Geometric conditioning in belief calculus

Fabio Cuzzolin

Department of Computing, Oxford Brookes University
Wheatley campus, Oxford OX33 1HX, United Kingdom

Abstract

Conditioning is crucial in applied science when inference involving time series is involved. Belief calculus is an effective way of handling such inference in the presence of uncertainty, but different approaches to conditioning in that framework have been proposed in the past, leaving the matter unsettled. We propose here an approach to the conditioning of belief functions based on geometrically projecting them onto the simplex associated with the conditioning event in the space of all belief functions. Two different such simplices can be defined, as each belief function can be represented as either the vector of its basic probability values or the vector of its belief values. We show here that such a geometric approach to conditioning often produces simple results with straightforward interpretations in terms of degrees of belief. The question of whether classical approaches, such as for instance Dempster's conditioning, can also be reduced to some form of distance minimization remains open: the study of families of combination rules generated by (geometric) conditioning rules appears to be the natural prosecution of the presented research.

Keywords: theory of evidence, belief function, conditioning, geometric approach, L_p norms

1. Introduction

Decision making and estimation are common problems in applied science, as people or machines need to make inferences about the state of the external world, and take appropriate actions. Such state is typically assumed to be described by a probability distribution over a set of alternative hypotheses, which in turn needs to be inferred from the available data. Sometimes, however, as in the case of extremely rare events (e.g., a volcanic eruption), few statistics are available to drive such inference. Part of the data can be missing. In addition, under the law of large numbers, probability distributions are the outcome of an infinite process of evidence accumulation, while in all practical cases the available evidence can only provide some sort of constraint on the unknown, "true" probability governing the process.

Different kinds of constraints are associated with different generalizations of probabilities, formulated to model uncertainty at the level of probability distributions. The simplest such generalizations are, possibly, interval probabilities [1] and convex sets of probabilities or "credal sets" [2]. A whole battery of different uncertainty theories [3], however, have been developed in the last century or so, starting from De Finetti's pioneering work [4]. In particular, Shafer's theory of belief functions (b.f.s) [5], based on A. Dempster's [6] seminal work, allows us to express partial belief by providing lower and upper bounds to probability values. The widespread influence of uncertainty at different levels explains why belief functions have been increasingly applied to fields as diverse as robotics, economics, or machine vision. Powerful tools for decision making with b.f.s have also been proposed [7, 8, 9].

When observations come from time series, however, or when conditional independence assumptions are necessary to simplify the structure of the joint (belief) distribution to estimate, the need for generalizing the classical results on total probability to belief calculus arises. This is the case, for instance, in image segmentation [10], where conditional independence is crucial to make optimization problems tractable. In target tracking, conditional constraints on the targets' future positions given their past locations are available. If such constraints are described as belief functions, predicting current target locations requires combining conditional b.f.s in a total function [11, 12].

1.1. Conditioning in belief calculus

Different definitions of conditional belief functions have been proposed in the past. The original proposal is due to Dempster himself [6]. He formulated it in his original model, in which belief functions are induced by multi-valued mappings $\Gamma : \Omega \rightarrow 2^{\Theta}$ 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 belief calculus have been proposed [13, 14, 15, 16, 17, 18, 19, 20, 21] along the years, 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 [22] 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 [23] a multi-valued extension of conditional b.f.s was introduced [24], and its properties examined. Klopotek and Wierzchon [25] provided instead a frequency-based interpretation for conditional belief functions.

Another way of dealing with the classical Bayesian criticism of Dempster's rule is to abandon all notions of multivalued mapping, and define belief directly on the power set of the frame as in Smets' Transferable Belief Model [26], as a sum function $b : 2^\Theta \rightarrow [0,1]$ of the form $b(A) = \sum_{B \subseteq A} m_b(B)$, induced by a "basic belief assignment" (b.b.a) $m_b : 2^\Theta \rightarrow [0,1]$. In particular, the conditional b.f. $b_U(B|A)$ with b.b.a. $m_U(B|A) = \sum_{C \subseteq A^c} m(B \cup C)$, $B \subseteq A$ is the minimal commitment specialization of b such that the plausibility of the complementary event A^c is nil [27]. In [28], Smets pointed out the distinction between "revision" and "focussing" in the conditional process, and the way they lead to unnormalized and geometric [29] conditioning, respectively. In these two scenarios he proposed some generalizations of Jeffrey's rule of conditioning [30, 31] to belief calculus.

1.2. A geometric approach to conditioning

Quite recently, the idea of formulating the problem geometrically has emerged. Lehrer [32], in particular, has proposed such a geometric approach to determine the conditional expectation of non-additive probabilities (such as belief functions). The notion of generating conditional b.f.s by minimizing a suitable distance function between the original b.f. b and the "conditioning region" $\mathcal{B}_A$ associated with the conditioning event A, i.e., the set of belief functions whose b.b.a. assigns mass to subsets of A only

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

has a clear potential. It expands our arsenal of 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 indeed been studied by different authors [33, 34, 35]. A similar approach has been developed and applied by the author more specifically to the theory of evidence [36, 37, 38]. Most relevantly, Jousselme et al [39] have conducted a very interesting survey of all the distances and similarity measures so far introduced in belief calculus, and proposed a number of generalizations. Other similarity measures between belief functions have been proposed by Shi et al [40], Jiang et al [41], and others [42, 43, 41]. Many of these measures could be in principle plugged in the above minimization problem (1) to define conditional belief functions. In [38] the author has computed some conditional belief functions generated via minimization of L_p norms in the "mass space", where b.f.s are represented by the vectors of their basic probabilities.

1.3. Contribution

In this paper we explore the geometric conditioning approach in both the mass space $\mathcal{M}$ and 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.

Mass space. 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* subset of "focal element" $\{B \subseteq A\}$. In turn, the L_2 conditional b.f. is the barycenter of this polytope, i.e., the belief function obtained by re-assigning the mass $\sum_{B \not\subset A} m(B)$ to each focal element $\{B \subseteq A\}$ *on an equal basis*. Such results can be interpreted as a generalization of Lewis' *imaging* approach to belief revision, originally formulated in the context of probabilities [44]. 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 [45]. Peter Gärdenfors [46] extended 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.
Belief space. Conditional belief functions in the belief space seem to have rather less straightforward interpretations than the corresponding quantities in the mass space. The barycenter of the set of L_∞ conditional belief functions can be interpreted 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*. 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). As a result, the obtained mass function is not necessarily non-negative: again, such version of geometrical conditioning may generated pseudo belief functions. Furthermore, while being quite similar to it, the L_2 conditional belief function in $\mathcal{B}$ is distinct from the barycenter of the L_∞ conditional b.f.s.
In the L_1 case, not only the resulting conditional pseudo belief functions are not guaranteed to be proper belief functions, but it appears difficult to find simple interpretations for these results in terms of degrees of belief.
A number of interesting cross relations between conditional b.f.s of the two representation domains appear to exist from an empirical comparison, and remain to be investigated further.

1.4. Paper outline

We commence by illustrating in Section 2 the crucial role of conditioning in a real world scenario drawn from an important computer vision application, data association, and an approach to this problem based on belief calculus. We then move on to recalling, in Section 3, the geometric approach to belief functions. In particular, we show how each b.f. b can be represented by either the vector $\vec{b}$ of its belief values in the belief space $\mathcal{B}$ or the vector $\vec{m}$ of its mass values in the mass space $\mathcal{M}$. In Section 4 we pick the latter representation. In Sections 4.1, 4.2 and 4.3 we prove the analytical forms of the L_1, L_2 and L_∞ conditional belief functions in $\mathcal{M}$, respectively. We discuss their interpretation in terms of degrees of belief (Section 4.5), and hint to an interesting link with Lewis' imaging [44], generalized to b.f.s (Section 4.6). Section 5 is dedicated to the derivation of geometric conditional belief functions in the belief space. In Section 5.1 we prove the analytical form of L_2 conditional b.f.s in $\mathcal{B}$, and propose a preliminary interpretation for them. We do the same in Sections 5.2 and 5.3 for L_1 and L_∞ conditional belief functions in the belief space, respectively. We conclude the paper with a critical discussion of the obtained results. In Section 6 a comparison of conditioning in mass and belief space is run, also with the help of the ternary case study. Finally, in Section 7 we prospect a number of future developments for the geometric approach to conditioning, and the application of similar minimization techniques to the consonant and consistent approximation problems.

2. Conditioning in belief calculus: a concrete scenario

2.1. Model-based data association

The *data association* problem is one of the more intensively studied computer vision applications for its important role in the implementation of automated defense systems, and its connections to the classical field of *structure from motion*, i.e. the reconstruction of a rigid scene from a sequence of images.

A number of feature points moving in the 3D space are followed by one or more cameras and appear in an image sequence as "unlabeled" points (i.e. we do not know the correspondences between points appearing in two consecutive frames). A typical example consists of a set of markers set at fixed positions on a moving articulated body: in order to reconstruct the trajectory of the cloud of "targets" (or of the underlying body) we need to associate feature points belonging to pairs of consecutive images, I_k and I_{k+1}.

A classical approach to the data association problem (the *joint probabilistic data association* filter, [47]) is based on tuning a number of Kalman filters (each associated with a single feature point), whose aim is to predict the future position of the target, in order to produce the most probable labeling of the cloud of points in the next image.

Unfortunately, the JPDA method suffers from a number of drawbacks: for example, when several features converge to a small region ("coalescence" [48]) the algorithm cannot tell them apart. Several techniques have been proposed to overcome this sort of problems, such as the *condensation* algorithm [49]. However, assume that the feature points represent fixed locations $\{M_i, i = 1, ..., M\}$ on an articulated body, and that we know the rigid motion constraints between pairs of markers. This is equivalent to possessing a *topological model* of the articulated body, represented by an undirected graph whose edges correspond to rigid motion constraints (see Figure 1-left). We can then exploit such *a-priori* information to solve the association task in critical situations where several points fall into the validation region of a single filter. A topological model of the body to track, for instance, can provide:

- a *prediction* constraint, encoding the likelihood of a measurement m_i^k at time k of being associated with a measurement m_i^{k-1} of the previous image;
- an *occlusion* constraint, expressing the chance that a given marker of the model is occluded in the current image;
- a *metric* constraint, representing the knowledge of the lengths of the links, which can be learned from the history of the past associations;
- a *rigid motion* constraint on pairs of markers.

Belief calculus provides a coherent framework in which to combine all these sources of information, and cope with possible conflicts. Indeed, all these constraints can be expressed as belief functions over a suitable "frame", or set of possible associations $m_i \leftrightarrow m_j$ between feature points.

