Geometric conditional belief functions in the belief space

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

Abstract

In this paper we study the problem of conditioning a belief function (b.f.) b with respect to an event A by geometrically projecting such belief function onto the simplex associated with A in the space of all belief functions. Defining geometric conditional b.f.s by minimizing L_p distances between b and the conditioning simplex in such "belief" space (rather than in the "mass" space) produces complex results with less natural interpretations in terms of degrees of belief. The question of weather classical approaches, such as Dempster's conditioning, can be themselves reduced to some form of distance minimization remains open: the generation of families of combination rules generated by (geometrical) conditioning appears to be the natural prosecution of this line of research.

Keywords. Belief functions, conditioning, geometric approach, belief space, Lp norms.

1 Introduction

The original proposal for a conditioning operator in Dempster-Shafer's theory of belief functions (b.f.s) [30] is due to Dempster himself [13], who formulated it in his original model in which belief functions are induced by multi-valued mappings $\Gamma : \Omega \to 2^\Theta$ of probability distributions defined on a set Ω onto the power set of another set ("frame") Θ. However, Dempster's conditioning was almost immediately and strongly criticized from a Bayesian point of view.
In response to these objections a number of approaches to conditioning in the framework of belief functions (b.f.s) have been proposed along the years [25, 3, 15, 20, 17, 14, 41, 19, 39], in different mathematical setups. In the framework of credal sets and lower probabilities Fagin and Halpern defined a notion of *conditional belief* [15] as the lower envelope of a family of conditional probability functions, and provided a closed-form expression for it. In the context of multi-valued mappings, Spies [37] defined conditional events as sets of equivalent events under conditioning. By applying multi-valued mapping to such events, conditional belief functions were introduced. In Slobodova's work [33] a multi-valued extension of conditional b.f.s was introduced [32], and its properties examined. Klopotek and Wierzchon [23] provided a frequency-based interpretation for conditional belief functions.
Another way of dealing with the classical Bayesian criticism of Dempster's rule of conditioning is to abandon all notions of multivalued mapping to define belief directly on the power set of the frame as in Smets' Transferable Belief Model [34]. The (unnormalized) conditional b.f. $b_U(B|A) = \sum_{C \subseteq B} m_U(C|A)$ s.t.

$$m_U(B|A) = \begin{cases} \sum_{C \subseteq A^c} m(B \cup C) & if\ B \subseteq A \\ 0 & elsewhere \end{cases}$$

is the minimal commitment specialization of b such that $pl_b(A^c|A) = 0$ [22]. In [35], Smets pointed out the distinction between "revision" and "focussing" in the conditional process, and the way they lead to unnormalized and geometric [38] conditioning, respectively. In these two scenarios he proposed generalizations of Jeffrey's rule of conditioning [21, 31] to the framework of belief functions.

A geometric approach to conditioning. Quite recently, the idea of formulating the problem geometrically has emerged. Lehrer [26], in particular, proposed such a geometric approach to determine the conditional expectation of non-additive probabilities (such as belief functions).
The notion of generating conditional belief functions by minimizing a suitable distance function between the original b.f. b and the "conditioning region" $\mathcal{B}_A$, i.e., the set of belief functions whose b.b.a. assigns mass to subsets of A only

$$b_d(.|A) = \arg\min_{b' \in \mathcal{B}_A} d(b, b') \tag{1}$$

has a clear potential. It expands our arsenal of possible approaches to the problem, and is a promising candidate to the role of general framework for conditioning. The geometry of set functions and other uncertainty measures has been studied by different authors [2, 12, 28]. A similar approach has been developed and applied by the author more specifically to the theory of evidence [4, 6, 8]. Most interestingly, Jousselme et al [1] have 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 plugged in the above minimization problem (1) to define conditional belief functions. In [8] the author computed all the conditional belief functions generated via minimization of L_p norms in the "mass space", where b.f.s are represented by the vector of their basic probabilities.

Contribution and paper outline. In this paper we explore geometric conditioning in the *belief space* $\mathcal{B}$, in which belief functions are represented by the vectors of their belief values $b(A)$. We adopt once again distance measures d of the classical L_p family, as a further step towards a complete analysis of the geometric approach to conditioning. We show that geometric conditional b.f.s in $\mathcal{B}$ are more complex than in the mass space, less naive objects whose interpretation in terms of degrees of belief is however less natural.
We first briefly recall in Section 2 the geometric approach to belief functions. We later compute the analytical form of the L_2 (Section 3) and L_∞ (Section 4) conditional b.f.s in $\mathcal{B}$, and propose a preliminary interpretation for them. We extensively compare the obtained results with those of [8] in Section 5. We conclude in Section 6 by prospecting a number of developments for the geometric approach to conditioning. As some of the proof are rather combinatorial, we moved them to an Appendix.

2 Geometric representation

2.1 Belief functions as vectors

As belief functions $b : 2^\Theta \rightarrow [0,1]$, $b(A) = \sum_{B\subseteq A} m_b(B)$, are set functions defined on a the power set 2^Θ of a finite space Θ, they are obviously completely defined by the associate set of $2^{|\Theta|}-2$ belief values $\{b(A), \emptyset \subsetneq A \subsetneq \Theta\}$ (since $b(\emptyset) = 0$, $b(\Theta) = 1$ for all b.f.s). They can therefore be represented as points of $\mathbb{R}^{N-2}, N = 2^{|\Theta|}$ [6]. The set $\mathcal{B}$ of points of $\mathbb{R}^{N-2}$ which correspond to a belief function is a *simplex* called *belief space*, namely:

$$\mathcal{B} = Cl(\vec{b}_A,\ \emptyset \subsetneq A \subseteq \Theta),$$

where Cl denotes the convex closure operator[1] and $\vec{b}_A$ is the vector associated with the categorical [34] belief function b_A assigning all the mass to a single subset $A \subseteq \Theta$: $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0$ for all $B \neq A$. The vector $\vec{b} \in \mathcal{B}$ that corresponds to a belief function b has in $\mathcal{B}$ coordinates $m_b(A)$:

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

2.2 Conditioning simplex

Similarly, the vector $\vec{a}$ associated with any belief function whose mass supports only focal elements $\{\emptyset \subsetneq B \subseteq A\}$ included in a given conditioning event A can be decomposed as:

$$\vec{a} = \sum_{\emptyset \subsetneq B \subseteq A} m_a(B)\vec{b}_B. \tag{2}$$

The set of such vectors form a simplex

$$\mathcal{B}_A \doteq Cl(\vec{b}_B,\ \emptyset \subsetneq B \subseteq A)$$

which we call the *conditioning simplex* (in the belief space).

