Geometric conditioning of belief functions

Fabio Cuzzolin
Department of Computing
Oxford Brookes University
Oxford, United Kingdom
Email: (fabio.cuzzolin)@brookes.ac.uk

***Abstract*—In this paper we study the problem of conditioning a belief function b with respect to an event A by geometrically projecting such belief function onto the simplex associated with A in the simplex of all belief functions. Two different such simplices can be defined, as each belief function can be represented as the vector of its basic probability values or the vector of its belief values. We show here that defining geometric conditional b.f.s by minimizing L_p distances between b and the conditioning simplex in the mass space produces simple, elegant results with straightforward semantics in terms of degrees of belief. Such results can be interpreted in the light of a generalization to belief functions of the notion of imaging introduced by Lewis.**

Keywords: Belief functions, conditioning, geometric approach, L_p norms.

I. Introduction

Several theories of and approaches to conditioning in the framework of belief functions (b.f.s) [1], [2] have been proposed along the years [3]–[9]. In the original model, in which belief functions are induced by multi-valued mappings of probability distributions, Dempster's conditioning can indeed be judged inappropriate from a Bayesian point of view. Spies [10] defined conditional events as sets of equivalent events under conditioning. By applying multi-valued mapping to such events, conditional belief functions were introduced. An updating rule generalizing the total probability theorem was derived from them. Kyburg [11] analyzed the links between Dempster conditioning of belief functions and Bayesian conditioning of closed, convex sets of probabilities, of which belief functions are a special case. He arrived at the conclusion that the probability intervals generated by Dempster updating were included in those generated by Bayesian updating.

One way of dealing with such criticism is to abandon all notions of multivalued mapping to define belief directly in terms of basis belief assignments, as in Smets' transferable belief model [12]. The unnormalized conditional belief function $b_U(.|B)$ with b.b.a. $m_U(.|B)$[1]

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

is the minimal commitment specialization of b such that $pl_b(B^c|B) = 0$ [13]. In [14], Xu and Smets used conditional belief functions to represent relations between variables in evidential networks, and presented a propagation algorithm for such networks. In [15], Smets pointed out the distinction between revision and focussing in the conditional process, and how they lead to unnormalized and geometric [16] conditioning $b_G(A|B) = \frac{b(A\cap B)}{b(B)}$, respectively. In these two scenarios he proposed generalizations of Jeffrey's rule of conditioning [17], [18] $P(A|P',\mathbb{B}) = \sum_{B\in\mathbb{B}} \frac{P(A\cap B)}{P(B)} P'(B)$ to the framework of belief functions.

Slobodova also conducted some early studies on the issue of conditioning. In particular, a multi-valued extension of conditional b.f.s was introduced [19], and its properties examined. More recently, Tang and Zheng [20] also discussed the issue of conditioning in a multi-dimensional space. Klopotek and Wierzchon [21] provided a frequency-based interpretation for conditional belief functions.

Quite recently, Lehrer [22] proposed a geometric approach to determine the conditional expectation of non-additive probabilities. Such conditional expectation was then applied for updating, whenever information became available, and to introduce a notion of independence. Early attempts of studying conditioning in a geometric framework appeared in [23], where the simplicial geometry of the set $\langle b\rangle$ of all belief functions obtained by Dempster combination with a given b.f. b, or *conditional subspace*, was described.

A. Contribution

Along this line, in this paper we propose indeed to define the notion itself of conditioning by geometric methods. The idea is simple: as the collection of events $\{B \subseteq A\}$ included in a given conditioning event A determine a simplex in the space of belief functions, conditional belief functions can be defined geometrically by minimizing a certain distance between the original b.f. b and the conditioning simplex. Such geometric conditioning can take place in two different spaces $\mathcal{M}$ and $\mathcal{B}$, according to whether we represent belief functions are vectors of mass values or belief values. We show here that defining geometric conditional b.f.s by minimizing L_p distances between b and the conditioning simplex in the mass space $\mathcal{M}$ produces simple, elegant results with straightforward interpretations in terms of degrees of belief.

In summary, L_1-conditional belief functions in $\mathcal{M}$ form a polytope in which each vertex is the b.f. obtained by re-assigning the entire mass not contained in A to a *single* focal element $\{B \subseteq A\}$. In turn, the L_2 conditional b.f. is the barycenter of this polytope, i.e., the belief function with core in A obtained by re-assigning the mass $\sum_{B\not\subset A} m(B)$ to each focal element $\{B \subseteq A\}$ *on an equal basis*.

[1] Author's notation.