2.2. Rigid motion constraints as conditional belief functions

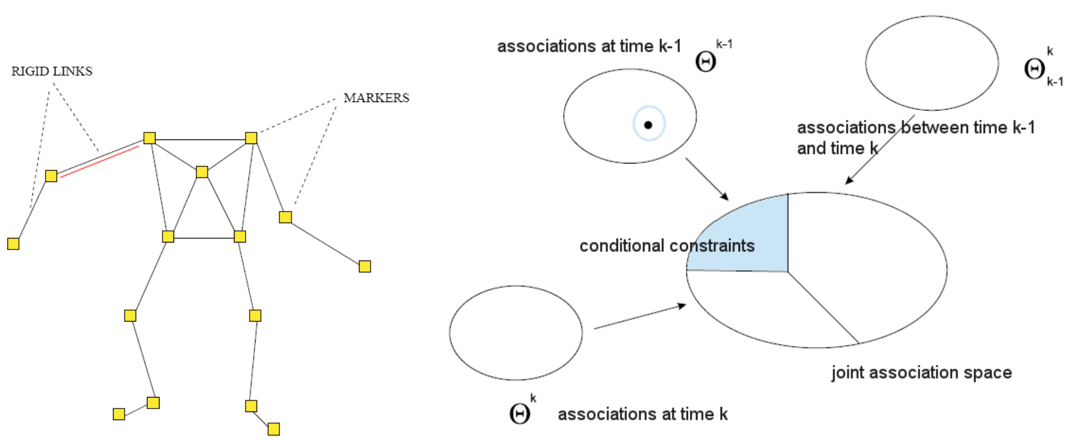

Figure 1: Left: topological model of a human body: the adjacency relations between the markers are shown. Right: rigid motion constraints in the data association problem involve the integration of a set of conditional belief functions in each partition of the joint association space in a single total function.

However, by observing the nature of the above constraints we can note that the information carried by predictions of filters and occlusions inherently concerns *associations between feature points belonging to consecutive images*. Other conditions, instead, such as the metric constraint, can be expressed *instantaneously* in the frame of the current time-k associations. Finally, a number of bodies of evidence depend on the *model-measurement associations* $m_i^{k-1} \leftrightarrow M_j$ at the previous time step. This is the case of belief functions encoding the information carried by the motion of the body, expression of rigid motion constraints.
We can then introduce the set or "frame" of *past model-to-feature* associations,

$$\Theta_M^{k-1} \doteq \{m_i^{k-1} \leftrightarrow M_j, \ \forall i = 1, ..., n^{k-1} \ \forall j = 1, ..., M\}$$

the set of *feature-to-feature* associations,

$$\Theta_k^{k-1} \doteq \{m_i^{k-1} \leftrightarrow m_j^k, \ \forall i = 1, ..., n^{k-1} \ \forall j = 1, ..., n^k\}$$

and *current model-to-feature* associations

$$\Theta_M^k \doteq \{m_i^k \leftrightarrow M_j, \ \forall i = 1, ..., n^k \ \forall j = 1, ..., M\}, \tag{2}$$

where n^k is the number of feature points $\{m_i^k\}$ appearing in image I_k.
All the available pieces of evidence can be combined on the "minimal refinement" [5] of all these frames, the *product association* frame $\Theta_M^{k-1} \otimes \Theta_k^{k-1}$. The result is later projected onto the *current* association set Θ_M^k in order to yield the best current estimate.
However, as we mentioned above, rigid motion constraints can be expressed in a *conditional* way only: hence, the computation of a belief estimate of the desired current model-measurement associations (2) involves *combining conditional belief functions* defined over the product association frame.
The purpose of this paper is to study how such conditional b.f.s can be induced by minimizing geometric distances between belief measures. We first need to recall the bases of belief calculus.

3. Belief functions and their geometric representation

Belief functions. A *basic probability assignment* (b.p.a.) over a finite set *(frame of discernment [5])* Θ is a function $m_b : 2^\Theta \rightarrow [0,1]$ on its power set $2^\Theta = \{A \subseteq \Theta\}$ such that $m_b(\emptyset) = 0$, $\sum_{A \subseteq \Theta} m_b(A) = 1$, and $m_b(A) \geq 0$

$\forall A \subseteq \Theta$. Subsets of Θ associated with non-zero values of m_b are called *focal elements*. The *core* $\mathcal{C}_b$ of a b.f. b is the union of its focal elements. The *belief function* $b : 2^\Theta \rightarrow [0,1]$ associated with a basic probability assignment m_b on Θ is defined as: $b(A) = \sum_{B \subseteq A} m_b(B)$. A dual mathematical representation of the evidence encoded by a belief function b is the *plausibility function* $pl_b : 2^\Theta \rightarrow [0,1]$, $A \mapsto pl_b(A)$, where the plausibility value $pl_b(A)$ of an event A is given by $pl_b(A) \doteq 1 - b(A^c) = \sum_{B \cap A \neq \emptyset} m_b(B)$ and expresses the amount of evidence *not against* A.

Belief functions as vectors. 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 vector of $\mathbb{R}^{N-2}$: $\vec{b} = [b(A), \emptyset \subsetneq A \subsetneq \Theta]'$. If we denote by b_A the *categorical* [50] belief function assigning all the mass to a single subset $A \subseteq \Theta$, $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0 \; \forall B \subseteq \Theta$, $B \neq A$, we can prove that [51, 37] the set of points of $\mathbb{R}^{N-2}$ which correspond to a b.f. or "belief space" $\mathcal{B}$ coincides with the convex closure Cl of all the vectors representing categorical belief functions: $\mathcal{B} = Cl(\vec{b}_A, \emptyset \subsetneq A \subseteq \Theta)$[1].
The belief space $\mathcal{B}$ is a simplex [37], and each vector $\vec{b} \in \mathcal{B}$ representing a belief function b can be written as a convex sum as:

$$\vec{b} = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B) \vec{b}_B. \tag{3}$$

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

$$\vec{m}_b = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B) \vec{m}_B, \tag{4}$$

where $\vec{m}_B$ is the vector of mass values associated with the categorical belief function b_B. Note that in $\mathbb{R}^{N-1}$ $\vec{m}_\Theta = [0, ..., 0, 1]'$ cannot be neglected.

Conditioning simplices. Similarly, the vector $\vec{m}_a$ associated with any belief function a whose mass supports only focal elements $\{\emptyset \subsetneq B \subseteq A\}$

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

included in a given event A can be decomposed as:

$$\vec{m}_a = \sum_{\emptyset \subsetneq B \subseteq A} m_a(B) \vec{m}_B. \tag{5}$$

The set of such vectors is a simplex $\mathcal{M}_A \doteq Cl(\vec{m}_B,\ \emptyset \subsetneq B \subseteq A)$. The same is true in the belief space, where (the vector $\vec{a}$ associated with) each b.f. a assigning mass to focal elements included in A only is decomposable 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 simplices* in the mass and the belief space, respectively.

Geometric conditional belief functions. 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) $b_{d,\mathcal{M}}(.|A)$ ($b_{d,\mathcal{B}}(.|A)$) which minimize(s) the distance $d(\vec{m}_b, \mathcal{M}_A)$ ($d(\vec{b}, \mathcal{B}_A)$) between the mass (belief) vector representing b and the conditioning simplex associated with A in $\mathcal{M}$ ($\mathcal{B}$).
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 imagined or designed [39].
We consider here as distance functions the three major L_p norms $d = L_1$, $d = L_2$ and $d = L_\infty$. This is not to claim that these are *the* distance functions of choice for this problem. The geometric approach to conditioning is potentially a wide research field which will take much time to explore properly. In recent times, however, L_p norms have been successfully employed in different problems such as probability [52] and possibility [53, 54] transformation/approximation, or conditioning [36, 38].

L_p norms in the mass and belief spaces. For vectors $\vec{m}_b, \vec{m}_{b'} \in \mathcal{M}$ representing the b.p.a.s of two belief functions b, b', such norms read as:

$$\|\vec{m}_b - \vec{m}_{b'}\|_{L_1} \doteq \sum_{\emptyset \subsetneq B \subseteq \Theta} |m_b(B) - m_{b'}(B)|, \quad \|\vec{m}_b - \vec{m}_{b'}\|_{L_\infty} \doteq \max_{\emptyset \subsetneq B \subseteq \Theta} |m_b(B) - m_{b'}(B)|$$
$$\|\vec{m}_b - \vec{m}_{b'}\|_{L_2} \doteq \sqrt{\sum_{\emptyset \subsetneq B \subseteq \Theta} (m_b(B) - m_{b'}(B))^2}, \tag{6}$$

while the same norms in the belief space are given by:

$$\|\vec{b}-\vec{b}'\|_{L_1} \doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |b(B)-b'(B)|; \quad \|\vec{b}-\vec{b}'\|_{L_2} \doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta} (b(B)-b'(B))^2};$$
$$\|\vec{b}-\vec{b}'\|_{L_\infty} \doteq \max_{\emptyset\subsetneq B\subseteq\Theta} |b(B)-b'(B)|. \tag{7}$$

We first run our analysis in the mass space $\mathcal{M}$.

4. Geometric conditional belief functions in $\mathcal{M}$

4.1. Conditioning by L_1 norm

Given a belief function b with basic probability assignment m_b collected in a vector $\vec{m}_b \in \mathcal{M}$, its L_1 conditional version(s) $b_{L_1,\mathcal{M}}(.|A)$ has/have basic probability assignment $m_{L_1,\mathcal{M}}(.|A)$ such that:

$$\vec{m}_{L_1,\mathcal{M}}(.|A) \doteq \arg\min_{\vec{m}_a\in\mathcal{M}_A} \|\vec{m}_b-\vec{m}_a\|_{L_1}. \tag{8}$$

Using the expression (6) of the L_1 norm in the mass space $\mathcal{M}$, (8) becomes:

$$\arg\min_{\vec{m}_a\in\mathcal{M}_A} \|\vec{m}_b-\vec{m}_a\|_{L_1} = \arg\min_{\vec{m}_a\in\mathcal{M}_A} \sum_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B)-m_a(B)|.$$

By exploiting the fact that the candidate solution $\vec{m}_a$ is an element of $\mathcal{M}_A$ (Equation (5)) we can greatly simplify this expression.

Lemma 1. *The difference vector $\vec{m}_b-\vec{m}_a$ in $\mathcal{M}$ has the form:*

$$\vec{m}_b-\vec{m}_a = \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)\vec{m}_B + \Big(b(A)-1-\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\Big)\vec{m}_A + \sum_{B\not\subset A} m_b(B)\vec{m}_B \tag{9}$$

where $\beta(B) \doteq m_b(B)-m_a(B)$.