2.3 Geometric conditional belief functions

Given a belief function b, we call *geometric conditional belief function induced by a distance function d in $\mathcal{B}$* the b.f.(s) $b_{d,\mathcal{B}}(.|A)$ which minimize(s) the distance $d(b, \mathcal{B}_A)$ between b and the conditioning simplex $\mathcal{B}$. As recalled above, a large number of proper distance functions or mere dissimilarity measures between belief functions have been proposed in the past, and many others can be designed [1].
We consider here as distance functions the major L_p norms $d = L_2$ and $d = L_\infty$. This is not to claim that these are *the* distance function of choice for this problem. The geometric approach to conditioning is potentially a huge research direction which will take much time to explore exhaustively. However, L_p norms have been used in the probability and the consistent approximation problems [4, 10], where they have produced rather simple and elegant results. For vectors $\vec{b}$, $\vec{b'} \in \mathcal{B}$ representing two b.f.s b, b', such norms read as:

$$\|\vec{b} - \vec{b'}\|_2 \doteq \Big[\sum_{\emptyset \subsetneq B \subseteq \Theta} (b(B) - b'(B))^2\Big]^{1/2}$$
$$\|\vec{b} - \vec{b'}\|_\infty \doteq \max_{\emptyset \subsetneq B \subseteq \Theta} |b(B) - b'(B)|.$$

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

3 L_2 conditioning in the belief space

The problem of projecting a belief function b represented by the corresponding vector $\vec{b}$ of belief values onto a conditioning simplex $\mathcal{B}_A = Cl(\vec{b}_B, \emptyset \subsetneq B \subseteq A)$ starts by explicitly writing the difference vector $\vec{b} - \vec{a}$ between $\vec{b}$ and an arbitrary point $\vec{a}$ of $\mathcal{B}_A$, i.e.,

$$\begin{array}{l} \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B)\vec{b}_B - \sum_{\emptyset \subsetneq B \subseteq A} m_a(B)\vec{b}_B = \\ = \sum_{\emptyset \subsetneq B \subseteq A} (m_b(B) - m_a(B))\vec{b}_B + \sum_{B \not\subseteq A} m_b(B)\vec{b}_B \\ = \sum_{\emptyset \subsetneq B \subseteq A} \beta(B)\vec{b}_B + \sum_{B \not\subseteq A} m_b(B)\vec{b}_B \end{array} \quad (3)$$

where $\beta(B) = m_b(B) - m_a(B)$. We exploit the fact that the minimal L_2 distance between a point and a vector space V is attained by the point of the vector space such that the difference vector is orthogonal to all the generators $\vec{g}_i$ of V:

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

whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$ (Figure 1).

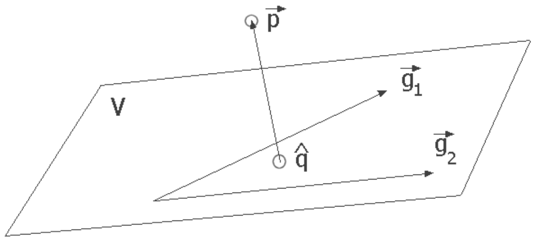

Figure 1: The point $\vec{q}$ of a vector space V at minimal L_2 distance from a given point $\vec{p} \notin V$ is such that the difference vector $\vec{p} - \vec{q}$ is orthogonal to all the generators $\vec{g}_i$ of V.

3.1 A case study: the ternary frame

In the case of a size-3 frame $\Theta = \{x, y, z\}$, and a conditioning event of size 2 (for instance $A = \{x, y\}$) the difference vector (3) particularizes as:

$$\begin{array}{l} \vec{b} - \vec{a} = \beta(x)\vec{b}_x + \beta(y)\vec{b}_y + \beta(x,y)\vec{b}_{\{x,y\}} + m_b(z)\vec{b}_z \\ \qquad + m_b(x,z)\vec{b}_{\{x,z\}} + m_b(y,z)\vec{b}_{\{y,z\}}, \end{array} \quad (4)$$

as $\vec{b}_\Theta = \vec{0}$. To find the vector $\vec{a} \in \mathcal{B}_A$ which minimizes the L_2 distance from $\vec{b}$ (its L_2 conditional belief function $b_{L_2,\mathcal{B}}(.|A)$ with respect to A in $\mathcal{B}$) we need to enforce the orthogonality conditions:

$$\langle \vec{b} - \vec{a}, \vec{b}_x - \vec{b}_{\{x,y\}} \rangle = 0, \quad \langle \vec{b} - \vec{a}, \vec{b}_y - \vec{b}_{\{x,y\}} \rangle = 0$$

as $\mathcal{B}_{\{x,y\}} = Cl(\vec{b}_x, \vec{b}_y, \vec{b}_{\{x,y\}})$ has two generators: $\vec{b}_x - \vec{b}_{\{x,y\}}$ and $\vec{b}_y - \vec{b}_{\{x,y\}}$. Remembering that

$$\langle \vec{b}_A, \vec{b}_B \rangle = \langle \vec{b}_{A\cup B}, \vec{b}_{A\cup B} \rangle = 2^{|(A\cup B)^c|} - 1, \quad (5)$$

after simple maths we obtain the following system

$$\left\{ \begin{array}{l} 2\beta(x) + m_b(z) + m_b(x,z) = 0 \\ 2\beta(y) + m_b(z) + m_b(y,z) = 0 \end{array} \right. \quad (6)$$

whose solution is clearly (in $m_a(x)$, $m_a(y)$)

$$\begin{array}{l} m_a(x) = m_b(x) + \frac{m_b(z) + m_b(x,z)}{2}, \\ m_a(y) = m_b(y) + \frac{m_b(z) + m_b(y,z)}{2}, \\ m_a(x,y) = m_b(x,y) + m_b(\Theta) + \frac{m_b(x,z) + m_b(y,z)}{2}. \end{array} \quad (7)$$

At a first glance, each focal element $B \subseteq A$ seems to be assigned to a fraction of the original mass $m_b(X)$ of all focal elements X of b such that $X \subseteq B \cup A^c$. This contribution seems proportional to the size of $X \cap A^c$, i.e., how much the focal element of b falls outside the conditioning event A. Notice that Dempster's conditioning $b_\oplus(.|A) = b \oplus b_A$ yields in this case:

$$\begin{array}{lcl} m_\oplus(x|A) & = & \frac{m_b(x) + m_b(x,z)}{1 - m_b(z)}, \\ m_\oplus(x,y|A) & = & \frac{m_b(x,y) + m_b(\Theta)}{1 - m_b(z)}. \end{array}$$

L_2 conditioning in the belief space differs from its "sister" operation in the mass space [8] in that it makes use of the set-theoretic relations between focal elements, as Dempster's rule does. However, contrarily to Dempster's conditioning it does not apply any normalization, as *even subsets of* A^c ($\{z\}$ in this case) *contribute as addenda* to the mass of the resulting conditional belief function.

