3 L_∞ approximation

3.3.1 $\mathbb{R}^{N-1}$ representation

In the $N - 1$ representation, the L_∞ norm of the difference vector is

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} = \max\Big\{ \max_{A\in\mathcal{C}} |\beta(A)|, \max_{B\notin\mathcal{C}} m_b(B) \Big\},$$

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

$$|\beta(\Theta)| = \Big| \sum_{B\notin\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B) \Big|$$

and the norm to minimize becomes

$$\begin{aligned}\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} = \max\Big\{ &\max_{A\in\mathcal{C},A\neq\Theta} |\beta(A)|, \\ &\Big|\sum_{B\notin\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B)\Big|, \max_{B\notin\mathcal{C}} m_b(B)\Big\}.\end{aligned} \qquad (18)$$

This is a function of the form

$$\max\Big\{|x_1|, |x_2|, |x_1 + x_2 + k_1|, k_2\Big\} \qquad (19)$$

with $0 \leq k_2 \leq k_1 \leq 1$. If $|\mathcal{C}| = 2$, for instance, $x_1 = \beta(A_1)$, $x_2 = \beta(A_2)$, $k_1 = \sum_{B\notin\mathcal{C}} m_b(B)$ and $k_2 = \max_{B\notin\mathcal{C}} m_b(B)$. 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 3. 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:

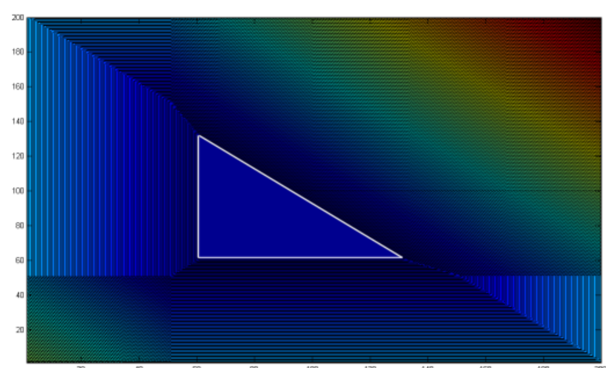

Figure 3: Contour function (level sets) and minimal points (white triangle) of a function of the form (19), when $k_1 \leq 3k_2$. In the example $k_2 = 0.4$ and $k_1 = 0.5$.

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 (18), in the first case

$$\max_{B \not\in \mathcal{C}} m_b(B) \geq \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B) \qquad (20)$$

the set of partial L_∞ approximations is given by

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

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

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

or, in terms of their basic probability assignments,

$$\begin{cases} \vec{m}^{L_\infty}_{\bar{A}}(A) = m_b(A) + \max_{B \not\in \mathcal{C}} m_b(B) \quad A \in \mathcal{C}, A \neq \bar{A} \\ \vec{m}^{L_\infty}_{\bar{A}}(\bar{A}) = m_b(\bar{A}) + \displaystyle\sum_{B \not\in \mathcal{C}} m_b(B) + \\ \qquad\qquad -(n-1) \max_{B \not\in \mathcal{C}} m_b(B). \end{cases} \qquad (21)$$

Note that such quantity is not guaranteed to be positive, as, for instance, when there exists a single subset B s.t. $m_b(B) \neq 0$ outside $\mathcal{C}$, $\vec{m}^{L_\infty}_{\bar{A}}(\bar{A})$ is negative unless $n \leq 2$. The barycenter of this simplex can be computed as follows:

$$m^{\overline{L_\infty}}(A) = \frac{\displaystyle\sum_{\bar{A} \in \mathcal{C}} \vec{m}^{L_\infty}_{\bar{A}}(A)}{n} = m_b(A) + \frac{\sum_{B \not\in \mathcal{C}} m_b(B)}{n},$$

i.e., the L_2 partial approximation. The corresponding minimal L_∞ norm of the difference vector is, according to (18), equal to $\max_{B \not\in \mathcal{C}} m_b(B)$.

Case 2. In the second case $k_1 > 3k_2$, or for us

$$\max_{B \not\in \mathcal{C}} m_b(B) < \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B), \qquad (22)$$

the contour function of (19) is as in Figure 4. 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 $[(-1/m)k_1, \cdots, (-1/m)k_1]'$, i.e., for system (18), $\beta(A) = -\frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B)$ for all $A \in \mathcal{C}, A \neq \Theta$ or, in terms of b.p.a.s,

$$m_{co_{L_\infty}[m_b]}(A) = m_b(A) + \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}.$$

The mass of Θ is obtained by normalization.
The corresponding minimal L_∞ norm of the difference vector is $\frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B)$.