In the L_1 case therefore

$$\|\vec{m}_b-\vec{m}_a\|_{L_1} = \sum_{\emptyset\subsetneq B\subsetneq A} |\beta(B)| + \Big|b(A)-1-\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\Big|, \tag{10}$$

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

$$\sum_i |x_i| + \Big|-\sum_i x_i - k\Big|, \quad k\geq 0 \tag{11}$$

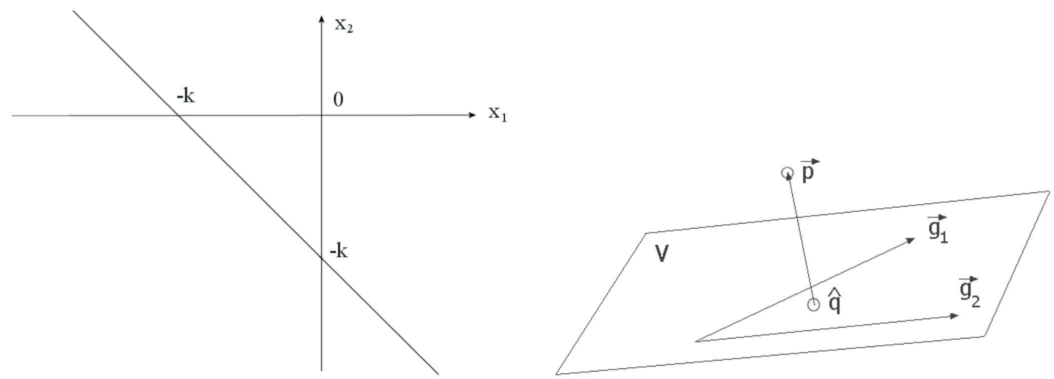

Figure 2: Left: the minima of a function of the form (11) with two variables x_1, x_2 form the triangle $x_1 \leq 0$, $x_2 \leq 0$, $x_1 + x_2 \geq -k$ depicted here. Right: the point $\vec{q}$ of a vector space V at minimal L_2 distance from a given point $\vec{p}$ external to it is such that the difference vector $\vec{p} - \vec{q}$ is orthogonal to all the generators $\vec{g}_i$ of V.

which has an entire simplex of minima, namely: $x_i \leq 0 \ \forall i$, $\sum_i x_i \geq -k$. See Figure 2-left for the case of two variables, x_1 and x_2 (corresponding to the L_1 conditioning problem on an event A of size $|A| = 2$). A similar behavior takes place in the general case too.

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 proper subsets of A:*

$$b_{L_1,\mathcal{M}}(.|A) = \Big\{ a : 2^\Theta \rightarrow [0,1] : \mathcal{C}_a \subseteq A, \ m_a(B) \geq m_b(B) \ \forall \emptyset \subsetneq B \subseteq A \Big\}. \tag{12}$$

As in the toy example of Figure 2-left, the set of L_1 conditional belief function in $\mathcal{M}$ has geometrically the form of a simplex.

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*

$$\mathcal{M}_{L_1,A}[b] = Cl(\vec{m}[b]|^B_{L_1} A, \emptyset \subsetneq B \subseteq A)$$

whose vertex $\vec{m}[b]|^B_{L_1} A$, $\emptyset \subsetneq B \subseteq A$, has coordinates $\{m_a(B)\}$ such that

$$\begin{cases} m_a(B) = m_b(B) + 1 - b(A) = m_b(B) + pl_b(A^c), \\ m_a(X) = m_b(X) \qquad \forall \emptyset \subsetneq X \subsetneq A, X \neq B. \end{cases} \tag{13}$$

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 $m_a(B)$ could have been negative. This is indeed the case for geometrical b.f.s induced by other norms, as we will see in the following.

4.2. Conditioning by L_2 norm

Let us now compute the analytical form of the L_2 conditional belief function(s) in the mass space. We make use of the form (9) of the difference vector $\vec{m}_b - \vec{m}_a$, where again $\vec{m}_a$ is an arbitrary vector of the conditional simplex $\mathcal{M}_A$. In this case, though, it is convenient to recall that the minimal L_2 distance between a point $\vec{p}$ and a vector space is attained by the point $\hat{q}$ of the vector space V s.t. the difference vector $\vec{p} - \hat{q}$ is orthogonal to all the generators $\vec{g}_i$ of V:

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

whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\vec{g}_i, i)$ (Figure 2-right).
This fact is used in the proof of Theorem 3.

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 function $b_{L_2,\mathcal{M}}(.|A)$ 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: $\forall \emptyset \subsetneq B \subseteq A$*

$$m_{L_2,\mathcal{M}}(B|A) \;=\; m_b(B) + \frac{1}{2^{|A|}-1} \sum_{B \not\subset A} m_b(B) = m_b(B) + \frac{pl_b(A^c)}{2^{|A|}-1}. \quad (14)$$

According to Equation (14) the L_2 conditional belief function is unique, and corresponds to the mass function which *redistributes the mass the original belief function assigns to focal elements not included in A to each and all the subsets of A in an equal, even way.*
L_2 and L_1 conditional belief functions in $\mathcal{M}$ display a strong relationship.

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 (13), is the vector

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

whose entry B is given by $\frac{1}{2^{|A|}-1}\Big[m_b(B)(2^{|A|}-1)+(1-b(A))\Big]$, i.e., (14). □

4.3. Conditioning by L_∞ norm

Similarly, we can use Equation (9) to minimize the L_∞ distance between the original mass vector $\vec{m}_b$ and the conditioning subspace $\mathcal{M}_A$. Let us recall it here for sake of readability:

$$\vec{m}_b-\vec{m}_a=\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\vec{m}_B+\sum_{B\not\subset A}m_b(B)\vec{m}_B+\Big(b(A)-1-\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\Big)\vec{m}_A.$$

Its L_∞ norm reads as $\|\vec{m}_b-\vec{m}_a\|_{L_\infty}=$

$$=\max\Big\{|\beta(B)|,\emptyset\subsetneq B\subsetneq A;|m_b(B)|,B\not\subset A;\Big|b(A)-1-\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\Big|\Big\}.$$

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

$$\max\Big\{|\beta(B)|,\emptyset\subsetneq B\subsetneq A;\ \max_{B\not\subset A}\{m_b(B)\};\ \Big|\sum_{B\not\subset A}m_b(B)+\sum_{\emptyset\subsetneq B\subsetneq A}\beta(B)\Big|\Big\}. \tag{15}$$

This is a function of the form

$$f(x_1,...,x_{m-1}) \ = \ \max\Big\{|x_i|\ \forall i,\Big|\sum_i x_i+k_1\Big|,k_2\Big\}, \tag{16}$$

with $0\leq k_2\leq k_1\leq 1$. Consider the case $m=3$. Such a function has two possible behaviors in terms of its minimal region in the plane x_1,x_2.
If $k_1\leq 3k_2$ its contour function has the form rendered in Figure 3-left. The set of minimal points is given by $x_i\geq -k_2$, $x_1+x_2\leq k_2-k_1$. In the opposite case $k_1>3k_2$ the contour function of (16) is as in Figure 3-right. There is a single minimal point, located in $[-1/3k_1,-1/3k_1]$. For an arbitrary number $m-1$ of variables $x_1,...,x_{m-1}$, the first case is such that $k_2\geq k_1/m$, in which

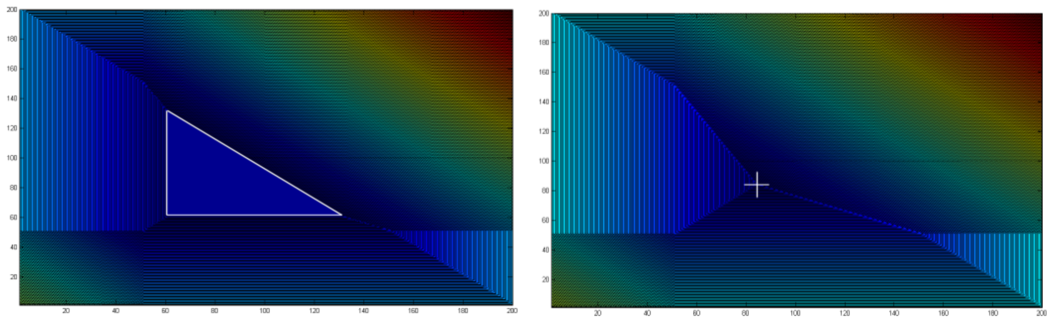

Figure 3: Left: contour function (level sets) and minimal points (white triangle) of a function of the form (16), when $m = 3$, $k_1 \leq 3k_2$. In the example $k_2 = 0.4$ and $k_1 = 0.5$. Right: the case $k_1 \geq 3k_2$. In this example $k_2 = 0.1$ and $k_1 = 0.5$.

situation the set of minimal points of a function of the form (16) is such that $x_i \geq -k_2$, $\sum_i x_i \leq k_2 - k_1$, and forms a simplex with m vertices. Each vertex v^i, $i \neq m$ has components $v^i(j) = -k_2 \ \forall j \neq i$, $v^i(i) = -k_1 + (m-1)k_2$, while obviously $v^m = [-k_2, \cdots, -k_2]'$. In the opposite case the unique minimal point is located in $[(-1/m)k_1, \cdots, (-1/m)k_1]'$.
This analysis applies to the norm (15) as follows.

Theorem 5. *Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, and an arbitrary non-empty focal element $\emptyset \subsetneq A \subseteq \Theta$, the set of L_∞ conditional belief functions $m_{L_\infty,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ forms the simplex*

$$\mathcal{M}_{L_\infty,A}[b] = Cl(\vec{m}[b]|^{\bar{B}}_{L_\infty} A, \ \bar{B} \subseteq A)$$

with vertices

$$\begin{cases} \vec{m}[b]|^{\bar{B}}_{L_\infty}(B|A) = m_b(B) + \max\limits_{C \not\subseteq A} m_b(C) & \forall B \subseteq A, B \neq \bar{B} \\ \vec{m}[b]|^{\bar{B}}_{L_\infty}(\bar{B}|A) = m_b(\bar{B}) + \sum_{C \not\subseteq A} m_b(C) - (2^{|A|} - 2) \max_{C \not\subseteq A} m_b(C). \end{cases} \tag{17}$$

when $\max_{C \not\subseteq A} m_b(C) \geq \frac{1}{2^{|A|}-1} \sum_{C \not\subseteq A} m_b(C)$.
It reduces to the single belief function

$$m_{L_\infty,\mathcal{M}}(B|A) = m_b(B) + \frac{1}{2^{|A|} - 1} \sum_{C \not\subseteq A} m_b(C) \quad \forall B \subseteq A$$

when $\max_{C \not\subseteq A} m_b(C) < \frac{1}{2^{|A|}-1} \sum_{C \not\subseteq A} m_b(C)$. *The latter is the barycenter of the simplex of L_∞ conditional b.f.s in the former case, and coincides with the L_2 conditional belief function (14).*

Note that, as (17) is not guaranteed to be non-negative, the simplex of L_∞ conditional belief functions in $\mathcal{M}$ does not necessarily fall entirely inside the conditioning simplex $\mathcal{M}_A$, i.e., it may include pseudo belief functions. Looking at (17) we can observe that vertices are obtained by assigning the maximum mass not in the conditioning event to all its subsets indifferently. Normalization is then achieved, rather than by normalization (as in Dempster's rule) *by subtracting* of the total mass in excess of 1 in the specific component $\bar{B}$. This behavior is exhibited by other geometric conditional b.f. as shown in the following.