3.2 General case

In the general case the orthogonality of the difference vector w.r.t. the generators $\vec{b}_C - \vec{b}_A$, $\emptyset \subsetneq C \subsetneq A$ of the conditional simplex

$$\langle \vec{b} - \vec{a}, \vec{b}_C - \vec{b}_A \rangle = 0 \quad \forall \emptyset \subsetneq C \subsetneq A$$

(where $\vec{b} - \vec{a}$ is given by Equation (3)) reads as

$$\left\{ \begin{array}{l} \sum_{B \not\subseteq A} m_b(B) \Big[\langle \vec{b}_B, \vec{b}_C \rangle - \langle \vec{b}_B, \vec{b}_A \rangle \Big] + \\ \sum_{B \subseteq A} \beta(B) \Big[\langle \vec{b}_B, \vec{b}_C \rangle - \langle \vec{b}_B, \vec{b}_A \rangle \Big] = 0 \end{array} \right.$$

for all $\emptyset \subsetneq C \subsetneq A$. Now, by (5), $\langle \vec{b}_B, \vec{b}_C \rangle = |\{Y \supseteq B \cup C, Y \neq \Theta\}| = 2^{|(B\cup C)^c|} - 1$ and $\langle \vec{b}_B, \vec{b}_A \rangle = 2^{|(B\cup A)^c|} - 1$. As $(B \cup A)^c = A^c$ when $B \subseteq A$, the system of orthogonality conditions is equivalent to, $\emptyset \subsetneq C \subsetneq A$,

$$\left\{ \begin{array}{l} \sum_{B \not\subseteq A} m_b(B) \Big[2^{|(B\cup C)^c|} - 2^{|(B\cup A)^c|} \Big] \\ + \sum_{B \subsetneq A} \beta(B) \Big[2^{|(B\cup C)^c|} - 2^{|A^c|} \Big] = 0 \end{array} \right. \quad (8)$$

This is a system of $2^{|A|}-2$ equations in the $2^{|A|}-2$ variables $\{\beta(B), B \subsetneq A\}$.

Theorem 1. *The L_2 conditional belief function $b_{L_2,\mathcal{B}}(.|A)$ with respect to A in the belief space $\mathcal{B}$ is unique, and is given by $\beta(C) =$*

$$= -\sum_{B \subseteq A^c} m_b(B \cup C) 2^{-|B|} + (-1)^{|C|} \sum_{B \subseteq A^c} m_b(B) 2^{-|B|} \tag{9}$$

i.e.,

$$m_{L_2,\mathcal{B}}(C|A) = m_b(C) + \sum_{B \subseteq A^c} m_b(B \cup C) 2^{-|B|} + \\ +(-1)^{|C|+1} \sum_{B \subseteq A^c} m_b(B) 2^{-|B|}$$

for each proper subset $\emptyset \subsetneq C \subsetneq A$ of the event A.

As a verification, in the case $\Theta = \{x, y, z\}$, $A = \{x, y\}$ we have, as in (7),

$$\begin{aligned} m_a(x) &= m_b(x) + m_b(\{z\} \cup \{x\}) 2^{-1} + (-1)^2 m_b(z) 2^{-1} \\ &= m_b(x) + \frac{m_b(z) + m_b(x,z)}{2}; \\ m_a(y) &= m_b(y) + m_b(\{z\} \cup \{y\}) 2^{-1} + (-1)^2 m_b(z) 2^{-1} \\ &= m_b(y) + \frac{m_b(z) + m_b(y,z)}{2}. \end{aligned}$$

We can notice that the (unique) L_2 conditional belief function in the belief space is not always a proper belief function, as some masses can be negative, due to the addendum $(-1)^{|C|+1} \sum_{B \subseteq A^c} m_b(B) 2^{-|B|}$. It shows, however, an interesting connection with the redistribution process associated with the orthogonal projection $\pi[b]$ of a belief function onto the probability simplex [5]. Such projection can indeed be expressed as $\pi[b] = \overline{\mathcal{P}}(1-k) + kO[b]$, where $\mathcal{P}$ is the uniform probability distribution, $k \in [0,1]$, and $O[b]$ is another probability obtained by first re-distributing the mass of each set A among all its subsets $B \subseteq A$ on an equal basis, to later compute its relative belief of singletons $p(x) = b(x)/\sum_y b(y)$. Here (9) the mass of each focal element not included in A is also broken into $2^{|B|}$ parts, proportionally to the number of its subsets. Only one such part is re-attributed to $C = B \cap A$, while the rest is re-attributed to A itself.

L_2 minimization seems to be related to factors of the form $2^{-|B|}$, which imply mass redistribution in the power set 2^{Θ}.

4 L_∞ conditioning in the belief space

We proceed to computing the L_∞ conditional belief functions in the belief space, starting with the ternary case study.

4.1 The ternary case

There $\|\vec{b} - \vec{a}\|_{L_\infty} = \max_{\emptyset \subsetneq B \subsetneq \Theta} |b(B) - a(B)| =$

$$\max\Big\{|b(x)-a(x)|, |b(y)-a(y)|, |b(x,y)-a(x,y)|, \\ |b(z)|, |b(x,z)-a(x,z)|, |b(y,z)-a(y,z)|\Big\} = \\ \max\Big\{m_b(z), 1-b(x,y), |\beta(x)+m_b(z)+m_b(x,z)|, \\ |\beta(x)|, |\beta(y)|, |\beta(y)+m_b(z)+m_b(y,z)|\Big\}$$

which is minimized for, as $1 - b(x,y) \geq m_b(z)$,

$$\max\Big\{|\beta(x)|, |\beta(x)+m_b(z)+m_b(x,z)|\Big\} \leq 1 - b(x,y) \\ \max\Big\{|\beta(y)|, |\beta(y)+m_b(z)+m_b(y,z)|\Big\} \leq 1 - b(x,y).$$

This involves functions of the form $\max\{|x|, |x+k|\}$,

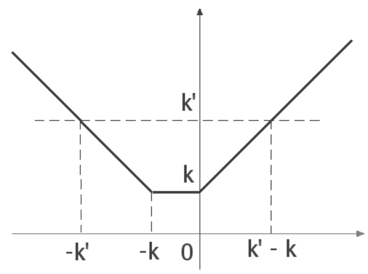

Figure 2: Graph of a function of the form $\max\{|x|, |x+k|\}$, and the interval of values below a threshold $k' >= k$.

whose shape is plotted in Figure 2. The interval of values in which such a function is below a certain threshold $k' \geq k$ is $[-k', k'-k]$, yielding

$$b(x,y) - 1 \leq \beta(x) \leq 1 - b(x,y) - (m_b(z) + m_b(x,z)) \\ b(x,y) - 1 \leq \beta(y) \leq 1 - b(x,y) - (m_b(z) + m_b(y,z)). \tag{10}$$

The **solution in the masses** of the sought L_∞ conditional b.f. reads therefore as

$$m_b(x) - m_b(y,z) - m_b(\Theta) \leq m_a(x) \leq \\ \leq 1 - (m_b(y) + m_b(x,y)) \\ m_b(y) - m_b(x,z) - m_b(\Theta) \leq m_a(y) \leq \\ \leq 1 - (m_b(x) + m_b(x,y)). \tag{11}$$

Its barycenter is clearly given by

$$\begin{aligned} m_a(x) &= m_b(x) + \tfrac{m_b(z)+m_b(x,z)}{2} \\ m_a(y) &= m_b(y) + \tfrac{m_b(z)+m_b(y,z)}{2} \\ m_a(x,y) &= m_b(x,y) + m_b(\Theta) + \tfrac{m_b(x,z)+m_b(y,z)}{2} \end{aligned} \tag{12}$$

i.e., it coincides with the L_2 conditional belief function (7) as computed in the ternary case.