B. Interpretation

Such results can be interpreted as generalization of Lewis' *imaging* approach to belief revision originally formulated in the context of probabilities [24]. The idea behind imaging is that, upon observing that some state $x \in \Theta$ is impossible, you transfer the probability initially assigned to x completely towards the remaining state you deem the most similar to x [25]. Peter Gärdenfors [26] extends Lewis' idea by allowing a fraction λ_i of the probability of such state x to be re-distributed to all remaining states x_i ($\sum_i \lambda_i = 1$).

In the case of belief functions, the mass $m(C)$ of each focal element not included in A should be re-assigned to the "closest" focal element in $\{B \subseteq A\}$. If no information on the similarity between focal elements is available or make sense in a particular context, ignorance translates into allowing all possible set of weights $\lambda(B)$ for Gärdenfors' (generalized) belief revision by imaging. This yields the set of L_1 conditional b.f.s. If such ignorance is expressed by assigning instead equal weight $\lambda(B)$ to each $B \subseteq A$, the resulting revised b.f. is the unique L_2 conditional b.f., the barycenter of the L_1 polytope.

C. Paper outline

We first briefly recall in Section II the geometric approach to belief functions. In particular, we show how each b.f. corresponds to both the vector of its belief values in the belief space $\mathcal{B}$, and the vector of its mass values in the mass space $\mathcal{M}$. In this paper we pick the latter, and propose to measure distances there by L_p norms. Therefore, in Sections III, IV and V we prove the analytical forms of the L_1, L_2 and L_∞ conditional belief functions in $\mathcal{M}$, and discuss their interpretation in terms of degrees of belief.

We conclude in Section VI by discussing the obtained results, hinting to the case of geometric conditioning *in the belief space*, and providing an interpretation of geometric conditional b.f.s in the mass space in terms of the notion of imaging introduced by Lewis [24], generalized to belief functions.

II. Geometric conditional belief functions

A. Belief functions as vectors

As belief functions $b : 2^\Theta \to [0,1]$, $b(A) = \sum_{B \subseteq A} m_b(B)$, are set functions defined on a the power set 2^Θ of a finite space Θ, they are obviously completely defined by the associated 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 $\mathbb{R}^{N-2}, N = 2^{|\Theta|}$ [23]. The set $\mathcal{B}$ of points of $\mathbb{R}^{N-2}$ which correspond to a belief function is a simplex, namely: $\mathcal{B} = Cl(\vec{b}_A,\ \emptyset \subsetneq A \subseteq \Theta)$, where Cl denotes the convex closure operator and $\vec{b}_A$ is the vector associated with the categorical [12] belief function 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: $\vec{b} = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A) \vec{b}_A$.

In the same way, though, 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 $\vec{m}$ of its $N-1$ mass components:

$$\vec{m} = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B) \vec{m}_B. \tag{1}$$

Of course such vectors $\vec{m}$ will live in the subspace $\mathcal{M}$ of vectors whose components sum to 1.

B. Conditioning simplex and L_p norms

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

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

The set of such vectors form a simplex $\mathcal{M}_A \doteq Cl(\vec{m}_B,\ \emptyset \subsetneq B \subseteq A)$. The same is true in the belief space, where each belief function $\vec{a}$ assigning mass to focal elements included in A can be decomposed as: $\vec{a} = \sum_{\emptyset \subsetneq B \subseteq A} a(B) \vec{b}_B$. These vectors live in a simplex $\mathcal{B}_A \doteq Cl(\vec{b}_B,\ \emptyset \subsetneq B \subseteq A)$. We call $\mathcal{M}_A$ and $\mathcal{B}_A$ the *conditioning simplex* in the mass and the belief space, respectively.

Given a belief function b, we call *geometric conditional belief function induced by a distance function* d in $\mathcal{M}$ ($\mathcal{B}$) the b.f.(s) $b_d(.|A)$ which minimize(s) the distance $d(b, \mathcal{M}_A)$ ($d(b, \mathcal{B}_A)$) between b and the conditioning simplex in $\mathcal{M}$ ($\mathcal{B}$). In this paper we are mainly concerned with such geometric conditional b.f.s in the mass space $\mathcal{M}$.