4.4. A case study: the ternary frame

If $|A| = 2$, $A = \{x, y\}$, the conditional simplex is 2-dimensional, with three vertices $\vec{m}_x$, $\vec{m}_y$ and $\vec{m}_{x,y}$. For a b.f. b on $\Theta = \{x, y, z\}$ Theorem 1 states that the vertices of the simplex $\mathcal{M}_{L_1,A}$ of L_1 conditional belief functions in $\mathcal{M}$ are:

$$\begin{array}{lcllll} \vec{m}[b]|_{L_1}^{\{x\}}\{x,y\} & = & [m_b(x) + pl_b(z), & m_b(y), & m_b(x,y) &]', \\ \vec{m}[b]|_{L_1}^{\{y\}}\{x,y\} & = & [m_b(x), & m_b(y) + pl_b(z), & m_b(x,y) &]', \\ \vec{m}[b]|_{L_1}^{\{x,y\}}\{x,y\} & = & [m_b(x), & m_b(y), & m_b(x,y) + pl_b(z) &]'. \end{array}$$

Figure 4 shows such simplex in the case of a belief function b on the ternary frame $\Theta = \{x, y, z\}$ and basic probability assignment

$$\vec{m} = [0.2,\ 0.3,\ 0,\ 0,\ 0.5,\ 0]', \tag{18}$$

i.e., $m_b(x) = 0.2$, $m_b(y) = 0.3$, $m_b(x, z) = 0.5$.

In the case of the belief function (18) of the above example, by Equation (14) its L_2 conditional belief function in $\mathcal{M}$ has b.p.a.:

$$\begin{array}{lcl} m(x) & = & m_b(x) + \dfrac{1 - b(x,y)}{3} = m_b(x) + \dfrac{pl_b(z)}{3}, \quad m(y) = m_b(y) + \dfrac{pl_b(z)}{3}, \\ m(x,y) & = & m_b(x,y) + \dfrac{pl_b(z)}{3}. \end{array} \tag{19}$$

Figure 4 visually confirms that such L_2 conditional belief function lies in the barycenter of the simplex of the related L_1 conditional b.f.s.

For what concerns L_∞ conditional belief functions, the b.f. (18) is such that

$$\begin{array}{lcl} \max\limits_{C \not\subset A} m_b(C) & = & \max\Big\{m_b(z), m_b(x,z), m_b(y,z), m_b(\Theta)\Big\} = m_b(x,z) \\ & = & 0.5 \geq \dfrac{1}{2^{|A|} - 1} \displaystyle\sum_{C \not\subset A} m_b(C) = \dfrac{1}{3} m_b(x,z) = \dfrac{0.5}{3}. \end{array}$$

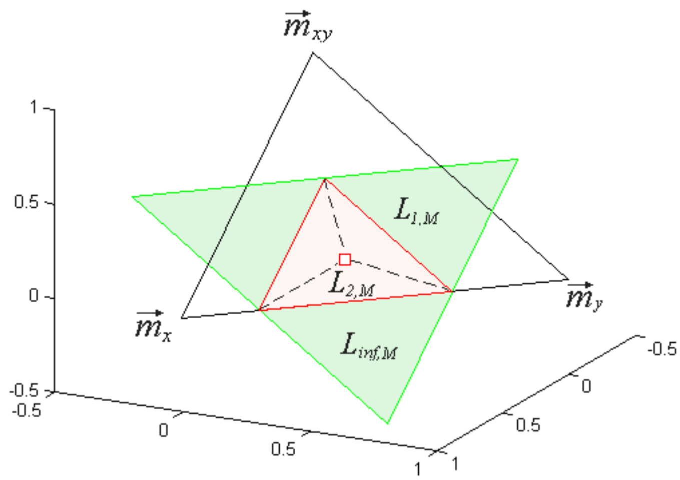

Figure 4: The simplex (solid red triangle) of L_1 conditional belief functions in $\mathcal{M}$ associated with the belief function with mass assignment (18) in $\Theta = \{x, y, z\}$. The related unique L_2 conditional belief function in $\mathcal{M}$ is also plotted as a red square. It coincides with the center of mass of the L_1 set. The set of L_∞ conditional (pseudo) belief functions is also depicted (green triangle).

We hence fall within Case 1, and there is a whole simplex of L_∞ conditional belief function (in $\mathcal{M}$). According to Equation (17) such simplex has $2^{|A|} - 1 = 3$ vertices, namely (taking into account the nil masses in (18)):

$$\begin{array}{lcllll} \vec{m}[b]|^{\{x\}}_{L_\infty,\mathcal{M}}\{x,y\} & = & [m_b(x) - m_b(x,z), & m_b(y) + m_b(x,z), & m_b(x,z) &]', \\ \vec{m}[b]|^{\{y\}}_{L_\infty,\mathcal{M}}\{x,y\} & = & [m_b(x) + m_b(x,z), & m_b(y) - m_b(x,z), & m_b(x,z) &]', \\ \vec{m}[b]|^{\{x,y\}}_{L_\infty,\mathcal{M}}\{x,y\} & = & [m_b(x) + m_b(x,z), & m_b(y) + m_b(x,z), & -m_b(x,z) &]'. \end{array} \quad (20)$$

We can notice that the set of L_∞ conditional (pseudo) b.f.s is not entirely admissible, but its admissible part contains the set of L_1 conditional b.f.s, which amounts therefore a *more conservative* approach to conditioning. Indeed, the latter is the triangle inscribed in the former, determined by its median points. Note also that both the L_1 and L_∞ simplices have the same barycenter in the L_2 conditional b.f. (19).

4.5. Features of geometric conditional belief functions in $\mathcal{M}$

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

- L_p conditional b.f.s, albeit 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 entire 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;
- in particular, while distinct, both the L_1 and L_∞ simplices have barycenter in (or coincide with, in case 2) the L_2 conditional b.f.;
- 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-based 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 $\vec{m}_b$ the mass of a given focal element appears only in the corresponding entry of $\vec{m}_b$. 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. We will see this in detail in Section 5.

4.6. 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* [45]. We will illustrate this notion and how it relates to our results using the example proposed in [45].
Suppose we 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 also that we 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 us the vague impression of having seen approximately the same number of black and white balls, we would probably deem the states a and c equally likely, but at the same time we would tend to deem the event "a or c" twice as likely as the state b. Hence, we assign probability 1/3 to each of the states. Now, we are told that state c is false. How do we revise the probabilities of the two remaining states a and b?

Lewis [44] argued that, upon observing that a certain state $x \in \Theta$ is impossible, we should transfer the probability originally allocated to x to the remaining state deemed 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. We 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 the 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 not difficult to see that this coincides with the set L_1 conditional belief functions $b_{L_1,\mathcal{M}}(.|A)$ of Theorem 1.

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. Both the L_1 and L_2 geometric conditioning operators presented here meet such axiom, supporting the case for their rationality.

5. Geometric 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 $\mathcal{B}_A = Cl(\vec{b}_B, \emptyset \subsetneq B \subseteq A)$ can be solved 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.,

$$\sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B)\vec{b}_B - \sum_{\emptyset \subsetneq B \subseteq A} m_a(B)\vec{b}_B = \\ \sum_{\emptyset \subsetneq B \subseteq A} (m_b(B) - m_a(B))\vec{b}_B + \sum_{B \not\subseteq A} m_b(B)\vec{b}_B = \sum_{\emptyset \subsetneq B \subseteq A} \beta(B)\vec{b}_B + \sum_{B \not\subseteq A} m_b(B)\vec{b}_B, \tag{21}$$

where once again $\beta(B) = m_b(B) - m_a(B)$.

5.1. L_2 conditioning in $\mathcal{B}$

We start with the L_2 norm, as this seems to have a more direct interpretation in the belief space. 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 (21)) reads as

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

for all $\emptyset \subsetneq C \subsetneq A$. Now, categorical b.f.s are such that

$$\langle \vec{b}_B, \vec{b}_C \rangle = |\{Y \supseteq B \cup C, Y \neq \Theta\}| = 2^{|(B \cup C)^c|} - 1 \tag{22}$$

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, $\emptyset \subsetneq C \subsetneq A$,

$$\left\{ \sum_{B \not\subseteq A} m_b(B)\Big[2^{|(B \cup C)^c|} - 2^{|(B \cup A)^c|}\Big] + \sum_{B \subsetneq A} \beta(B)\Big[2^{|(B \cup C)^c|} - 2^{|A^c|}\Big] = 0 \right. \tag{23}$$

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

Theorem 6. *Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, and an arbitrary non-empty focal element $\emptyset \subsetneq A \subseteq \Theta$, the L_2 conditional b.f. $b_{L_2,\mathcal{B}}(.|A)$ with respect to A in the belief space $\mathcal{B}$ is unique, and has b.p.a.*

$$m_{L_2,\mathcal{B}}(C|A) = m_b(C) + \sum_{B \subseteq A^c} m_b(B \cup C) 2^{-|B|} + (-1)^{|C|+1} \sum_{B \subseteq A^c} m_b(B) 2^{-|B|} \tag{24}$$

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

5.1.1. A case study: the ternary frame

In the ternary case (23) reduces to the following system

$$\begin{cases} 2\beta(x) + m_b(z) + m_b(x,z) = 0 \\ 2\beta(y) + m_b(z) + m_b(y,z) = 0 \end{cases} \tag{25}$$

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

$$\begin{array}{l} m_a(x) = m_b(x) + \frac{m_b(z)+m_b(x,z)}{2}, \quad m_a(y) = m_b(y) + \frac{m_b(z)+m_b(y,z)}{2}, \\ m_a(x,y) = m_b(x,y) + m_b(\Theta) + \frac{m_b(x,z)+m_b(y,z)}{2}. \end{array} \tag{26}$$

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:

$$m_\oplus(x|A) = \frac{m_b(x) + m_b(x,z)}{1 - m_b(z)}, \quad m_\oplus(x,y|A) = \frac{m_b(x,y) + m_b(\Theta)}{1 - m_b(z)}.$$

L_2 conditioning in the belief space differs from its "sister" operation in the mass space 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.

5.1.2. Interpretation

In the general case (24) 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 [52], in which the mass of each subset A is re-distributed among all its subsets $B \subseteq A$ on an equal basis. Here (24) the mass of each focal element not included in A is also broken into $2^{|B|}$ parts, equal to the number of its subsets. Only one such part is re-attributed to $C = B \cap A$, while the rest is re-distributed to A itself.

5.2. L_1 conditioning in $\mathcal{B}$

To discuss L_1 conditioning in the belief space we need to write explicitly the difference vector $\vec{b} - \vec{a}$. We need to recall that the vector $\vec{b}_A$ associated with a categorical b.f. b_A has as entries $\vec{b}_A(B) = 1$ if $B \supseteq A$, 0 otherwise.