4.2 The general case

From (19) the norm becomes, after introducing the variable $\gamma(C)=\sum_{X\subseteq C}\beta(X)$, $\max_{\emptyset\subsetneq B\subsetneq\Theta}|\vec{b}-\vec{a}(B)|=$

$$
\begin{aligned}
&= \max_{\emptyset\subsetneq B\subsetneq\Theta}\Big|\sum_{C\subseteq A\cap B}\beta(C)+\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big| \\
&= \max_{\emptyset\subsetneq B\subsetneq\Theta}\Big|\gamma(A\cap B)+\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big| \\
&= \max\Bigg\{\max_{B:B\cap A=\emptyset}\Big|\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big|, \\
&\qquad \max_{B:B\cap A\neq\emptyset, A}\Big|\gamma(A\cap B)+\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big|, \\
&\qquad \max_{B:B\cap A=A, B\neq\Theta}\Big|\gamma(A)+\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big|\Bigg\},
\end{aligned}
\tag{13}
$$

where clearly

$$
\begin{aligned}
\gamma(A) &= \sum_{B\subseteq A}\beta(B)=\sum_{B\subseteq A}\Big(m_b(B)-m_a(B)\Big) \\
&= b(A)-\sum_{B\subseteq A}m_a(B)=b(A)-1.
\end{aligned}
$$

Now, the first term in the above max problem is s.t.

$$
\max_{B:B\cap A=\emptyset}\Big|\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big|=\max_{B\subseteq A^c}\Big|\sum_{C\subseteq B}m_b(C)\Big|=
$$

$=\max_{B\subseteq A^c}b(B)=b(A^c)$. For the third one we have

$$
\sum_{C\subseteq B, C\not\subset A}m_b(C)\leq\sum_{C\cap A^c}m_b(C)=pl_b(A^c)=1-b(A),
$$

which is maximized when $B=A$, in which case it is equal to

$$
\Big|b(A)-1+\sum_{C\subseteq A, C\not\subset A}m_b(C)\Big|=|b(A)-1+0|=1-b(A).
$$

Therefore, the L_∞ norm (13) of the difference $\vec{b}-\vec{a}$ reduces to $\max_{\emptyset\subsetneq B\subsetneq\Theta}|\vec{b}-\vec{a}(B)|=$

$$
\max\Bigg\{\max_{B:B\cap A\neq\emptyset, A}\Big|\gamma(A\cap B)+\sum_{\substack{C\subseteq B,\\ C\not\subset A}}m_b(C)\Big|, 1-b(A)\Bigg\}
\tag{14}
$$

and is obviously minimized by all the values of $\gamma^*(X)$ such that

$$
\max_{B:B\cap A\neq\emptyset, A}\Big|\gamma^*(A\cap B)+\sum_{C\subseteq B, C\not\subset A}m_b(C)\Big|\leq 1-b(A).
$$

Lemma 1. *The values $\gamma^*(X)$ which minimize the norm (13) are, $\forall\ \emptyset\subsetneq X\subsetneq A$:*

$$
-(1-b(A))\leq\gamma^*(X)\leq(1-b(A))-\sum_{\substack{C\cap A^c\neq\emptyset,\\ C\cap A\subseteq X}}m_b(C).
\tag{15}
$$

Solution in the masses. It is not difficult to see by induction that in the original auxiliary variables β the set of L_∞ conditional b.f.s in $\mathcal{B}$ is determined by the following constraints:

$$
\begin{aligned}
&-K(X)+(-1)^{|X|}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X}m_b(C)\leq \\
&\leq\beta(X)\leq K(X)-\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X}m_b(C)
\end{aligned}
\tag{16}
$$

where

$$
K(X)=(2^{|X|}-1)(1-b(A))-\sum_{C\cap A^c\neq\emptyset, \emptyset\subseteq C\cap A\subsetneq X}m_b(C).
$$

In the masses of the sought L_∞ conditional b.f. this reads as follows.

Theorem 2. *Given a belief function $b:2^\Theta\to[0,1]$ and an arbitrary non-empty focal element $\emptyset\subsetneq A\subseteq\Theta$, the set of L_∞ conditional belief functions $b_{L_\infty,\mathcal{B}}(.|A)$ with respect to A in $\mathcal{B}$ is the set of b.f.s with focal elements in $\{X\subseteq A\}$ which meet the following constraints for all $\emptyset\subsetneq X\subseteq A$:*

$$
\begin{aligned}
&m_b(X)+\sum_{\substack{C\cap A^c\neq\emptyset,\\ \emptyset\subseteq C\cap A\subseteq X}}m_b(C)+(2^{|X|}-1)(1-b(A)) \\
&\leq m_a(X)\leq m_b(X)+(2^{|X|}-1)(1-b(A))+ \\
&-\sum_{C\cap A^c\neq\emptyset, \emptyset\subseteq C\cap A\subsetneq X}m_b(C)-(-1)^{|X|}\sum_{B\subseteq A^c}m_b(B).
\end{aligned}
\tag{17}
$$

4.3 Barycenter of the L_∞ solution

Clearly the barycenter of the optimal region (15) is

$$
\gamma^*(X)=-\frac{1}{2}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X}m_b(C),
$$

a solution which corresponds, in the set of variables $\{\beta(X)\}$, to the system of equations: $\forall\emptyset\subsetneq X\subsetneq A$

$$
\Bigg\{\sum_{\emptyset\subsetneq C\subseteq X}\beta(C)+\frac{1}{2}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X}m_b(C)=0.
\tag{18}
$$

Theorem 3. *The center of mass of the set of L_∞ conditional belief functions $b_{L_\infty,\mathcal{B}}(.|A)$ with respect to A in the belief space $\mathcal{B}$ is the unique solution of the system (18), and is given by:*

$$
\beta(C)=\frac{1}{2}\sum_{\emptyset\subsetneq B\subseteq A^c}\Big[(-1)^{|C|}m_b(B)-m_b(B\cup C)\Big].
\tag{19}
$$

i.e., in terms of masses,

$$
\begin{aligned}
m_{\overline{L_\infty},\mathcal{B}}(C|A)&=m_b(C)+\frac{1}{2}\sum_{\emptyset\subsetneq B\subseteq A^c}m_b(B+C)+ \\
&+\frac{1}{2}(-1)^{|C|+1}b(A^c).
\end{aligned}
\tag{20}
$$

The mass of A can be easily obtained by normalization.