We consider as distance functions the three major L_p norms: $d = L_1$ (Section III); $d = L_2$ (Section IV); $d = L_\infty$ (Section V). For vectors $\vec{m}$, $\vec{m'} \in \mathcal{M}$ representing the b.p.a.s of two belief functions b, b', such norms read as:

$$\begin{array}{lcl} \|\vec{m} - \vec{m'}\|_1 & \doteq & \sum_{\emptyset \subsetneq B \subseteq \Theta} |m(B) - m'(B)|; \\ \|\vec{m} - \vec{m'}\|_2 & \doteq & \sqrt{\sum_{\emptyset \subsetneq B \subseteq \Theta} (m(B) - m'(B))^2}; \\ \|\vec{m} - \vec{m'}\|_\infty & \doteq & \max_{\emptyset \subsetneq B \subseteq \Theta} |m(B) - m'(B)|. \end{array} \tag{3}$$

III. Conditional belief functions by L_1 norm

Given any belief function b with basic probability assignment m_b collected in a vector $\vec{m}_b \in \mathcal{M}$, the set of L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ is the set of b.f.s whose basic probability assignments $m_{L_1,\mathcal{M}}(.|A)$ satisfy:

$$\vec{m}_{L_1,\mathcal{M}}(.|A) \doteq \arg\min_{\vec{a} \in \mathcal{M}_A} \|\vec{m}_b - \vec{a}\|_1. \tag{4}$$

Using the expression (3) of the L_1 norm in the mass space $\mathcal{M}$, (4) can be written as:

$$\arg\min_{\vec{a} \in \mathcal{M}_A} \|\vec{m}_b - \vec{a}\|_1 = \arg\min_{\vec{a} \in \mathcal{M}_A} \sum_{\emptyset \subsetneq B \subseteq \Theta} |m_b(B) - a(B)|.$$

A. A change of variables

By exploiting the fact that the candidate solution $\vec{a}$ is an element of $\mathcal{M}_A$ (Equation (2)) we can greatly simplify this expression. Namely,

$$\begin{aligned}\vec{m}_b - \vec{a} &= \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B)\vec{m}_B - \sum_{\emptyset \subsetneq B \subseteq A} a(B)\vec{m}_B \\ &= \sum_{\emptyset \subsetneq B \subseteq A} (m_b(B) - a(B))\vec{m}_B + \sum_{B \not\subset A} m_b(B)\vec{m}_B.\end{aligned}$$

The following change of variables

$$\beta(B) \doteq m_b(B) - a(B) \tag{5}$$

further yields:

$$\vec{m}_b - \vec{a} = \sum_{\emptyset \subsetneq B \subseteq A} \beta(B)\vec{m}_B + \sum_{B \not\subset A} m_b(B)\vec{m}_B. \tag{6}$$

We need to observe, though, that the variables $\{\beta(B), \emptyset \subsetneq B \subseteq A\}$ are not independent. Indeed,

$$\sum_{\emptyset \subsetneq B \subseteq A} \beta(B) = \sum_{\emptyset \subsetneq B \subseteq A} m_b(B) - \sum_{\emptyset \subsetneq B \subseteq A} a(B) = b(A) - 1$$

as $\sum_{\emptyset \subsetneq B \subseteq A} a(B) = 1$ by definition, since $\vec{a} \in \mathcal{M}_A$. Therefore in the optimization problem there are just $2^{|A|} - 2$ independent variables, while: $\beta(A) = b(A) - 1 - \sum_{\emptyset \subsetneq B \subsetneq A} \beta(B)$. By replacing the above equality into (6) we finally get:

$$\begin{aligned}&\vec{m}_b - \vec{a} = \sum_{\emptyset \subsetneq B \subsetneq A} \beta(B)\vec{m}_B + \\ &+\Big(b(A) - 1 - \sum_{\emptyset \subsetneq B \subsetneq A} \beta(B)\Big)\vec{m}_A + \sum_{B \not\subset A} m_b(B)\vec{m}_B.\end{aligned} \tag{7}$$

B. L_1-conditional belief functions and dominating masses

In the L_1 case we get then

$$\|\vec{m}_b - \vec{a}\|_1 = \sum_{\emptyset \subsetneq B \subsetneq A} |\beta(B)| + \Big|b(A) - 1 - \sum_{\emptyset \subsetneq B \subsetneq A} \beta(B)\Big|, \tag{8}$$

plus the constant $\sum_{B \not\subset A} |m_b(B)|$ which does not depend on β. This is a function of the form

$$\sum_i |x_i| + \Big| - \sum_i x_i - k\Big|, \quad k \geq 0 \tag{9}$$

which has an entire simplex of minima, namely: $x_i \leq 0 \ \forall i$, $\sum_i x_i \geq -k$. See Figure 1 for the case of two variables, x_1 and x_2 (corresponding to the L_1 conditioning problem on an event A of size $|A| = 2$). The minima of the L_1 norm (8) are therefore given by the following system of constraints:

$$\begin{cases} \beta(B) \leq 0 & \forall \emptyset \subsetneq B \subsetneq A, \\ \sum_{\emptyset \subsetneq B \subsetneq A} \beta(B) \geq b(A) - 1. \end{cases} \tag{10}$$