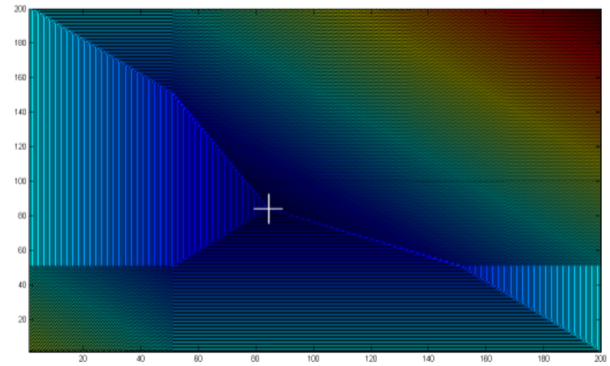

Figure 4: Contour function (level sets) and minimal point (white cross) of a function of the form (19), when $k_1 \geq 3k_2$. In the example $k_2 = 0.1$ and $k_1 = 0.5$.

3.3.2 $\mathbb{R}^{N-2}$ representation

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

$$\begin{aligned} \|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} &= \max_{\emptyset \subsetneq A \subsetneq \Theta} |m_b(A) - m_{co}(A)| \\ &= \max\Big\{ \max_{A \in \mathcal{C}, A \neq \Theta} |\beta(A)|, \max_{B \not\in \mathcal{C}} m_b(B) \Big\} \end{aligned} \qquad (23)$$

which is minimized by

$$|\beta(A)| \leq \max_{B \not\in \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta \qquad (24)$$

i.e., in the original mass coordinates,

$$\begin{aligned} & m_b(A) - \max_{B \not\in \mathcal{C}} m_b(B) \leq m_{co}(A) \leq \\ & \qquad \leq m_b(A) + \max_{B \not\in \mathcal{C}} m_b(B) \quad \forall A \in \mathcal{C}, A \neq \Theta. \end{aligned} \qquad (25)$$

According to (23) the corresponding minimal L_∞ norm is equal to $\max_{B \not\in \mathcal{C}} m_b(B)$.
Clearly, the vertices of the set (24) 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.:

$$m_{co}(B) = \begin{cases} m_b(B) & B \in \mathcal{C}, B \neq \Theta \\ m_b(B) + \sum_{B \not\in \mathcal{C}} m_b(B) & B = \Theta. \end{cases} \qquad (26)$$

Summarizing:

Theorem 3. *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*

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

with vertices (21) 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 (16). In both cases, the global L_∞ consonant *approximation is associated with the maximal chain of focal elements:*

$$\mathcal{C}^* = \arg\min_{\mathcal{C}} \max_{B \not\in \mathcal{C}} m_b(B). \quad (27)$$

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 (25). Its barycenter reassigns all the mass outside the chain to Θ, leaving the masses of the other elements untouched. The related global approximations of b are associated with the same optimal chain (27).

4 Semantics

Let us interpret the results we obtained in terms of basic probability assignments of the various consonant approximations, and compare those results with the outer consonant approximations [11] whose geometry has been described in [7].

Summary of approximations in $\mathcal{M}$. We can summarize all the results obtained here in the following tables. In the $\mathbb{R}^{N-1}$ mass representation the partial L_p approximations are:

$$\begin{array}{rcl} co^{\mathcal{C}}_{L_1,N-1}[m_b] & = & Cl(\vec{m}^{L_1}_A, A \in \mathcal{C}) \\ & : & m_{co}(A) \geq m_b(A) \ \forall A \in \mathcal{C}; \\ co^{\mathcal{C}}_{\overline{L_1},N-1}[m_b] & = & co^{\mathcal{C}}_{L_2,N-1}[b] \\ & : & m_{co}(A) = m_b(A) + \frac{\sum_{B \not\in \mathcal{C}} m_b(B)}{n}. \end{array} \quad (28)$$

Concerning the L_∞ approximation, if (20) holds

$$\begin{array}{rcl} co^{\mathcal{C}}_{L_\infty,N-1}[m_b] & = & Cl(\vec{m}^{L_\infty}_{\bar{A}}, \bar{A} \in \mathcal{C}); \\ co^{\mathcal{C}}_{\overline{L_\infty},N-1}[m_b] & = & co^{\mathcal{C}}_{L_2,N-1}[m_b], \end{array}$$

while if (22) 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 (12));
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, it 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 [6]. There,