5 Comparison with conditioning in $\mathcal{M}$

It is interesting to compare these results with those obtained by L_p conditioning in the mass space [8].

5.1 L_p conditioning in the mass space

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 set of L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is the set of b.f.s with core in A such that their mass dominates that of b over all the subsets of A: $b_{L_1,\mathcal{M}}(.|A) =$*

$$= \Big\{ b' : \mathcal{C}_{b'} \subseteq A, m_{b'}(B) \geq m_b(B) \; \forall \emptyset \subsetneq B \subseteq A \Big\}. \tag{21}$$

Such a set is a simplex $\mathcal{M}_{L_1,A}[b] = Cl(\vec{m}[b]|^B_{L_1} A, \emptyset \subsetneq B \subseteq A)$ whose vertices $\vec{a} = \vec{m}[b]|^B_{L_1} A$ have b.p.a.:

$$\begin{cases} a(B) = m_b(B) + 1 - b(A) = m_b(B) + pl_b(A^c), \\ a(X) = m_b(X) \quad \forall \emptyset \subsetneq X \subsetneq A, X \neq B \end{cases} \tag{22}$$

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:

$$m_{L_2,\mathcal{M}}(B|A) = m_b(B) + \frac{pl_b(A^c)}{2^{|A|}-1}, \tag{23}$$

$\forall \emptyset \subsetneq B \subseteq A$, and corresponds to the center of mass of the simplex $\mathcal{M}_{L_1,A}[b]$ of L_1 conditional b.f.s.

L_1 and L_2 conditioning are strictly related in the mass space, the latter being the barycenter of the former, and they have a compelling interpretation in terms of general imaging [29, 16]. L_∞ conditioning in $\mathcal{M}$ is of more difficult interpretation.
The L_2 and L_∞ conditioning in the belief space computed here

$$\begin{array}{rcl} m_{L_2,\mathcal{B}}(B|A) & = & m_b(B) + \displaystyle\sum_{C \subseteq A^c} m_b(B+C) 2^{-|C|} \\ & & +(-1)^{|B|+1} \displaystyle\sum_{C \subseteq A^c} m_b(C) 2^{-|C|} \\ m_{\overline{L_\infty},\mathcal{B}}(B|A) & = & m_b(B) + \displaystyle\frac{1}{2} \sum_{\emptyset \subsetneq C \subseteq A^c} m_b(B+C) \\ & & + \displaystyle\frac{1}{2}(-1)^{|B|+1} b(A^c). \end{array}$$

According to Theorem 3 we can interpret the barycenter of the set of L_∞ conditional belief functions as follows: the mass of all the subsets whose intersection with A is $C \subsetneq A$ is re-assigned by the conditioning process *half to C*, and *half to A itself*. In the case of $C = A$ itself, by normalizarion, all the subsets $D \supseteq A$ including A have their whole mass re-assigned to A, consistently with the above interpretation. The mass of the subsets $B \subseteq A^c$ which have no relation with the conditioning event A is used to guarantee the normalization of the resulting mass distribution. As a result, the obtained mass function is not necessarily non-negative: again, such version of geometrical conditioning may generated pseudo belief functions. Furthermore, if we compare Theorem 3 with Theorem 1 we will notice that, while being quite similar to each other, the coincidence of $b_{L_\infty,\mathcal{B}}(.|A)$ and $b_{L_2,\mathcal{B}}(.|A)$ was really an artifact of the ternary case, for which $\emptyset \subsetneq B \subseteq A^c \equiv \emptyset \subsetneq B \subseteq \{z\}^c \equiv B = \{z\}, |B| = 1$.
Conditional b.f.s in the belief space seem to have rather less straightforward interpretations than the corresponding quantities in the mass space. A comparison between Equations (23) and (20) illustrates a clear difference: while in the $\mathcal{M}$ case the barycenter of L_1 conditional b.f.s is obtained by reassigning the mass of all $B \not\subset A$ to each $B \subsetneq A$ on equal grounds, for the barycenter of L_∞ conditional b.f.s in $\mathcal{B}$ normalization is achieved by adding or subtracting their masses according to the cardinality of C (even or odd).

5.2 Comparison on the ternary example

We conclude by illustrating the different approximations in the toy case of a ternary frame, $\Theta = \{x, y, z\}$, for sake of completeness. Assuming we want the conditioning event approximation to be $A = \{x, y\}$, the L_2 (partial) conditional belief function is given by Equation (7), while the L_∞ partial conditional b.f.s form the set determined by Equation (11), with barycenter in (12). By Proposition 1 the vertices of $\mathcal{M}_{L_1,A}[b]$ are

$$\begin{array}{l} \vec{m}^{\{x\}}_{L_1,\mathcal{M}} = [m_b(x) + pl_b(z), m_b(y), m_b(x,y)]', \\ \vec{m}^{\{y\}}_{L_1,\mathcal{M}} = [m_b(x), m_b(y) + pl_b(z), m_b(x,y)]', \\ \vec{m}^{\{x,y\}}_{L_1,\mathcal{M}} = [m_b(x), m_b(y), m_b(x,y) + pl_b(z)]'. \end{array}$$

while the L_2 conditional b.f. in $\mathcal{M}$ has b.b.a.:

$$\begin{array}{l} m(x) = m_b(x) + \dfrac{1 - b(x,y)}{3} = m_b(x) + \dfrac{pl_b(z)}{3}, \\ m(y) = m_b(y) + \dfrac{pl_b(z)}{3}, m(x,y) = m_b(x,y) + \dfrac{pl_b(z)}{3}. \end{array} \tag{24}$$

In this case the conditional simplex is 2-dimensional, with three vertices $\vec{b}_x$, $\vec{b}_y$ and $\vec{b}_{x,y}$. Figure 3 illustrates the different geometric conditional belief functions given $A = \{x, y\}$ for the belief function with masses

$$\vec{m} = [0.2,\ 0.3,\ 0,\ 0,\ 0.5,\ 0]', \tag{25}$$

i.e., $m_b(x) = 0.2$, $m_b(y) = 0.3$, $m_b(x,z) = 0.5$. It confirms that $m_{L_2,\mathcal{M}}(.|A)$ lies in the barycenter of the simplex of the related L_1 conditional b.f.s. The same is true (in the ternary case) for $m_{L_2,\mathcal{B}}(.|A)$ which is the barycenter of the (green) polytope of $m_{L_\infty,\mathcal{B}}(.|A)$ conditional b.f.s. We can notice that the latter does not fall entirely in the admissible conditional simplex $C(\vec{b}_x, \vec{b}_y, \vec{b}_{x,y})$, but a good portion of it does. A number of interesting cross relations between conditional b.f. of the two representation domain appear to exist.