In the original simplicial coordinates $\{a(B), \emptyset \subsetneq B \subseteq A\}$ of the candidate solution $\vec{a}$ in $\mathcal{M}_A$ this reads as:

$$\begin{cases} m_b(B) - a(B) \leq 0 & \forall \emptyset \subsetneq B \subsetneq A, \\ \sum_{\emptyset \subsetneq B \subsetneq A} (m_b(B) - a(B)) \geq b(A) - 1 \end{cases}$$

i.e., $a(B) \geq m_b(B) \ \forall \emptyset \subsetneq B \subseteq A$.

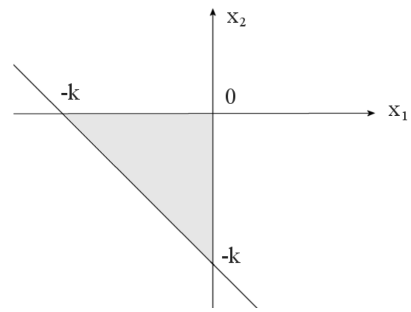

Figure 1. The minima of a function of the form (9) with two variables x_1, x_2 form the triangle $x_1 \leq 0$, $x_2 \leq 0$, $x_1 + x_2 \geq -k$ depicted here.

Recall that the *core* $\mathcal{C}_b$ of a b.f. b is the union of its focal elements $B : m_b(B) \neq 0$.

Theorem 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:

$$b_{L_1,\mathcal{M}}(.|A) = \{b' : \mathcal{C}_{b'} \subseteq A, m_{b'}(B) \geq m_b(B) \ \forall \emptyset \subsetneq B \subseteq A\}. \tag{11}$$

C. Simplex of L_1-conditional belief functions

As we may observe in Figure 1, the set of L_1 conditional belief functions in $\mathcal{M}$ has geometrically the form of a simplex. It is easy to see that by Equation (10) the $2^{|A|} - 2$ vertices of such simplex are associated with the following solutions:

$$\begin{cases} \beta(X) &= 0 \quad \forall \emptyset \subsetneq X \subsetneq A; \\ \beta(B) &= b(A) - 1, \ \beta(X) = 0 \ \forall \emptyset \subsetneq X \subsetneq A, X \neq B \\ & \quad \forall \emptyset \subsetneq B \subsetneq A. \end{cases}$$

Such solutions read in the $\{a(B)\}$ coordinates as the vectors $\vec{m}[b]|^B_{L_1} A = \vec{a} \in \mathcal{M}_A$ such that:

$$\begin{array}{l} a(B) = m_b(B) + 1 - b(A), \\ a(X) = m_b(X) \quad \forall \emptyset \subsetneq X \subsetneq A, X \neq B \end{array} \tag{12}$$

for all $\emptyset \subsetneq B \subseteq A$ (A included).

Theorem 2: Given a b.f. $b : 2^\Theta \rightarrow [0, 1]$ 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 simplex $\mathcal{M}_{L_1,A}[b] = Cl(\vec{m}[b]|^B_{L_1} A)$ with vertices (12).

It is important to notice that all the vertices of the L_1 conditional simplex fall inside $\mathcal{M}_A$ proper. In principle, some of them could have fallen in the linear space generated by $\mathcal{M}_A$ but outside the simplex $\mathcal{M}_A$, i.e., some of the solutions $a(B)$ could have been negative. This is not the case, somehow supporting the validity of such an approach to conditioning based on geometric projections onto the appropriate simplices.

IV. Conditional belief functions by L_2 norm

We can proceed to find the L_2 conditional belief function(s) by using again the form (6) of the difference vector $\vec{m}_b - \vec{a}$, where again $\vec{a}$ is an arbitrary vector of the conditional simplex $\mathcal{M}_A$. In this case 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:

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

whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$ (the vector space generated by the vectors $\{\vec{g}_i, i\}$).