Lemma 2. *The L_1 norm of the difference vector $\vec{b} - \vec{a}$ can be written as*

$$\|\vec{b} - \vec{a}\|_{L_1} = \sum_{\emptyset \subsetneq B \cap A \subsetneq A} \Big|\gamma(B \cap A) + b(B) - b(B \cap A)\Big|$$

so that the L_1 conditional belief functions in $\mathcal{B}$ are the solutions of the following minimization problem:

$$\arg\min_\gamma \|\vec{b} - \vec{a}\|_{L_1} \quad = \quad \arg\min_\gamma \sum_{\emptyset \subsetneq B \cap A \subsetneq A} \Big|\gamma(B \cap A) + b(B) - b(B \cap A)\Big|,$$

where $\beta(B) = m_b(B) - m_a(B)$ and $\gamma(B) = \sum_{C \subseteq B} \beta(B)$.

As we also noticed in the L_1 minimization problem in the mass space, each group of addenda which depend on the same variable $\gamma(X)$, $\emptyset \subsetneq X \subsetneq A$, can be minimized separately. Therefore, the set of L_1 conditional belief functions in the belief space $\mathcal{B}$ is determined by the following minimization problem:

$$\arg\min_{\gamma(X)} \sum_{B: B \cap A = X} \Big|\gamma(X) + b(B) - b(X)\Big| \qquad \forall \emptyset \subsetneq X \subsetneq A. \tag{27}$$

The functions appearing in (27) are of the form $|x + k_1| + ... + |x + k_m|$, where m is even. Such functions are minimized by the interval determined by the two central "nodes" $-k_{int_1} \leq -k_{int_2}$ (see Figure 5 for an example). In the case of system (27) this yields

$$b(X) - b(B^X_{int_1}) \leq \gamma(X) \leq b(X) - b(B^X_{int_2}) \tag{28}$$

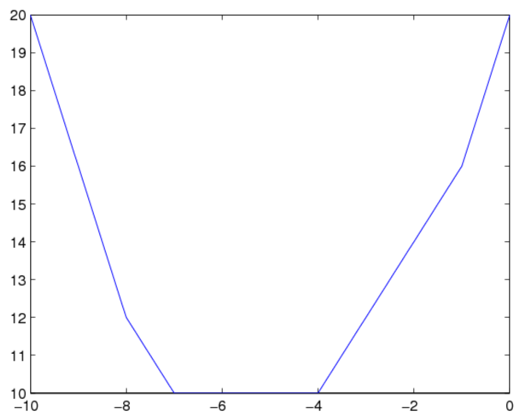

Figure 5: The function $|x+1|+|x+4|+|x+7|+|x+8|$ is minimized by the interval of values delimited by the two central nodes -4 and -7.

where $B^X_{int_1}$ and $B^X_{int_2}$ are the central, median values of the collection $\{b(B), B \cap A = X\}$. Unfortunately, it is not possible, in general, to determine the median values of such a collection of belief values, as belief functions are defined on a partially (rather than totally) ordered set (the power set 2^Θ).

5.2.1. The special case $|A^c| = 1$

This is possible, however, in the special case in which $|A^c| = 1$ (i.e., the conditioning event is of cardinality $n-1$). In this case

$$B^X_{int_1} = b(X + A^c), \quad B^X_{int_2} = b(X)$$

so that the solution in the variables $\{\gamma(X)\}$ is

$$b(X) - b(X + A^c) \leq \gamma(X) \leq 0, \quad \emptyset \subsetneq X \subsetneq A.$$

It is not difficult to see that, in the variables $\{\beta(X)\}$, the solution reads as

$$b(X) - b(X + A^c) \leq \beta(X) \leq - \sum_{\emptyset \subsetneq B \subsetneq X} \Big(b(B) - b(B + A^c) \Big),$$

$\emptyset \subsetneq X \subsetneq A$, i.e., in the mass of the desired L_1 conditional belief function,

$$m_b(X) + \sum_{\emptyset \subsetneq B \subsetneq X} \Big(b(B) - b(B + A^c) \Big) \leq m_a(X) \leq m_b(X) + \Big(b(X + A^c) - b(X) \Big). \tag{29}$$

Not only the resulting conditional (pseudo) belief functions are not guaranteed to be proper belief functions (see Equation (29)), but it difficult to find straightforward interpretations for these results in terms of degrees of belief. On these grounds, we would be tempted to conclude that the L_1 norm is not suitable to induce conditioning in belief calculus. However, the numerical example of Section 6.1 seems to hint otherwise.
In the ternary case, in particular, the L_1 conditional b.f.s $b_{L_1,\mathcal{B}_3}(.|\{x,y\})$ with respect to $A = \{x,y\}$ in $\mathcal{B}_3$ are all those b.f.s with core included in A such that the conditional mass of $B \subsetneq A$ falls between $b(B)$ and $b(B \cup A^c)$:

$$b(B) \leq m_{L_1,\mathcal{B}}(B|A) \leq b(B \cup A^c).$$

One could be tempted to conjecture that this simple and elegant behavior is indeed a general feature of L_1 conditioning on general belief spaces.
Another fact one can remark is that the barycenter of the L_1 solutions is

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

i.e., the L_2 conditional b.f. (26), just like in the case of geometric conditioning in the mass space $\mathcal{M}$. The same can be easily proved for all $A \subseteq \{x,y,z\}$.

Theorem 7. *For every belief function $b : 2^{\{x,y,z\}} \to [0,1]$, the unique L_2 conditional b.f. $b_{L_1,\mathcal{B}_3}(.|\{x,y\})$ with respect to $A \subseteq \{x,y,z\}$ in $\mathcal{B}_3$ is the barycenter of the polytope of L_1 conditional b.f.s with respect to A in $\mathcal{B}_3$.*

5.3. L_∞ conditioning in $\mathcal{B}$

Let us finally approach the problem of finding L_∞ conditional belief functions given an event A, starting with the ternary case study.

5.3.1. The ternary case

In the ternary case, $\|\vec{b} - \vec{a}\|_{L_\infty} = \max_{\emptyset \subsetneq B \subsetneq \Theta} |b(B) - a(B)| =$

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

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

$$\begin{array}{rcl} \beta(x) & : & \max\Big\{ |\beta(x)|, |\beta(x) + m_b(z) + m_b(x,z)| \Big\} \leq 1 - b(x,y) \\ \beta(y) & : & \max\Big\{ |\beta(y)|, |\beta(y) + m_b(z) + m_b(y,z)| \Big\} \leq 1 - b(x,y). \end{array}$$

On the left hand side we have functions of the form $\max\{|x|, |x+k|\}$, whose shape is plotted in Figure 6. The interval of values in which such a function is below a certain threshold $k' \geq k$ is $[-k', k'-k]$. This yields:

$$\begin{array}{rcccl} b(x,y) - 1 & \leq & \beta(x) & \leq & 1 - b(x,y) - (m_b(z) + m_b(x,z)) \\ b(x,y) - 1 & \leq & \beta(y) & \leq & 1 - b(x,y) - (m_b(z) + m_b(y,z)). \end{array} \tag{30}$$

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

$$\begin{array}{rcccl} m_b(x) - m_b(y,z) - m_b(\Theta) & \leq & m_a(x) & \leq & 1 - (m_b(y) + m_b(x,y)) \\ m_b(y) - m_b(x,z) - m_b(\Theta) & \leq & m_a(y) & \leq & 1 - (m_b(x) + m_b(x,y)). \end{array} \tag{31}$$

Its barycenter is clearly given by

$$\begin{array}{rcl} m_a(x) & = & m_b(x) + \frac{m_b(z) + m_b(x,z)}{2} \quad m_a(y) = m_b(y) + \frac{m_b(z) + m_b(y,z)}{2} \\ m_a(x,y) & = & 1 - m_a(x) - m_a(y) = m_b(x,y) + m_b(\Theta) + \frac{m_b(x,z) + m_b(y,z)}{2} \end{array} \tag{32}$$

i.e., the L_2 conditional belief function (26) as computed in the ternary case.

5.3.2. The general case

From the expression (19) of the difference $\vec{b} - \vec{a}$ we get, after introducing the variables $\gamma(C) = \sum_{X \subseteq C} \beta(X)$, $\emptyset \subsetneq C \subseteq A$: $\max_{\emptyset \subsetneq B \subsetneq \Theta} |\vec{b} - \vec{a}(B)| =$

$$\begin{array}{l} = \max_{\emptyset \subsetneq B \subsetneq \Theta} \Big| \sum_{C \subseteq A \cap B} \beta(C) + \sum_{C \subseteq B, C \not\subseteq A} m_b(C) \Big| = \max_{\emptyset \subsetneq B \subsetneq \Theta} \Big| \gamma(A \cap B) + \sum_{C \subseteq B, C \not\subseteq A} m_b(C) \Big| \\ = \max \Bigg\{ \max_{B: B \cap A = \emptyset} \Big| \sum_{C \subseteq B, C \not\subseteq A} m_b(C) \Big|, \max_{B: B \cap A \neq \emptyset, A} \Big| \gamma(A \cap B) + \sum_{C \subseteq B, C \not\subseteq A} m_b(C) \Big|, \\ \qquad \max_{B: B \cap A = A, B \neq \Theta} \Big| \gamma(A) + \sum_{C \subseteq B, C \not\subseteq A} m_b(C) \Big| \Bigg\}, \end{array} \tag{33}$$

where, once again, $\gamma(A) = b(A) - 1$.

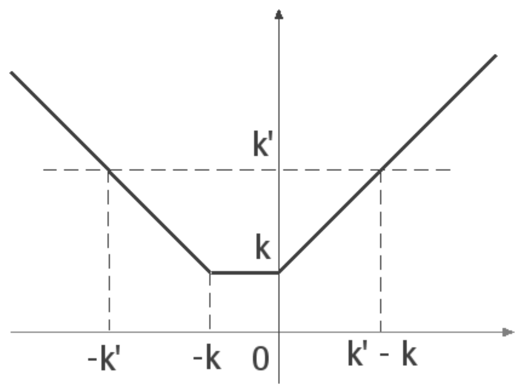

Figure 6: Graph of a function of the form $\max\{|x|, |x+k|\}$, and the interval of values for which the function is below a threshold $k' >= k$.

Lemma 3. *The values $\gamma^*(X)$ which minimize (33) are, $\forall\ \emptyset \subsetneq X \subsetneq A$:*

$$-(1-b(A)) \leq \gamma^*(X) \leq (1-b(A)) - \sum_{C\cap A^c \neq \emptyset, C\cap A \subseteq X} m_b(C). \tag{34}$$

Lemma 3 can be used to prove the following form of the set of L_∞ conditional belief functions in $\mathcal{B}$.

Theorem 8. *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_∞ 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 $\emptyset \subsetneq X \subseteq A$:*

$$\begin{array}{l} m_b(X) + \displaystyle\sum_{C\cap A^c \neq \emptyset, \emptyset \subseteq C\cap A \subseteq X} m_b(C) + (2^{|X|}-1)(1-b(A)) \leq m_a(X) \leq m_b(X) \\ +(2^{|X|}-1)(1-b(A)) - \displaystyle\sum_{C\cap A^c \neq \emptyset, \emptyset \subseteq C\cap A \subsetneq X} m_b(C) - (-1)^{|X|} \sum_{B\subseteq A^c} m_b(B). \end{array} \tag{35}$$

This result appears of rather difficult interpretation in terms of mass allocation. Nevertheless, the ternary example we will see in Section 6.1 seems to suggest that this set, or at least its admissible part, has some nice properties that would be worth to explore.
For instance, its barycenter has a much simpler form.