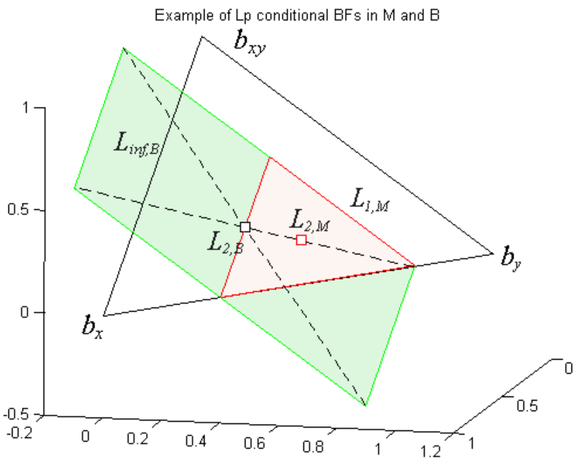

Figure 3: The simplex (red triangle) of $L_1, \mathcal{M}$ conditional belief functions associated with the belief function with mass assignment (25) in $\Theta = \{x,y,z\}$, with conditioning event $A = \{x,y\}$. The related $L_2, \mathcal{M}$ conditional belief function is plotted as a red square, and coincides with the center of mass of the L_1 set. The set of $L_\infty, \mathcal{B}$ conditional b.f.s is drawn as a green rectangle, and falls in part outside the conditioning simplex (black triangle). In the ternary case $L_2, \mathcal{B}$ is the barycenter of this rectangle. Interesting cross - relations between conditional functions in $\mathcal{M}$ and $\mathcal{B}$ seem to emerge.

1. $m_{L_\infty,\mathcal{B}}(.|A)$ seems to contain $m_{L_1,\mathcal{M}}(.|A)$, while
2. the two L_2 conditional b.f.s $m_{L_2,\mathcal{M}}(.|A)$ and $m_{L_2,\mathcal{B}}(.|A)$ appear to lie on the same line.

There is probably more to these conditioning approaches than the simple comparison run here. For instance, it remains an open problem to find the admissible part of $m_{L_\infty,\mathcal{B}}(.|A)$. We will investigate them further in the near future.

6 Conclusions and perspectives

In this paper we showed how the notion of conditional belief function $b(.|A)$ can be introduced by geometric means, by projecting any belief function onto the the simplicial subspace associated with the event A. The result will obviously depend on the choice of the vectorial representation for b, and of the distance function to minimize. We thoroughly analyzed the case of conditioning a belief vector by means of the norms L_2, and L_∞, and showed that the results often have somehow complex interpretations in terms of degrees of belief. We compared such results with those obtained in [8] for in the mass space. The present analysis opens a number of interesting questions.

6.1 Geometrical rules of combination

We may wonder, for instance, what classes of conditioning rules can be generated by such a process. Do they span all known definitions of conditioning? In particular, is Dempster's conditioning itself a special case of geometric conditioning? We already mentioned Jousselme et al [1] and their survey of the distance or similarity measures so far introduced between belief functions. This line of research could possibly be very useful in such a quest.
A related question links geometric conditioning with combination rules. In the case of Dempster's rule it can be easily proven that [4],

$$b \oplus b' = b \oplus \sum_{A \subseteq \Theta} m'(A) b_A = \sum_{A \subseteq \Theta} \mu(A) b \oplus b_A,$$

where as usual b' is decomposed as a convex combination of categorical belief functions b_A, and $\mu(A) \propto m'(A) pl_b(A)$. This means that Dempster's combination can be decomposed into a convex combination of Dempster's conditioning with respect to all possible events A. We can imagine to reverse this link, and generate combination rules starting from conditioning rules. Additional conditions have to be imposed in order to obtain a unique such combination rule. For instance, by imposing commutativity with affine combination (*linearity*, in Smets' terminology [36]), any (geometrical) conditioning rule $b|_A^{\uplus}$ implies necessarily:

$$b \uplus b' = \sum_{A \subseteq \Theta} m'(A) b \uplus b_A = \sum_{A \subseteq \Theta} m'(A) b|_A^{\uplus}.$$

In the near future we plan to explore the world of combination rules induced by conditioning rules, starting from the different geometrical conditional processes introduced here.

6.2 L_p consonant approximation

In addition, the same techniques used to project a belief function onto a conditional simplex can be used to solve the approximation problem.

It has been proven in [9] that consonant belief functions (b.f.s whose focal elements are nested), for instance, live in a "simplicial complex", i.e., a structured collection of simplices. The same holds of consistent belief functions, i.e., belief functions whose focal elements have non-empty intersection [10].
Consonant or consistent approximations of belief functions can therefore be obtained by minimizing L_p distances between the original b.f. and the relevant simplicial complex. The analysis has been conducted in the case of consistent belief functions, but only in the belief space, in [10, 7]: the present results suggest L_p approximations in the mass space should turn out to be more elegant and easier to interpret. Given the practical importance of possibility measures and fuzzy sets among practitioners, we plan to do the same for consonant belief functions.

Appendix

Proof of Theorem 1. To prove Theorem 1 we just need to replace (9) into the system of constraints (8) for the L_2 solution in the belief space case. We obtain, for all $\emptyset \subsetneq C \subsetneq A$:

$$\sum_{B \not\subseteq A} m_b(B)\Big[2^{|(B\cup C)^c|} - 2^{|(B\cup A)^c|}\Big] + \sum_{B\subsetneq A}\Big[-\sum_{X\subseteq A^c} m_b(X\cup B)2^{-|X|} + (-1)^{|B|}\sum_{X\subseteq A^c} m_b(X)2^{-|X|}\Big]\cdot \Big[2^{|(B\cup C)^c|} - 2^{|A^c|}\Big] = 0.$$

Now, whenever $B \not\subset A$ it can be decomposed as $B = X + Y$, with $\emptyset \subsetneq X \subseteq A^c$, $\emptyset \subseteq Y \subseteq A$. Therefore $B \cup C = (Y \cup C) + X$, $B \cup A = A + X$ and, since

$$2^{-|X|}(2^{|(Y\cup C)^c|} - 2^{|A^c|}) = 2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|},$$

we can write the constraints as

$$\sum_{\emptyset\subsetneq X\subseteq A^c}\sum_{\emptyset\subseteq Y\subseteq A} m_b(X+Y)\Big(2^{|[(Y\cup C)+X]^c|} + -2^{|(A+X)^c|}\Big) + \sum_{\emptyset\subsetneq X\subseteq A^c}\sum_{\emptyset\subsetneq Y\subsetneq A}\Big((-1)^{|Y|}m_b(X) + -m_b(X+Y)\Big)\Big(2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|}\Big) = 0.$$