In our case $\vec{p} = \vec{m}_b$ is the original mass function, $\vec{q} = \vec{a}$ is an arbitrary point in $\mathcal{M}_A$, while the generators of $\mathcal{M}_A$ are all the vectors $\vec{g}_B = \vec{m}_B - \vec{m}_A$, $\forall \emptyset \subsetneq B \subsetneq A$. Such generators are vectors of the form

$$[0, \cdots, 0, 1, 0, \cdots, 0, -1, 0, \cdots, 0]'$$

with all zero entries but entry B (equal to 1) and entry A (equal to -1). Making use of Equation (6), the condition $\langle \vec{m}_b - \vec{a}, \vec{m}_B - \vec{m}_A \rangle = 0$ assumes then a very simple form

$$\beta(B) - b(A) + 1 + \sum_{\emptyset\subsetneq X\subsetneq A, X\neq B} \beta(X) = 0$$

for all possible generators of $\mathcal{M}_A$, i.e.:

$$2\beta(B) + \sum_{\emptyset\subsetneq X\subsetneq A, X\neq B} \beta(X) = b(A) - 1 \quad \forall \emptyset \subsetneq B \subsetneq A. \quad (13)$$

A. The unique solution of the L_2 problem

The system (13) is a linear system of $2^{|A|} - 2$ equations in $2^{|A|} - 2$ variables (the $\beta(X)$). It can therefore be written as

$$\mathcal{A}\vec{\beta} = (b(A) - 1)\vec{1},$$

where $\vec{1}$ is the vector of the appropriate size with all entries at 1. Its unique solution is trivially $\vec{\beta} = (b(A) - 1) \cdot \mathcal{A}^{-1}\vec{1}$. The matrix $\mathcal{A}$ is of the form

$$\mathcal{A} = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ & & \cdots & \\ 1 & 1 & \cdots & 2 \end{bmatrix}.$$

Simple linear algebra proves that the inverse of such matrix has the form

$$\mathcal{A}^{-1} = \frac{1}{d+1} \begin{bmatrix} d & -1 & \cdots & -1 \\ -1 & d & \cdots & -1 \\ & & \cdots & \\ -1 & -1 & \cdots & d \end{bmatrix}$$

where d is the number of rows (or columns) of $\mathcal{A}$. It is easy to see that $\mathcal{A}^{-1}\vec{1} = \frac{1}{d+1}\vec{1}$, where in our case $d = 2^{|A|} - 2$. The solution of the system (13) is then given by

$$\vec{\beta} = \mathcal{A}^{-1}\vec{1}(b(A) - 1) = \frac{1}{2^{|A|} - 1}\vec{1}(b(A) - 1),$$

i.e., $\arg\min_{\vec{\beta}} \|\vec{m}_b - \vec{a}\|_2$ is such that

$$\beta(B) = \frac{b(A) - 1}{2^{|A|} - 1} \quad \forall \emptyset \subsetneq B \subsetneq A. \quad (14)$$

In the $\{a(B)\}$ coordinates the solution reads as

$$a(B) = m_b(B) + \frac{1 - b(A)}{2^{|A|} - 1} \quad \forall \emptyset \subsetneq B \subseteq A, \quad (15)$$

A included.

B. L_2 conditional b.f. as barycenter of dominating masses

According to Equation (15) then, the L_2 conditional belief function is unique, and corresponds to the mass function which *redistributes in an equal, even way the mass the original belief function assign to focal elements not included in A to each and all the subsets of A.*

Theorem 3: Given a belief function $b : 2^\Theta \to [0, 1]$ 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:

$$m_{L_2,\mathcal{M}}(B|A) = m_b(B) + \frac{1}{2^{|A|} - 1} \sum_{B\not\subset A} m_b(B) \;\forall \emptyset \subsetneq B \subseteq A.$$

Besides, L_2 and L_1 conditional belief functions in $\mathcal{M}$ display a strong relationship, as:

Theorem 4: Given a belief function $b : 2^\Theta \to [0, 1]$ and an arbitrary non-empty focal element $\emptyset \subsetneq A \subseteq \Theta$, the L_2 conditional belief function $b_{L_2,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is the center of mass of the simplex $\mathcal{M}_{L_1,A}[b]$ of L_1 conditional belief functions with respect to A in $\mathcal{M}$.