5.3.3. Barycenter of the L_∞ solution

The barycenter of (34) 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:

$$\left\{ \sum_{\emptyset\subsetneq C\subseteq X} \beta(C) + \frac{1}{2} \sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C) = 0, \quad \forall \emptyset \subsetneq X \subsetneq A. \right. \tag{36}$$

We can prove the following result, which we will discuss in the subsequent summary section.

Theorem 9. *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 (36), and has basic probability assignment:*

$$\begin{aligned} m_{L_\infty,\mathcal{B}}(C|A) &= m_b(C) + \frac{1}{2} \sum_{\emptyset\subsetneq B\subseteq A^c} \Big[m_b(B\cup C) + (-1)^{|C|+1} m_b(B) \Big] \\ &= m_b(C) + \frac{1}{2} \sum_{\emptyset\subsetneq B\subseteq A^c} m_b(B+C) + \frac{1}{2}(-1)^{|C|+1} b(A^c). \end{aligned} \tag{37}$$

6. Comparison with conditioning in $\mathcal{M}$

It is interesting to compare the outcomes of L_p conditioning in the mass space (Section 4) and the belief space (Section 5). Given a belief function $b : 2^\Theta \to [0,1]$ and an arbitrary non-empty focal element $\emptyset \subsetneq A \subseteq \Theta$:

1. 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:

$$b_{L_1,\mathcal{M}}(.|A) = \Big\{ a : \mathcal{C}_a \subseteq A, m_a(B) \geq m_b(B) \quad \forall \emptyset \subsetneq B \subseteq A \Big\}.$$

   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 $\vec{m}_a = \vec{m}[b]|^B_{L_1} A$ have b.p.a.:

$$\left\{ \begin{array}{l} m_a(B) = m_b(B) + 1 - b(A) = m_b(B) + pl_b(A^c), \\ m_a(X) = m_b(X) \quad \forall \emptyset \subsetneq X \subsetneq A, X \neq B; \end{array} \right.$$

2. the unique L_2 conditional belief function $b_{L_2,\mathcal{M}}(.|A)$ with respect to A in $\mathcal{M}$ is the b.f. whose b.p.a. redistributes the mass $1-b(A)=pl_b(A^c)$ to each focal element $B \subseteq A$ in an equal way:

$$m_{L_2,\mathcal{M}}(B|A) = m_b(B) + \frac{pl_b(A^c)}{2^{|A|}-1}, \tag{38}$$

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

3. the L_∞ conditional b.f. either coincides with the L_2 one, or forms a simplex obtained by assigning the maximal mass outside A (rather than the sum of such masses $pl_b(A^c)$) to all subsets of A (but one) indifferently.

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 [45, 46].
The L_2 and $\overline{L_\infty}$ conditional b.f.s just computed in the belief space are instead:

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

As for the L_2 case, the result makes a lot of sense in the ternary case, but it is difficult to interpret in its general form (above). It seems to be related to the process of mass redistribution among all subsets, as it happens with the (L_2 induced) orthogonal projection of a belief function onto the probability simplex. In both expressions above we can note that normalization is achieved by alternatively subtracting and summing a quantity, rather than via a ratio or, as in Equation (38), by reassigning the mass of all $B \not\subset A$ to each $B \subsetneq A$ on equal grounds.

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 normalization, all the subsets $D \supseteq A$ including A have their whole mass re-assigned to A, consistently with the above interpretation. The mass $b(A^c)$ of the subsets 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 9 with Theorem 6 we will notice that, while being quite similar to each other, the coincidence of $b_{\overline{L_\infty},\mathcal{B}}(.|A)$ and $b_{L_2,\mathcal{B}}(.|A)$ was really an artifact of the ternary case, for which $\emptyset \subsetneq B \subseteq A^c \equiv \emptyset \subsetneq B \subseteq \{z\}^c \equiv B = \{z\}, |B| = 1$. The L_1 case is also intriguing, as in that case it appears impossible to obtain an analytic expression, while when this is possible the result is intriguing, as confirmed in the following empirical comparison. In general, though, conditional b.f.s in the belief space seem to have rather less straightforward interpretations than the corresponding quantities in the mass space.

6.1. Comparison on the ternary example

We conclude by comparing the different approximations in the case study of a ternary frame, $\Theta = \{x, y, z\}$, already introduced in Section 4.4. Assuming again that we want the conditioning event approximation to be $A = \{x, y\}$, the unique L_2 conditional belief function in $\mathcal{B}$ is given by Equation (26), while the L_∞ conditional b.f.s form the set determined by Equation (31), with barycenter in (32).
By Theorem 1 the vertices of $\mathcal{M}_{L_1,\{x,y\}}[b]$ are instead

$$\begin{array}{rcl} \vec{m}[b]|^{\{x\}}_{L_1}\{x,y\} & = & [m_b(x) + pl_b(z), m_b(y), m_b(x,y)]', \\ \vec{m}[b]|^{\{y\}}_{L_1}\{x,y\} & = & [m_b(x), m_b(y) + pl_b(z), m_b(x,y)]', \\ \vec{m}[b]|^{\{x,y\}}_{L_1}\{x,y\} & = & [m_b(x), m_b(y), m_b(x,y) + pl_b(z)]'. \end{array}$$

while by Theorem 3 the L_2 conditional b.f. given $\{x, y\}$ in $\mathcal{M}$ has b.p.a.:

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

In this case the conditional simplex is 2-dimensional, with three vertices $\vec{b}_x$, $\vec{b}_y$ and $\vec{b}_{x,y}$. Figure 7 illustrates the different geometric conditional belief functions given $A = \{x, y\}$ for the belief function with masses as in (18), 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. 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.

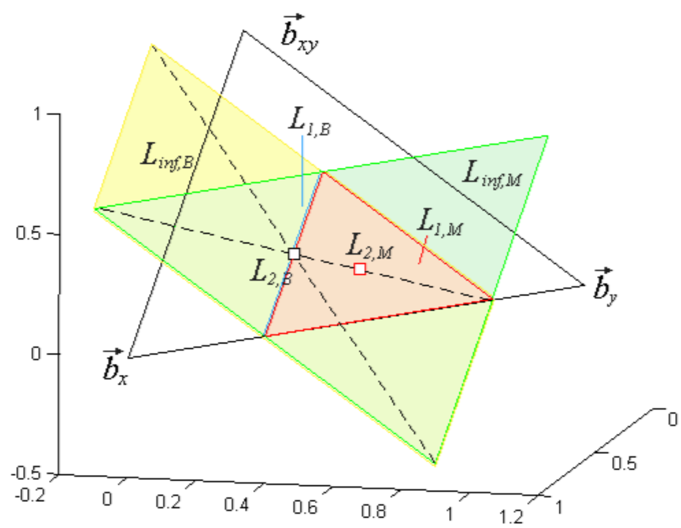

Figure 7: The simplex (red triangle) of $L_1, \mathcal{M}$ conditional belief functions associated with the belief function with mass assignment (18) in $\Theta = \{x, y, z\}$, with conditioning event $A = \{x, y\}$. The related $L_2, \mathcal{M}$ conditional belief function is plotted as a red square, and coincides with the center of mass of the L_1 set. The set of $L_\infty, \mathcal{M}$ conditional belief functions is represented as the green triangle containing $L_2, \mathcal{M}$. The set of $L_\infty, \mathcal{B}$ conditional b.f.s is drawn as a blue rectangle, and also falls in part outside the conditioning simplex (black triangle). The set of L_1 conditional b.f.s in $\mathcal{B}$ is a (light blue) line segment with barycenter in the L_2 conditional b.f. (black square). 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 which are not clearly reflected by their analytical expressions computed here.

The set of L_1 conditional b.f.s in $\mathcal{B}$ is a line segment with barycenter in the L_2 conditional b.f. $m_{L_2,\mathcal{B}}(.|A)$, which is:

- entirely included in the set of L_∞ approximations in both $\mathcal{B}$ and $\mathcal{M}$, i.e., a more conservative approach to conditioning;
- *entirely admissible.*

It seems that, hard as it is to compute, L_1 conditioning in the belief space produces interesting results. A number of interesting cross relations between conditional b.f. of the two representation domains appear to exist.

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 joining opposite vertices of $m_{L_\infty,\mathcal{B}}(.|A)$;
3. $m_{L_\infty,\mathcal{B}}(.|A)$ and $m_{L_\infty,\mathcal{M}}(.|A)$ have several vertices in common.

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 parts of $m_{L_\infty,\mathcal{B}}(.|A)$ and $m_{L_\infty,\mathcal{M}}(.|A)$.

7. 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 simplex 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_1, L_2 and L_∞. The present analysis opens a number of interesting questions.

7.1. Geometrical rules of combination

We may wonder, for instance, what classes of conditioning rules can be generated by such a distance minimization 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 [39] and their survey of the distance or similarity measures so far introduced between belief functions. Such as line of research could possibly be very useful in our quest. A related question links geometric conditioning with combination rules. Indeed, in the case of Dempster's rule it can be easily proven that [36],

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

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 from conditioning rules. Additional constraints have to be imposed in order to obtain a unique result. For instance, by imposing commutativity with affine combination (*linearity*, in Smets' terminology [55]), any (geometrical) conditioning rule $b|^{\uplus}_A$ implies:

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

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.

7.2. L_p consonant/consistent 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 [56] 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 [57]. 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. A preliminary analysis has been conducted in the case of consistent belief functions, but only in the mass space, in [57]. Given the paramount importance of possibility measures and fuzzy sets among practitioners, we plan to do the same for consonant belief functions.

Appendix

Proof of Lemma 1

By definition, $\vec{m}_b - \vec{m}_a = \sum_{\emptyset \subsetneq B \subseteq \Theta} m_b(B) \vec{m}_B - \sum_{\emptyset \subsetneq B \subseteq A} m_a(B) \vec{m}_B$.
The change of variables $\beta(B) \doteq m_b(B) - m_a(B)$ further yields:

$$\vec{m}_b - \vec{m}_a = \sum_{\emptyset \subsetneq B \subseteq A} \beta(B) \vec{m}_B + \sum_{B \not\subset A} m_b(B) \vec{m}_B. \tag{39}$$

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

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

as $\sum_{\emptyset\subsetneq B\subseteq A} m_a(B) = 1$ by definition, since $\vec{m}_a \in \mathcal{M}_A$. As a consequence, in the optimization problem (8) there are just $2^{|A|}-2$ independent variables (as $\emptyset$ is not included), while: $\beta(A) = b(A) - 1 - \sum_{\emptyset\subsetneq B\subsetneq A} \beta(B)$. By replacing the above equality into (39) we get Equation (9).