Some combinatorial operations lead to the simplified form:

$$\sum_{\emptyset\subsetneq X\subseteq A^c} m_b(X)\Big(2^{|(X+C)^c|} - 2^{|(X+A)^c|}\Big) + \sum_{\emptyset\subsetneq X\subseteq A^c}\sum_{\emptyset\subsetneq Y\subsetneq A}(-1)^{|Y|}m_b(X)\Big(2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|}\Big) = 0. \tag{26}$$

Now, the first addendum is easily reduced to

$$\sum_{\emptyset\subsetneq X\subseteq A^c} m_b(X)2^{-|X|}\Big(2^{|C^c|} - 2^{|A^c|}\Big),$$

while the second one becomes, after some manipulations,

$$\sum_{\emptyset\subsetneq X\subseteq A^c} m_b(X)2^{-|X|}\Big[\sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|}\Big(2^{|(Y\cup C)^c|} + -2^{|A^c|}\Big) - (2^{|C^c|} - 2^{|A^c|})\Big]. \tag{27}$$

But then we can notice that

$$\sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|}\Big(2^{|(Y\cup C)^c|} - 2^{|A^c|}\Big) = \sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|}2^{|(Y\cup C)^c|} - 2^{|A^c|}\sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|} = \sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|}2^{|(Y\cup C)^c|}$$

since $\sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|} = 0$ by Newton's binomial. As for the remaining term, using a common technique, we can decompose Y into the disjoint sum $Y = (Y \cap C) + (Y \setminus C)$ and rewrite it as

$$\sum_{|Y\cap C|=0}^{|C|}\binom{|C|}{|Y\cap C|}\sum_{|Y\setminus C|}^{|A\setminus C|}\binom{|A\setminus C|}{|Y\setminus C|}(-1)^{|Y\cap C|+|Y\setminus C|}\cdot 2^{n-|Y\setminus C|-|C|} = 2^{n-|C|}\sum_{|Y\cap C|=0}^{|C|}\binom{|C|}{|Y\cap C|}(-1)^{|Y\cap C|}\cdot \sum_{|Y\setminus C|}^{|A\setminus C|}\binom{|A\setminus C|}{|Y\setminus C|}(-1)^{|Y\setminus C|}2^{-|Y\setminus C|}$$

where

$$\sum_{|Y\setminus C|}^{|A\setminus C|}\binom{|A\setminus C|}{|Y\setminus C|}(-1)^{|Y\setminus C|}2^{-|Y\setminus C|} = \Big(-1+\frac{1}{2}\Big)^{|A\setminus C|} =$$

$= -2^{-|A\setminus C|}$ by Newton's binomial. Hence

$$\sum_{\emptyset\subseteq Y\subseteq A}(-1)^{|Y|}2^{|(Y\cup C)^c|} = -2^{|n-|C|-|A\setminus C|}\sum_{|Y\cap C|=0}^{|C|}\binom{|C|}{|Y\cap C|}(-1)^{|Y\cap C|} = 0,$$

again by Newton's binomial. By replacing this result in cascade into (27) and (26) the system of constraints reduces to $0 = 0$.

Proof of Lemma 1. The variable term in (14) can be decomposed into collections of terms which depend on the same individual variable $\gamma(X)$:

$$\max_{\emptyset\subsetneq X\subsetneq A}\max_{\emptyset\subseteq Y\subseteq A^c}\Big|\gamma(X) + \sum_{\emptyset\subsetneq Z\subseteq Y}\sum_{\emptyset\subseteq W\subseteq X} m_b(Z+W)\Big|$$

where $B = X + Y$, with $X = A \cap B$ and $Y = B \cap A^c$. Note that $Z \neq \emptyset$, as $C = Z + W \not\subset A$. Therefore the global optimal solution decomposes into a collection of solutions $\{\gamma^*(X), \emptyset \subsetneq X \subsetneq A\}$ for each individual problem, where $\gamma^*(X)$:

$$\max_{\emptyset \subseteq Y \subseteq A^c} \Big| \gamma^*(X) + \sum_{\emptyset \subsetneq Z \subseteq Y} \sum_{\emptyset \subseteq W \subseteq X} m_b(Z+W) \Big| \leq 1 - b(A). \tag{28}$$

We must distinguish three cases. **1.** If $\gamma(X) \geq 0$ we have that $\gamma^*(X)$:

$$\begin{aligned} &\max_{\emptyset \subseteq Y \subseteq A^c} \Big\{ \gamma^*(X) + \sum_{\emptyset \subsetneq Z \subseteq Y} \sum_{\emptyset \subseteq W \subseteq X} m_b(Z+W) \Big\} = \\ &= \gamma^*(X) + \sum_{\emptyset \subsetneq Z \subseteq A^c} \sum_{\emptyset \subseteq W \subseteq X} m_b(Z+W) = \\ &= \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \leq 1 - b(A) \end{aligned}$$

since when $\gamma(X) \geq 0$ the argument to maximize is non-negative, and its maximum is obviously achieved for $Y = A^c$. Hence:

$$\gamma^*(X) : \gamma^*(X) \leq 1 - b(A) - \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C). \tag{29}$$

2. If $\gamma(X) < 0$ the maximum in (28) can be achieved for either $Y = A^c$ or $Y = \emptyset$. Therefore we are left with the two corresponding terms in the max. We look then for $\gamma^*(X)$ such that

$$\max_{\emptyset \subseteq Y \subseteq A^c} \Big\{ \big| \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \big|, -\gamma^*(X) \Big\} \tag{30}$$

$\leq 1 - b(A)$. Now, either

$$|\gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C)| \geq -\gamma^*(X)$$

or viceversa. In the first case

$$\begin{aligned} &\Big| \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \Big| = \\ &= \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \geq -\gamma^*(X), \end{aligned}$$

as the argument of the absolute value has to be non-negative, i.e.,

$$\gamma^*(X) \geq -\frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C).$$

Furthermore, the optimality condition is met when

$$\gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \leq 1 - b(A),$$

equivalent to

$$\begin{aligned} &\gamma^*(X) \leq 1 - b(A) - \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) = \\ &= \sum_{C \cap A^c \neq \emptyset, C \cap (A \setminus X) \neq \emptyset} m_b(C) \geq 0 \end{aligned}$$

which is trivially true since $\gamma^*(X) < 0$. Therefore, the optimal interval is:

$$0 \geq \gamma^*(X) \geq -\frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C). \tag{31}$$

3. In the last case,

$$\Big| \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \Big| \leq -\gamma^*(X),$$

i.e., $\gamma^*(X) \leq -\frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C)$. Optimality is met for $-\gamma^*(X) \leq 1 - b(A)$ which is equivalent to $\gamma^*(X) \geq b(A) - 1$, which is in turn met for all

$$b(A) - 1 \leq \gamma^*(X) \leq -\frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C). \tag{32}$$

Equation (32) particularizes, in the ternary case $\Theta = \{x, y, z\}$, $A = \{x, y\}$, as

$$b(x,y) - 1 \leq \beta(x) \leq 1 - b(x,y) - \sum_{C \cap \{z\} \neq \emptyset, C \cap A \subseteq \{x\}} m_b(C).$$

This is indeed the solution (10) found in the ternary case. Putting (29), (31) and (32) together we have the thesis.