Proof: By definition the center of mass of $\mathcal{M}_{L_1,A}[b]$, whose vertices are given by (12), is the vector

$$\frac{1}{2^{|A|} - 1} \sum_{\emptyset\subsetneq B\subseteq A} \vec{m}[b]|^B_{L_1} A$$

whose entry B is given by

$$\frac{1}{2^{|A|} - 1}\Big[m_b(B)(2^{|A|} - 1) + (1 - b(A))\Big]$$

i.e., (15). ■

V. Conditional belief functions by L_∞ norm

Similarly, we can use Equation (7) to minimize the L_∞ distance between the original mass vector $\vec{m}_b$ and the conditioning subspace $\mathcal{M}_A$. Let us rewrite it for sake of readability:

$$\begin{aligned} \vec{m}_b - \vec{a} \;\; = & \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)\vec{m}_B + \sum_{B\not\subset A} m_b(B)\vec{m}_B \\ & + \Bigg(b(A) - 1 - \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)\Bigg)\vec{m}_A. \end{aligned}$$

Its L_∞ norm reads as:

$$\begin{aligned} \|\vec{m}_b - \vec{a}\|_\infty \;\; = & \max\Big\{|\beta(B)| \;\emptyset \subsetneq B \subsetneq A, \\ & |m_b(B)| \; B \not\subset A, \Big|b(A) - 1 - \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)\Big|\Big\}. \end{aligned}$$

As $|b(A) - 1 - \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)| = | - \sum_{B\not\subset A} m_b(B) - \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)| = |\sum_{B\not\subset A} m_b(B) + \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)|$, the above norms simplifies as:

$$\max\Big\{ \;\; |\beta(B)| \;\emptyset \subsetneq B \subsetneq A, \max_{B\not\subset A}\{m_b(B)\}, \Big|\sum_{B\not\subset A} m_b(B) + \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)\Big|\Big\}. \quad (16)$$

This is a function of the form

$$f(x_1, ..., x_n) = \max\left\{|x_i|\ \forall i, |\sum_i x_i + k_1|, k_2\right\} = \max\left\{g(x_1, ..., x_n), k_2\right\}$$

where $g(x_1, ..., x_n) = \max\left\{|x_i|\ \ \forall i, |\sum_i x_i + k_1|\right\}$, with $k_1, k_2 \geq 0$, $k_2 \leq k_1$. The level sets $g = const$ of a function of the form g is represented (in the case in which there are two variables, i.e., when $A = 2$) in Figure 2.

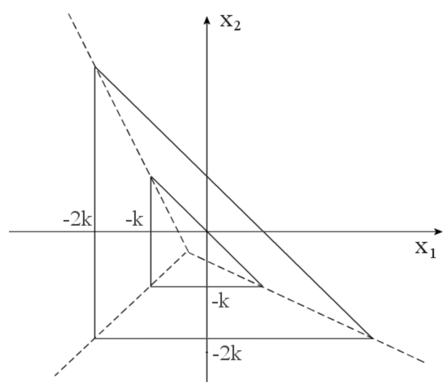

Figure 2. The level sets $g(x_1, x_2) = const$ of functions of the form $g(x_1, x_2) = \max\{|x_1|, |x_2|, |x_1 + x_2 + k_1|\}$, $k_1 \geq 0$ are triangles with vertices $[-x, -x]'$, $[-x, +x]'$, $[x, -x]'$. The minimum of such a function g lies in $[x_1, x_2]' = [-k_1/3, -k_1/3]'$.

The minimum of such a function lies in $\frac{1}{n+1}[-k_1, ..., -k_1]'$, where n is the number of variables. The function there assumes value $\min g = \frac{k_1}{n+1}$. Its level sets $g(x_1, ..., x_n) = const$ have the form of simplices of the kind depicted in Figure 2 for the case of 2 variables.
Now, what is the minimum of $f(x_1, ..., x_n)$? Clearly, if $k_2 \leq \min g = \frac{k_1}{n+1}$ then f has also minimum in $\frac{1}{n+1}[-k_1, ..., -k_1]'$.
Immediately, since (16) is a function of the form g in $n = 2^{|A|} - 2$ variables (the $\beta(B)$) we obtain the following result.

Lemma 1: If $\max_{B \not\subset A}\{m_b(B)\} \leq \frac{1-b(A)}{2^{|A|}-1}$ then $\arg\min_{\vec{\beta}} \|\vec{m}_b - \vec{a}\|_\infty$ is such that

$$\beta(B) = \frac{b(A) - 1}{2^{|A|} - 1} \quad \forall \emptyset \subsetneq B \subsetneq A. \qquad (17)$$

As an immediate consequence:

Theorem 5: If

$$\max_{B \not\subset A}\{m_b(B)\} \leq \frac{\sum_{B \not\subset A} m_b(B)}{2^{|A|} - 1} \qquad (18)$$

then the L_∞ conditional belief function $b_{L_\infty, \mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is unique, and corresponds to the L_2 conditional belief function (15), and the barycenter of the polytope of L_1 conditional belief functions.