Proposition 1. *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 is a consistent behavior of L_1 and L_2 minimization in the $\mathbb{R}^{N-1}$ representation of the mass space.
In the $\mathbb{R}^{N-2}$ mass representation the corresponding partial L_p approximations are:

$$\begin{array}{rcl} co^{\mathcal{C}}_{L_\infty,N-2}[b] & : & |m_{co}(A) - m_b(A)| \leq \max_{B \not\in \mathcal{C}} m_b(B) \\ & & \forall A \in \mathcal{C}, A \neq \Theta; \\ co^{\mathcal{C}}_{\overline{L_\infty},N-2}[b] & = & co^{\mathcal{C}}_{L_1,N-2}[b] = co^{\mathcal{C}}_{L_2,N-2}[b] \\ & : & \left\{ \begin{array}{l} m_{co}(A) = m_b(A), \ A \in \mathcal{C}, \neq \Theta \\ m_{co}(\Theta) = m_b(\Theta) + \sum_{B \not\in \mathcal{C}} m_b(B). \end{array} \right. \end{array} \quad (29)$$

We can notice a number of facts here too:

1. 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: $m_b(A) \leq \max_{B \not\in \mathcal{C}} m_b(B)$;
2. it forms a polytope in the mass space $\mathcal{M}$, whose size is determined by the largest mass outside the desired maximal chain;
3. the L_1 and L_2 partial approximations are uniquely determined, and coincide with the barycenter of the set of L_∞ partial approximations;
4. 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 [6].

Relation with outer consonant approximations. Let us recall the main results on the geometry of outer consonant approximations [7].

Proposition 2. *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

$$m_{o^{\vec{B}}[b]}(B_i) = \sum_{A \subseteq \Theta: \vec{B}(A) = A_i} m_b(A), \qquad (30)$$

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

The points (30) 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 $A_i = \{x_{\rho(1)}, \cdots, x_{\rho(i)}\}$, $i = 1, ..., n$, i.e.

$$m_{co^\rho}(A_i) = \sum_{B \subseteq A_i, B \not\subset A_{i-1}} m_b(B), \qquad (31)$$

is an actual vertex of $O_{\mathcal{C}}[b]$, and corresponds to the *maximal* outer consonant approximation with maximal chain $\mathcal{C}$.
Indeed, by Equation (11), 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).
It can be 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.

Comparison on a ternary example. It can therefore be useful to compare the different approximations in the toy case of a ternary frame, $\Theta = \{x, y, z\}$, to look for insights. Let us assume that we want the consonant approximation to have maximal chain $\mathcal{C} = \{\{x\}, \{x, y\}, \Theta\}$.
Figure 5 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

$$m_b(x) = 0.2, m_b(y) = 0.3, m_b(x, z) = 0.5 \qquad (32)$$

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 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 [5], the set of outer consonant approximations of (32) with chain $\{\{x\}, \{x, y\}, \Theta\}$ is the convex closure of the points:

$$\begin{array}{ll} \vec{m}_{B_1,B_2} & = [m_b(x), m_b(y), 1 - m_b(x) - m_b(y)]', \\ \vec{m}_{B_3,B_4} & = [m_b(x), 0, 1 - m_b(x)]', \\ \vec{m}_{B_5,B_6} & = [0, m_b(x) + m_b(y), 1 - m_b(x) - m_b(y)]', \\ \vec{m}_{B_7,B_8} & = [0, m_b(x), 1 - m_b(x)]', \\ \vec{m}_{B_9,B_{10}} & = [0, m_b(y), 1 - m_b(y)]', \\ \vec{m}_{B_{11},B_{12}} & = [0, 0, 1]', \end{array} \qquad (33)$$

These points are plotted as light blue squares in Figure 5. We can notice many interesting things.

1. 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 $O^{\mathcal{C}}[b]$;
2. on the contrary, outer approximations and $L_{1,N-1}$ approximations are mutually exclusive, as it can be inferred by Equation (11);
3. the maximal outer approximation co^ρ lies on the border between the two, where $m_{co}(x, y) = m_b(x, y)$.

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

5 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 the mass space of basic probability vectors. Interpretations for such approximations 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.
The nature of L_p-induced consonant approximations in the belief space remains an open problem, as is a comprehensive analysis of consonant and consistent approximations induced by distance minimization.

References

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

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

[3] F. Cuzzolin. Two new Bayesian approximations of belief functions based on convex geometry. *IEEE Transactions on SMC - Part B*, 37:4, 993–1008, 2007.