Proof of Theorem 3. We can prove it by substitution. By replacing (19) in (18) we get, since $\sum_{\emptyset \subsetneq C \subseteq X} (-1)^{|C|} = 0 - (-1)^0 = -1$ by Newton's binomial,

$$\begin{aligned} &\sum_{\emptyset \subsetneq C \subseteq X} \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \Big[(-1)^{|C|} m_b(B) - m_b(B \cup C) \Big] + \\ &+ \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) = \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} m_b(B) \cdot \\ &\cdot \Big(\sum_{\emptyset \subsetneq C \subseteq X} (-1)^{|C|} \Big) - \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subsetneq C \subseteq X} m_b(B \cup C) + \\ &+ \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) = -\frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} m_b(B) + \\ &- \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subsetneq C \subseteq X} m_b(B \cup C) + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} \\ &m_b(C) = -\frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subseteq C \subseteq X} m_b(B \cup C) + \\ &+ \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) = \\ &- \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \end{aligned}$$

$= 0$ for all $\emptyset \subsetneq X \subsetneq A$, and system (18) is met.

References

[1] Proceedings of BELIEF'10, Brest, France, 2010.

[2] P. Black, *Geometric structure of lower probabilities*, Random Sets: Theory and Applications (Goutsias, Malher, and Nguyen, eds.), Springer, 1997, pp. 361–383.

[3] A. Chateauneuf and J.Y. Jaffray, *Some characterization of lower probabilities and other monotone capacities through the use of mbius inversion*, Math. Soc. Sci. **17** (1989), 263–283.

[4] F. Cuzzolin, *Geometry of Dempster's rule of combination*, IEEE Tr. SMC-B **34:2** (2004), 961–977.

[5] F. Cuzzolin, *Two new Bayesian approximations of belief functions based on convex geometry*, IEEE Tr. SMC-B **37** (2007), no. 4, 993–1008.

[6] F. Cuzzolin, *A geometric approach to the theory of evidence*, IEEE Tr. SMC-C **38** (2008), no. 4, 522–534.

[7] F. Cuzzolin, *L_p consistent approximations of belief functions*, IEEE Tr. Fuzzy Systems (under review) (2008).

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

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

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

[11] Fabio Cuzzolin and Ruggero Frezza, *Geometric analysis of belief space and conditional subspaces*, Proceedings of ISIPTA2001, Cornell University, Ithaca, NY, June 2001.

[12] V.I. Danilov and G.A. Koshevoy, *Cores of cooperative games, superdifferentials of functions and the minkovski difference of sets*, Journal of Mathematical Analysis Applications **247** (2000), 1–14.

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

[14] D. Denneberg, *Conditioning (updating) non-additive probabilities*, Ann. Operations Res. **52** (1994), 21–42.

[15] R. Fagin and Joseph Y. Halpern, *A new approach to updating beliefs*, UAI, 6, 1991, pp. 347–374.

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

[17] I. Gilboa and D. Schmeidler, *Updating ambiguous beliefs*, Journal of economic theory **59** (1993), 33–49.

[18] V. Ha and P. Haddawy, *Geometric foundations for interval-based probabilities*, KR'98, San Francisco, California, 1998, pp. 582–593.

[19] M. Itoh and T. Inagaki, *A new conditioning rule for belief updating in the Dempster-Shafer theory of evidence*, Transactions of the Society of Instrument and Control Engineers **31:12** (1995), 2011–2017.

[20] J. Y. Jaffray, *Bayesian updating and belief functions*, IEEE Tr. SMC **22** (1992), 1144–1152.

[21] R.C. Jeffrey, *The logic of decision*, Mc Graw - Hill, 1965.

[22] F. Klawonn and Philippe Smets, *The dynamic of belief in the transferable belief model and specialization-generalization matrices*, UAI, 7, 1992, pp. 130–137.

[23] M.A. Klopotek and S.T. Wierzchon, *An interpretation for the conditional belief function in the theory of evidence*, LNCS, vol. 1609/1999, Springer, 1999, pp. 494–502.

[24] G.A. Koshevoy, *Distributive lattices and products of capacities*, Journal of Mathematical Analysis Applications **219** (1998), 427–441.

[25] H. E. Kyburg, *Bayesian and non-Bayesian evidential updating*, Artificial Intelligence **31** (1987), 271–293.

[26] E. Lehrer, *Updating non-additive probabilities - a geometric approach*, Games and Economic Behavior **50** (2005), 42–57.

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

[28] S. Maass, *A philosophical foundation of non-additive measure and probability*, Theory and decision **60** (2006), 175–191.

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

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

[31] G. Shafer, *Jeffrey's rule of conditioning*, Philosophy of Sciences **48** (1981), 337–362.

[32] A. Slobodova, *Multivalued extension of conditional belief functions*, Qualitative and quantitative practical reasoning, vol. 1244/1997, Springer Berlin/Heidelberg, 1997, pp. 568–573.

[33] A. Slobodova, *Conditional belief functions and valuation-based systems*, Tech. report, Slovak Academy of Sciences, Bratislava, SK, 1994.

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

[35] Ph. Smets, *Jeffrey's rule of conditioning generalized to belief functions*, Proc. of UAI'93, 1993, pp. 500–505.

[36] Ph. Smets, *The axiomatic justification of the transferable belief model*, Technical Report TR/IRIDIA/1995-8.1, Universite' Libre de Bruxelles, 1995.

[37] M. Spies, *Conditional events, conditioning, and random sets*, IEEE Tr. SMC **24** (1994), 1755–1763.

[38] P. Suppes and M. Zanotti, *On using random relations to generate upper and lower probabilities*, Synthese **36** (1977), 427–440.

[39] Y. Tang and J. Zheng, *Dempster conditioning and conditional independence in evidence theory*, AI 2005: Advance in Artificial Intelligence, vol. 3809/2005, Springer Berlin/Heidelberg, 2005, pp. 822–825.

[40] H. Xu and Philippe Smets, *Reasoning in evidential networks with conditional belief functions*, International Journal of Approximate Reasoning **14** (1996), 155–185.

[41] C. Yu and F. Arasta, *On conditional belief functions*, International Journal of Approxiomate Reasoning **10** (1994), 155–172.