Proof: It suffices to notice that (17) coincides with (14). ∎

In the case in which condition (18) does not hold, we get again an entire simplex of solutions, whose epistemic interpretation is difficult to understand. We will work on this point in the near future.

VI. Discussion and perspectives

A. Features of geometric conditional belief functions

From the analysis of geometric conditioning in the space of mass functions $\mathcal{M}$ a number of facts arise:

- geometric conditional b.f.s, even though obtained by minimizing purely geometric distances, possess very simple and elegant interpretations in terms of degrees of belief;
- while some of them correspond to pointwise conditioning, some others form whole polytopes of solutions whose vertices also have simple interpretations;
- conditional belief functions associated with the major L_1, L_2 and L_∞ norms are strictly related to each other;
- they are all characterized by the fact that, in the way they re-assign mass from focal elements $B \not\subset A$ not in A to focal elements in A, they do not distinguish between subsets which have non-empty intersection with A and those which have not.

The last point is quite interesting: mass geometric conditional b.f.s do not seem to care about the contribution focal elements make *to the plausibility* of the conditioning event A, but only to whether they contribute or not to the *degree of belief* of A. The reason is, roughly speaking, that in mass vectors $\vec{m}$ the mass of a given focal element appears only in the corresponding entry of $\vec{m}$. In opposition, belief vectors $\vec{b}$ are such that each entry $\vec{b}(B) = \sum_{X \subseteq B} m_b(X)$ of theirs contains information about the mass of all the subsets of B. As a result, it is to be expected that geometric conditioning *in the belief space* $\mathcal{B}$ will see the mass redistribution process function in a manner linked to the contribution of each focal element to the plausibility of the conditioning event A. This interesting matter is also left for future analysis.

B. Interpretation as general imaging for belief functions

The form of geometric conditional belief functions in the mass space can be naturally interpreted in the framework of an interesting approach to belief revision, known as *imaging* [25]. We will illustrate this notion and how it related to our results using the example proposed in [25].

Suppose you briefly glimpse at a transparent urn filled with black or white balls, and are asked to assign a probability value to the possible "configurations" of the urn. Suppose you are given three options: 30 black balls and 30 white balls (state a); 30 black balls and 20 white balls (state b); 20 black balls and 20 white balls (state c). Hence, $\Theta = \{a, b, c\}$. Since the observation only gave you the vague impression of having seen approximately the same number of black and white balls, you deem the states a and c equally likely, but at the same time you deem the event "a or c" twice as likely as the state b. Hence, you assign probability 1/3 to each of the states. Now, you are told that state c is false. How do you revise the probabilities of the two remaining states a and b?

Lewis [24] argued that, upon observing that a certain state $x \in \Theta$ is impossible, you should transfer the probability originally allocated to x to the remaining state which you deem the "most similar" to x. In this case, a is the state most similar to c, as they both consider an equal number of black and white balls. You obtain $(2/3, 1/3)$ as probability values of a and b, respectively. Peter Gärdenfors further extended Lewis' idea (*general imaging*) by allowing to transfer a part λ of the probability 1/3, initially assigned to c, towards state a, and the remaining part $1-\lambda$ to state b. These fractions should be independent of the initial probabilistic state of belief.

Now, what happens when our state of belief is described by a belief function, and we are told that A is true? In the general imaging framework we need to re-assign the mass $m(C)$ of each focal element not included in A to all the focal elements $B \subseteq A$, according to some weights $\{\lambda(B), B \subseteq A\}$.
Suppose there is no reason to attribute larger weights to any focal element in A, as, for instance, we have no meaningful similarity measure (in the given context for the given problem) between the states described by two different focal elements. We can then proceed in two different ways.
One option is to represent our complete ignorance about the similarities between C and each $B \subseteq A$ as a vacuous belief function on the set of weights. If applied to all focal elements C not included in A, this results in an entire polytope of revised belief functions, each associated with an arbitrary normalized weighting. It is trivial to see that this coincides with the set L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ of Theorem 2.
On the other hand, we can represent the same ignorance as a uniform probability distribution on the set of weights $\{\lambda(B), B \subseteq A\}$, for all $C \not\subseteq A$. Again, it is easy to see that general imaging produces in this case a single revised b.f., the L_2 conditional belief functions $b_{L_2,\mathcal{M}}(.|A)$ of Theorem 3.