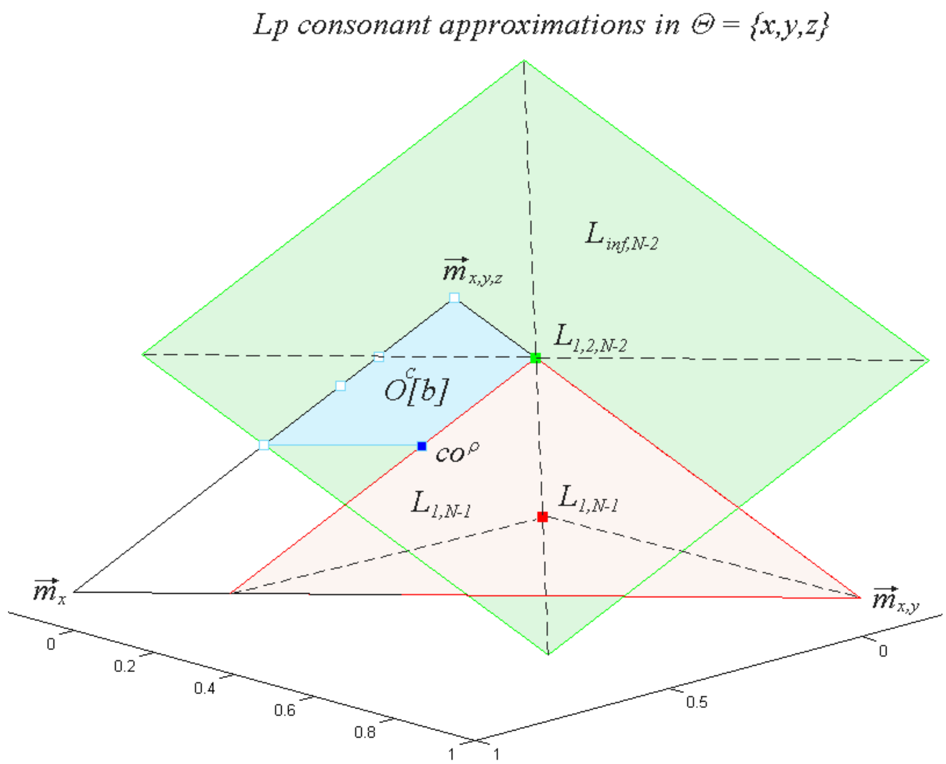

Figure 5: The simplex (solid black triangle) of consonant belief functions with maximal chain $\{\{x\}, \{x, y\}, \Theta\}$, and the L_p partial consonant approximations of the belief function with basic probabilities (32) on $\Theta = \{x, y, z\}$. The related set of outer consonant approximations (33) is also shown.

[4] F. Cuzzolin. A geometric approach to the theory of evidence. *IEEE Transactions on SMC - Part C*, 38:4, 522–534, 2008.

[5] F. Cuzzolin. Complexes of outer consonant approximations. *Proceedings of ECSQARU'09*, Verona, Italy, 2009.

[6] F. Cuzzolin. Geometric conditioning of belief functions. *Proceedings of BELIEF'10*, Brest, France, 2010.

[7] F. Cuzzolin. The geometry of consonant belief functions: simplicial complexes of necessity measures. *Fuzzy Sets and Systems*, 161:10, 1459–1479, 2010.

[8] F. Cuzzolin. Consistent approximation of belief functions. *Proceedings of ISIPTA'09*, Durham, UK, June 2009.

[9] J. Diaz, M. Rifqi and B. Bouchon-Meunier. A similarity measure between basic belief assignments. *Proceedings of FUSION'06*, 2006.

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

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

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

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

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

[15] C. Joslyn. Possibilistic normalization of inconsistent random intervals. *Advances in Systems Science and Applications*, 44–51, 1997.

[16] C. Joslyn and G. Klir. Minimal information loss possibilistic approximations of random sets, 1992.

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

[18] A.-L. Jousselme and P. Maupin. Distances in evidence theory: Comprehensive survey and generalizations. *International Journal of Approximate Reasoning*, 2011 (under review).

[19] V. Khatibi and G. A. Montazer. A new evidential distance measure based on belief intervals. *Scientia Iranica - Transactions D: Computer Science and Engineering and Electrical Engineering*, 17:2, 119–132, 2010.

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

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

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

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

[24] Ph. Smets. The transferable belief model and possibility theory. *Proceedings of NAFIPS-90*, 215–218, 1990.

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