Proof of Theorem 1

The minima of the L_1 norm (10) are given by the set of constraints:

$$\begin{cases} \beta(B) \leq 0 & \forall \emptyset \subsetneq B \subsetneq A \\ \displaystyle\sum_{\emptyset\subsetneq B\subsetneq A} \beta(B) \geq b(A) - 1. & \end{cases} \tag{40}$$

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

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

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

Proof of Theorem 2

It is easy to see that, by Equation (40), the $2^{|A|}-2$ vertices of the simplex of L_1 conditional belief function in $\mathcal{M}$ (denoted by $\vec{m}[b]|_{L_1}^{B} A$, where $\emptyset \subsetneq B \subseteq A$) are determined by the following solutions:

$$\begin{array}{l} \vec{m}[b]|_{L_1}^{A} A : \left\{ \begin{array}{lcl} \beta(X) & = & 0 \quad \forall \emptyset \subsetneq X \subsetneq A, \end{array} \right. \\ \vec{m}[b]|_{L_1}^{B} A : \left\{ \begin{array}{lcll} \beta(B) & = & b(A) - 1, & \\ \beta(X) & = & 0 & \forall \emptyset \subsetneq X \subsetneq A, X \neq B. \end{array} \right. \quad \forall \emptyset \subsetneq B \subsetneq A \end{array}$$

In the $\{m_a(B)\}$ coordinates, the vertex $\vec{m}[b]|_{L_1}^{B} A$ is the vector $\vec{m}_a \in \mathcal{M}_A$ which meets Equation (13).

Proof of Theorem 3

In the case that concerns us, $\vec{p} = \vec{m}_b$ is the original mass function, $\vec{q} = \vec{m}_a$ 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 (39), the condition $\langle \vec{m}_b - \vec{m}_a, \vec{m}_B - \vec{m}_A \rangle = 0$ 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.:

$$2\beta(B) + \sum_{\emptyset \subsetneq X \subsetneq A, X \neq B} \beta(X) = b(A) - 1 \quad \forall \emptyset \subsetneq B \subsetneq A. \tag{41}$$

System (41) is a linear system of $2^{|A|} - 2$ equations in $2^{|A|} - 2$ variables (the $\beta(X)$), that can 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 $\mathcal{A}$ and its inverse are

$$\mathcal{A} = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ & & \cdots & \\ 1 & 1 & \cdots & 2 \end{bmatrix}. \quad \mathcal{A}^{-1} = \tfrac{1}{d+1} \begin{bmatrix} d & -1 & \cdots & -1 \\ -1 & d & \cdots & -1 \\ & & \cdots & \\ -1 & -1 & \cdots & d \end{bmatrix},$$

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 to (41) is then

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

or, more explicitly, $\beta(B) = \dfrac{b(A) - 1}{2^{|A|} - 1}$ for all $\emptyset \subsetneq B \subsetneq A$. In the $\{m_a(B)\}$ coordinates the L_2 conditional belief function reads as

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

A included.

Proof of Theorem 5

For the norm (15) the condition $k_2 \geq k_1/m$ for functions of the form (16) reads as:

$$\max_{C \not\subset A} m_b(C) \geq \frac{1}{2^{|A|} - 1} \sum_{C \not\subset A} m_b(C). \tag{42}$$

In such a case the set of L_∞ conditional belief functions is given by the constraints $x_i \geq -k_2$, $\sum_i x_i \leq k_2 - k_1$, i.e.,

$$\begin{cases} \beta(B) \geq -\max_{C \not\subset A} m_b(C) & \forall B \subsetneq A, \\ \sum_{B \subsetneq A} \beta(B) \leq \max_{C \not\subset A} m_b(C) - \sum_{C \not\subset A} m_b(C). \end{cases}$$

This is a simplex $Cl(\vec{m}[b]|_{\bar{B}}^{L_\infty} A, \bar{B} \subseteq A)$, where each vertex $\vec{m}[b]|_{\bar{B}}^{L_\infty} A$ is characterized by the following values $\vec{\beta}_{\bar{B}}$ of the auxiliary variables:

$$\begin{cases} \vec{\beta}_{\bar{B}}(B) = -\max_{C \not\subset A} m_b(C) & \forall B \subseteq A, B \neq \bar{B} \\ \vec{\beta}_{\bar{B}}(\bar{B}) = -\sum_{C \not\subset A} m_b(C) + (2^{|A|} - 2) \max_{C \not\subset A} m_b(C) \end{cases}$$

or, in terms of their basic probability assignments, (17).
The barycenter of this simplex can be computed as follows:

$$\begin{aligned} m_{\overline{L_\infty}, \mathcal{M}}(B|A) &= \frac{\sum_{\bar{B} \subseteq A} \vec{m}[b]|_{\bar{B}}^{L_\infty}(B|A)}{2^{|A|} - 1} \\ &= \frac{(2^{|A|} - 1) m_b(B) + (2^{|A|} - 2) \max_{C \not\subset A} m_b(C) + \sum_{C \not\subset A} m_b(C) - (2^{|A|} - 2) \max_{C \not\subset A} m_b(C)}{2^{|A|} - 1} \\ &= \frac{(2^{|A|} - 1) m_b(B) + \sum_{C \not\subset A} m_b(C)}{2^{|A|} - 1} = m_b(B) + \frac{\sum_{C \not\subset A} m_b(C)}{2^{|A|} - 1}, \end{aligned}$$

i.e., the L_2 conditional belief function (14). The corresponding minimal L_∞ norm of the difference vector is, according to (15), equal to $\max_{C \not\subset A} m_b(C)$.
The opposite case reads as

$$\max_{C \not\subset A} m_b(C) < \frac{1}{2^{|A|} - 1} \sum_{C \not\subset A} m_b(C). \tag{43}$$

For system (15) the unique solution is $\beta(B) = -\frac{1}{2^{|A|} - 1} \sum_{C \not\subset A} m_b(C)$ for all $B \subsetneq A$ or, in terms of basic probability assignments,

$$m_{L_\infty, \mathcal{M}}(B|A) = m_b(B) + \frac{1}{2^{|A|} - 1} \sum_{C \not\subset A} m_b(C) \quad \forall B \subseteq A.$$

The corresponding minimal L_∞ norm of the difference vector is in this second case equal to $\frac{1}{2^{|A|}-1}\sum_{C\not\subset A} m_b(C)$.

Proof of Theorem 6

In the $\{\beta(B)\}$ variables (24) reads as:

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

To prove Theorem 6 we just need to replace the above expression in $\{\beta(B)\}$ into the system of constraints (23). We obtain, for all $\emptyset \subsetneq C \subsetneq A$:

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

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 above system of constraints as

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

As $2^{-|X|}(2^{|(Y\cup C)^c|} - 2^{|A^c|}) = 2^{n-|Y\cup C|-|X|} - 2^{n-|A|-|X|} = 2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|}$ the system of constraints further simplifies as:

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

After separating in the first sum the contributions of $Y = \emptyset$ and $Y = A$, noting that $A \cup C = A$ as $C \subset A$, and splitting the second one into the part which depends on $m_b(X)$ and the one depending on $m_b(X+Y)$, the system of constraints becomes, again for all $\emptyset \subsetneq C \subsetneq A$:

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

and by simplifying further:

$$\sum_{\emptyset \subsetneq X \subseteq A^c} m_b(X)\Big(2^{|(X+C)^c|} - 2^{|(X+A)^c|}\Big) + \\ + \sum_{\emptyset \subsetneq X \subseteq A^c} \sum_{\emptyset \subsetneq Y \subsetneq A} (-1)^{|Y|} m_b(X)\Big(2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|}\Big) = 0. \tag{44}$$

Now, the first addendum is easily reduced to

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

As for the second one, we have:

$$\sum_{\emptyset \subsetneq X \subseteq A^c} \sum_{\emptyset \subsetneq Y \subsetneq A} (-1)^{|Y|} m_b(X)\Big(2^{|[(Y\cup C)+X]^c|} - 2^{|(A+X)^c|}\Big)$$
$$= \sum_{\emptyset \subsetneq X \subseteq A^c} m_b(X) 2^{-|X|} \sum_{\emptyset \subsetneq Y \subsetneq A} (-1)^{|Y|}\Big(2^{|(Y\cup C)^c|} - 2^{|A^c|}\Big)$$
$$= \sum_{\emptyset \subsetneq X \subseteq A^c} m_b(X) 2^{-|X|}\Big[\sum_{\emptyset \subseteq Y \subseteq A} (-1)^{|Y|}\Big(2^{|(Y\cup C)^c|} - 2^{|A^c|}\Big) - (2^{|(\emptyset\cup C)^c|} - 2^{|A^c|})$$
$$-(-1)^{|A|}(2^{|(A\cup C)^c|} - 2^{|A^c|})\Big]$$
$$= \sum_{\emptyset \subsetneq X \subseteq A^c} m_b(X) 2^{-|X|}\Big[\sum_{\emptyset \subseteq Y \subseteq A} (-1)^{|Y|}\Big(2^{|(Y\cup C)^c|} - 2^{|A^c|}\Big) - (2^{|C^c|} - 2^{|A^c|})\Big]. \tag{45}$$

At this stage 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 in (45), using a common technique we can decompose Y into the disjoint sum $Y=(Y\cap C)+(Y\setminus C)$ and rewrite it as

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

where $\displaystyle\sum_{|Y\setminus C|}^{|A\setminus C|}\binom{|A\setminus C|}{|Y\setminus C|}(-1)^{|Y\setminus C|}2^{-|Y\setminus C|}=\big(-1+\frac{1}{2}\big)^{|A\setminus C|}=-2^{-|A\setminus C|}$ by Newton's binomial, so that we get

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

again by Newton's binomial. By replacing this result in cascade into (45) and (44) we have that the system of constraints is always met as it reduces to the equality $0=0$.