As a final remark, the "information order independence" axiom of belief revision states that the revised belief should not depend on the order in which the information is made available. In our case, the revised (conditional) b.f.s obtained by observing first an event A and later another event A' should be the same as the ones obtained by revising first with respect to A' and then A. It is not difficult to see that both the L_1 and L_2 geometric conditioning operators presented here meet such axiom, supporting the case for their rationality.

VII. Conclusions

In this paper we showed how the notion of conditional belief function $b(.|A)$ can be introduced by geometric ways, by projecting any belief function onto the the simplicial subspace associated with the event A. The result depends on the choice of the vectorial representation for b, and of the distance function to minimize. We thoroughly analyzed the case of conditioning a b.b.a. vector by means of the norms L_1, L_2, and L_∞, and showed that such results have simple interpretations in terms of degrees of belief. Interpretations in terms of general imaging of these results were also given. A complete analysis of geometric conditioning in the belief space is the next obvious step. A full understanding of how geometric conditional b.f.s relate to classical approaches to conditioning is also needed, and will be pursued in the future.

References

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

[2] A. P. Dempster, "Upper and lower probabilities induced by a multivariate mapping," *Annals of Mathematical Statistics*, vol. 38, pp. 325–339, 1967.

[3] A. Chateauneuf and J. Jaffray, "Some characterization of lower probabilities and other monotone capacities through the use of möbius inversion," *Math. Soc. Sci.*, vol. 17, pp. 263–283, 1989.

[4] R. Fagin and J. Y. Halpern, "A new approach to updating beliefs," in *Uncertainty in Artificial Intelligence, 6*, L. K. P.P. Bonissone, M. Henrion and J. Lemmer, Eds., 1991, pp. 347–374.

[5] J. Y. Jaffray, "Bayesian updating and belief functions," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, pp. 1144–1152, 1992.

[6] I. Gilboa and D. Schmeidler, "Updating ambiguous beliefs," *Journal of economic theory*, vol. 59, pp. 33–49, 1993.

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

[8] C. Yu and F. Arasta, "On conditional belief functions," *International Journal of Approxiomate Reasoning*, vol. 10, pp. 155–172, 1994.

[9] 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*, vol. 31:12, pp. 2011–2017, 1995.

[10] M. Spies, "Conditional events, conditioning, and random sets," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 24, pp. 1755–1763, 1994.

[11] H. E. Kyburg, "Bayesian and non-Bayesian evidential updating," *Artificial Intelligence*, vol. 31, pp. 271–293, 1987.

[12] P. Smets, "Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of Approximate Reasoning*, vol. 9, pp. 1–35, 1993.

[13] F. Klawonn and P. Smets, "The dynamic of belief in the transferable belief model and specialization-generalization matrices," in *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, D. B. Dubois D., Wellman M.P. and S. Ph., Eds., 1992, pp. 130–137.

[14] H. Xu and P. Smets, "Reasoning in evidential networks with conditional belief functions," *International Journal of Approximate Reasoning*, vol. 14, pp. 155–185, 1996.

[15] P. Smets, "Jeffrey's rule of conditioning generalized to belief functions," in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence (UAI93)*, M. A. Heckerman D., Ed., 1993, pp. 500–505.

[16] P. Suppes and M. Zanotti, "On using random relations to generate upper and lower probabilities," *Synthese*, vol. 36, pp. 427–440, 1977.

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

[18] G. Shafer, "Jeffrey's rule of conditioning," *Philosophy of Sciences*, vol. 48, pp. 337–362, 1981.

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

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

[21] M. Klopotek and S. Wierzchon, "An interpretation for the conditional belief function in the theory of evidence," in *Foundations of intelligent systems - Lecture Notes in Computer Science*. Springer Berlin/Heidelberg, 1999, vol. 1609/1999, pp. 494–502.

[22] E. Lehrer, "Updating non-additive probabilities - a geometric approach," *Games and Economic Behavior*, vol. 50, pp. 42–57, 2005.

[23] F. Cuzzolin, "A geometric approach to the theory of evidence," *IEEE Transactions on Systems, Man and Cybernetics part C*, vol. 38, no. 4, pp. 522–534, 2008.

[24] D. Lewis, "Probabilities of conditionals and conditional probabilities," *Philosophical Review*, vol. 85, pp. 297–315, 1976.

[25] A. Perea, "A model of minimal probabilistic belief revision," *Theory and Decision*, vol. 67, no. 2, pp. 163–222, August 2009. [Online]. Available: http://ideas.repec.org/a/kap/theord/v67y2009i2p163-222.html

[26] P. Gardenfors, *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, MA, 1988.