Proof of Lemma 2

After introducing the auxiliary variables $\beta(B)=m_b(B)-m_a(B)$ and $\gamma(B)=\sum_{C\subseteq B}\beta(B)$ the desired norm becomes

$$\|\vec{b}-\vec{a}\|_{L_1}=\sum_{\emptyset\subsetneq B\subsetneq\Theta}|b(B)-a(B)|=\sum_{\emptyset\subsetneq B\subsetneq\Theta}\Big|\sum_{\emptyset\subsetneq C\subseteq A\cap B}\beta(C)+\sum_{C\subseteq B,C\not\subset A}m_b(C)\Big|$$
$$=\sum_{\emptyset\subsetneq B\subsetneq\Theta}\Big|\gamma(A\cap B)+\sum_{C\subseteq B,C\not\subset A}m_b(C)\Big|$$
$$=\sum_{B:B\cap A=\emptyset}\Big|\sum_{C\subseteq B,C\not\subset A}m_b(C)\Big|+\sum_{B:B\cap A\neq\emptyset,\Theta}\Big|\gamma(A\cap B)+\sum_{C\subseteq B,C\not\subset A}m_b(C)\Big|$$
$$+\sum_{B:B\cap A=A,B\neq\Theta}\Big|\gamma(A)+\sum_{C\subseteq B,C\not\subset A}m_b(C)\Big|$$

where $\gamma(A) = \sum_{C\subsetneq A} \beta(C) = \sum_{C\subsetneq A} m_b(C) - \sum_{C\subsetneq A} m_a(C) = b(A) - 1$. Thus, the first and the third addenda above are constant, and since

$$\sum_{C\subseteq B, C\not\subset A} m_b(C) = b(B) - b(B\cap A)$$

we obtain, as desired:

$$\arg\min_{\gamma} \|\vec{b} - \vec{a}\|_{L_1} = \arg\min_{\gamma} \sum_{\emptyset\subsetneq B\cap A\subsetneq A} \Big|\gamma(B\cap A) + \sum_{C\subseteq B, C\not\subset A} m_b(C)\Big|.$$

Proof of Lemma 3

The first term in (33) is such that

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

For the third one we have instead

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

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

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

Therefore, the L_∞ norm (33) of the difference $\vec{b} - \vec{a}$ reduces to:

$$\max_{\emptyset\subsetneq B\subsetneq\Theta} |\vec{b} - \vec{a}(B)| = \max\left\{ \max_{B:B\cap A\neq\emptyset, A} \Big|\gamma(A\cap B) + \sum_{C\subseteq B, C\not\subset A} m_b(C)\Big|, 1 - b(A)\right\} \quad (46)$$

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

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

The variable term in (46) can be decomposed into collections of terms which depend on the same individual variable $\gamma(X)$:

$$\max_{B:B\cap A\neq\emptyset, A} \Big|\gamma(A\cap B) + \sum_{C\subseteq B, C\not\subset A} m_b(C)\Big|$$
$$= \max_{\emptyset\subsetneq X\subsetneq A} \max_{\emptyset\subseteq Y\subseteq A^c} \Big|\gamma(X) + \sum_{\emptyset\subsetneq Z\subseteq Y} \sum_{\emptyset\subseteq W\subseteq X} m_b(Z+W)\Big|$$

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:

$$\gamma^*(X) : \max_{\emptyset \subseteq Y \subseteq A^c} \left| \gamma^*(X) + \sum_{\emptyset \subsetneq Z \subseteq Y} \sum_{\emptyset \subseteq W \subseteq X} m_b(Z + W) \right| \leq 1 - b(A). \tag{47}$$

We must distinguish three cases.

1. If $\gamma^*(X) \geq 0$ we have that

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

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

$$\gamma^*(X) : \gamma^*(X) \leq 1 - b(A) - \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \tag{48}$$

   are optimal.

2. If $\gamma^*(X) < 0$ the maximum in (47) can be achieved for either $Y = A^c$ or $Y = \emptyset$. We are then left with the two corresponding terms in the max:

$$\gamma^*(X) : \max_{\emptyset \subseteq Y \subseteq A^c} \Big\{ \Big| \gamma^*(X) + \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \Big|, -\gamma^*(X) \Big\} \leq 1 - b(A). \tag{49}$$

   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

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

   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

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

which is equivalent to

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

in turn trivially true, since $\gamma^*(X) < 0$. Therefore, all

$$0 \geq \gamma^*(X) \geq -\frac{1}{2}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C) \tag{50}$$

are optimal as well.

3. In the last case, $\left|\gamma^*(X) + \sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C)\right| \leq -\gamma^*(X)$, 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

$$-\gamma^*(X) \leq 1 - b(A) \;\equiv\; \gamma^*(X) \geq b(A) - 1,$$

which is satisfied for all

$$b(A) - 1 \leq \gamma^*(X) \leq -\frac{1}{2}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C). \tag{51}$$

Putting (48), (50) and (51) together we have the thesis.

Proof of Theorem 8

Following Lemma 3 it is not difficult to see by induction that in the original auxiliary variables $\{\beta(X)\}$ the set of L_∞ conditional b.f.s in $\mathcal{B}$ is determined by the following constraints:

$$-K(X) + (-1)^{|X|}\sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C) \leq \beta(X) \leq K(X) - \sum_{C\cap A^c\neq\emptyset, C\cap A\subseteq X} m_b(C) \tag{52}$$

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.s (52) becomes

$$
\begin{array}{l}
m_b(X) - K(X) + \sum_{\emptyset \subsetneq B \subseteq A^c} m_b(X+B) \leq m_a(X) \leq \\
\qquad\qquad\qquad m_b(X) + K(X) - (-1)^{|X|} \sum_{B \subseteq A^c} m_b(B)
\end{array}
$$

which reads, after replacing the expression for $K(X)$,

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

By further trivial simplification we obtain the thesis.

Proof of Theorem 9

The proof is by substitution. In the $\{\beta(B)\}$ variables the thesis reads as:

$$
\beta(C) = \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \Big[(-1)^{|C|} m_b(B) - m_b(B \cup C) \Big]. \qquad (53)
$$

By replacing (53) in (36) we get, since $\sum_{\emptyset \subsetneq C \subseteq X} (-1)^{|C|} = 0 - (-1)^0 = -1$ by Newton's binomial,

$$
\begin{array}{rl}
& \sum_{\emptyset \subsetneq C \subseteq X} \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \Big[(-1)^{|C|} m_b(B) - m_b(B \cup C) \Big] + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \\
= & \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} m_b(B) \Big(\sum_{\emptyset \subsetneq C \subseteq X} (-1)^{|C|} \Big) - \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subsetneq C \subseteq X} m_b(B \cup C) + \\
& + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \\
= & -\frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} m_b(B) - \frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subsetneq C \subseteq X} m_b(B \cup C) + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \\
= & -\frac{1}{2} \sum_{\emptyset \subsetneq B \subseteq A^c} \sum_{\emptyset \subseteq C \subseteq X} m_b(B \cup C) + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) \\
= & -\frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) + \frac{1}{2} \sum_{C \cap A^c \neq \emptyset, C \cap A \subseteq X} m_b(C) = 0
\end{array}
$$

for all $\emptyset \subsetneq X \subsetneq A$, and system (36) is met.

References

[1] L. de Campos, J. Huete, S. Moral, Probability intervals: a tool for uncertain reasoning, IJUFKS 1 (1994) 167–196.

[2] I. Levi, The enterprise of knowledge, MIT Press, 1980.

[3] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, New York, 1991.

[4] B. D. Finetti, Theory of Probability, Wiley, London, 1974.

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

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

[7] T. Strat, Decision analysis using belief functions, International Journal of Approximate Reasoning 4 (1990) 391–417.

[8] P. Smets, Decision making in the TBM: the necessity of the pignistic transformation, International Journal of Approximate Reasoning 38 (2) (February 2005) 133–147.

[9] M. Troffaes, Decision making under uncertainty using imprecise probabilities, International Journal of Approximate Reasoning 45 (1) (2007) 17–29.

[10] P. Kohli, P. Torr, Efficiently solving dynamic Markov random fields using graph cuts, in: Proceedings of ICCV'05, Vol. 2, 2005, pp. 922–929.

[11] F. Cuzzolin, Visions of a generalized probability theory, PhD dissertation, Università di Padova (19 February 2001).

[12] S. Moral, N. Wilson, Importance sampling monte-carlo algorithms for the calculation of dempster-shafer belief, in: Proc. of IPMU'96, 1996.

[13] H. Kyburg, Bayesian and non-Bayesian evidential updating, Artificial Intelligence 31 (3) (1987) 271–294.

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

[15] R. Fagin, J. Halpern, A new approach to updating beliefs, in: Proc. of UAI, 1991, pp. 347–374.

[16] J. Y. Jaffray, Bayesian updating and belief functions, IEEE Transactions on Systems, Man and Cybernetics 22 (1992) 1144–1152.

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

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

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

[20] M. Itoh, 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.

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

[22] M. Spies, Conditional events, conditioning, and random sets, IEEE Transactions on Systems, Man, and Cybernetics 24 (1994) 1755–1763.

[23] A. Slobodova, Conditional belief functions and valuation-based systems, Tech. rep., Slovak Academy of Sciences (1994).

[24] A. Slobodova, Multivalued extension of conditional belief functions, in: Qualitative and quantitative practical reasoning, Vol. 1244/1997, Springer, pp. 568–573.

[25] M. Klopotek, S. Wierzchon, An interpretation for the conditional belief function in the theory of evidence, in: LNCS, Vol. 1609/1999, Springer, 1999, pp. 494–502.

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

[27] F. Klawonn, P. Smets, The dynamic of belief in the transferable belief model and specialization-generalization matrices, in: Proceedings of UAI'92, pp. 130–137.

[28] P. Smets, Jeffrey's rule of conditioning generalized to belief functions, in: Proceedings of UAI'93, pp. 500–505.

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

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

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

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

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

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

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

[36] F. Cuzzolin, Geometry of Dempster's rule of combination, IEEE Transactions on Systems, Man and Cybernetics part B 34 (2) (2004) 961–977.

[37] F. Cuzzolin, A geometric approach to the theory of evidence, IEEE Transactions on Systems, Man and Cybernetics part C 38 (4) (2008) 522–534.

[38] F. Cuzzolin, Geometric conditioning of belief functions, in: Proceedings of BELIEF'10, Brest, France, 2010.

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

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

[41] W. Jiang, A. Zhang, Q. Yang, A new method to determine evidence discounting coefficient, in: LNCS, Vol. 5226/2008, 2008, pp. 882–887.

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

[43] J. Diaz, M. Rifqi, B. Bouchon-Meunier, A similarity measure between basic belief assignments, in: Proceedings of FUSION'06.

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

[45] A. Perea, A model of minimal probabilistic belief revision, Theory and Decision 67 (2) (2009) 163–222.

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

[47] Y. Bar-Shalom, T. E. Fortmann, Tracking and Data Association, Academic Press, Inc., 1988.

[48] E. Bloem, H. Blom, Joint probabilistic data association methods avoiding track coalescence, in: Proceedings of CDC'95.

[49] M. Isard, A. Blake, Contour tracking by stochastic propagation of conditional density, in: Proceedings of ECCV'96, 1996, pp. 343–356.

[50] P. Smets, R. Kennes, The Transferable Belief Model, Artificial Intelligence 66 (1994) 191–234.

[51] F. Cuzzolin, R. Frezza, Geometric analysis of belief space and conditional subspaces, in: Proceedings of ISIPTA'01, Cornell University, June 2001.

[52] F. Cuzzolin, Two new Bayesian approximations of belief functions based on convex geometry, IEEE Transactions on Systems, Man, and Cybernetics - Part B 37 (4) (2007) 993–1008.

[53] F. Cuzzolin, Lp consonant approximations of belief functions, in: submitted to ECSQARU'11, Belfast, UK, 2011.

[54] F. Cuzzolin, Lp consonant approximations of belief functions in the mass space, in: submitted to ISIPTA'11, Innsbruck, Austria, 2011.

[55] P. Smets, The axiomatic justification of the transferable belief model, Tech. rep., TR/IRIDIA/1995-8.1, Universite' Libre de Bruxelles (1995).

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

[57] F. Cuzzolin, Consistent approximation of belief functions, in: Proceedings of ISIPTA'09, Durham, UK, June 2009.