L_p consonant approximations of belief functions

Fabio Cuzzolin

Abstract

In this paper we solve the problem of approximating a belief measure with a necessity measure or "consonant belief function" in a geometric framework. Consonant belief functions form a simplicial complex in both the space of all belief functions and the space of all mass vectors: partial approximations are first sought in each component of the complex, while global solutions are selected among them. As a first step in this line of study, we seek here approximations which minimize L_p norms. Approximations in the mass space can be interpreted in terms of mass redistribution, while approximations in the belief space generalize the maximal outer consonant approximation. We compare them with each other and with other classical approximations, and illustrate them with the help of a running example.

Index Terms

Theory of evidence, possibility theory, consonant belief functions, geometric approach, simplicial complex, (outer) consonant approximation, isopignistic function, L_p norms.

I. Introduction

The theory of evidence [1] is a popular approach to uncertainty description in which probabilities are replaced by *belief functions* (b.f.s), functions $b : 2^\Theta \to [0,1]$ on the power set $2^\Theta = \{A \subseteq \Theta\}$ of the sample space Θ of the form $b(A) = \sum_{B\subseteq A} m_b(B)$, where $m_b : 2^\Theta \to [0,1]$ is a non-negative, normalized set function called "basic probability assignment" (b.p.a.) or "mass assignment". Subsets of Θ associated with non-zero values of m_b are called *focal elements*. Possibility theory [2], instead, studies *possibility measures*, i.e., functions $Pos : 2^\Theta \to [0,1]$ on

F. Cuzzolin is with the Department of Computing and Communication Technologies, Oxford Brookes University, Oxford, UK. Email: Fabio.Cuzzolin@brookes.ac.uk.

the power set such that $Pos(\emptyset) = 0$, $Pos(\Theta) = 1$, and $Pos(\bigcup_i A_i) = \max_i Pos(A_i)$ for any family of subsets $\{A_i | A_i \in 2^\Theta\}$. Given a possibility measure Pos, the dual *necessity* measure is defined as $Nec(A) = 1 - Pos(A^c)$. Necessity measures have counterparts in the theory of evidence in the form of *consonant* belief functions (co.b.f.s), i.e., b.f.s whose non-zero mass subsets $m_b(A) \neq 0$ or "focal elements" (f.e.s) are nested [1] and form a chain (totally ordered collection) of subsets $A_1 \subset \cdots \subset A_m$, $A_i \subseteq \Theta$.

Reducing the complexity of belief calculus: an argument often raised against using belief functions in practice is their relatively high computational complexity, when compared to methods based on classical probability theory. To overcome these computational limitations, different approximation methods have been proposed. Some are based on Monte-Carlo techniques [3], [4], [5], while others seek to restrict the number of focal elements [6], [7], often by mapping belief functions to probability measures (*Bayesian approximation* or *probability transformation* [8], [9], [10], [11]), as the latter have a number of focal elements which is linear in the size of the frame of discernment.

As possibilities are completely determined by their values on the singletons ($Pos(\{x\})$, $x \in \Theta$), they are also less computationally expensive than belief functions, making a "possibility approximation" process attractive for many applications. Approximating a belief function with a possibility/necessity measure amounts, as we pointed out, to mapping it to a consonant b.f. [12], [13], [14], [15]. However, as explained by Dubois et al [16], possibility and probability do not capture the same facets of uncertainty: while probability theory offers a good quantitative model for randomness and undecisiveness, possibility theory better models partial ignorance. Bayesian and consonant approximation focus therefore on different aspects of the original belief function, while allowing us both to reduce its complexity. For this reason, possible mappings between possibilities and probabilities have also been investigated in the past [16], [17].

This *consonant approximation* problem has been studied by relatively few researchers: in [13] a "focused consistent transformation" of a random set was sought which minimized the information loss caused by the transformation. On their side, Dubois and Prade have developed the notion of "outer consonant approximation", which has received considerable attention in the past. Indeed, belief functions admit the following order relation: $b \leq b' \equiv \forall A \subseteq \Theta, b(A) \leq b'(A)$, called "weak inclusion". It is then possible to define the outer consonant approximations [12] of a belief function b as those co.b.f.s co such that $co(A) \leq b(A)\ \forall A \subseteq \Theta$. Dubois and

Prade's work has been later extended by Baroni [15] to capacities, while the author of this paper has provided a comprehensive description of the geometry of the set of outer consonant approximations [18]. Particularly interesting is, for each possible maximal chain $A_0 \subset \cdots \subset A_{|\Theta|}$, $|A_i| = i$ of focal elements, the maximal outer consonant approximation with mass assignment: $m'(A_i) = b(A_i) - b(A_{i-1})$, which mirrors the behavior of the vertices of the credal set of probabilities dominating a belief function or a 2-alternating capacity [19], [20].

Another interesting approximation emerges in the framework of Smets' Transferable Belief Model [21], where the "pignistic" probability $BetP(x) = \sum_{A \supseteq x} \frac{m_b(A)}{|A|}$ has a central role for decision making. The notion of an "isopignistic" approximation as the unique consonant belief function whose pignistic probability is identical to that of the original b.f. b can then be defined [22], [16]. The expression of the isopignistic consonant b.f. associated with a unimodal probability density has been derived in [23]. In [24], instead, consonant belief functions are constructed from sample data using confidence sets of pignistic probabilities.

A geometric approach to approximation: in more recent times the opportunity of seeking probability or consonant approximations / transformations of belief functions by minimizing appropriate distance functions has been explored. The author has himself introduced the notion of orthogonal projection $\pi[b]$ of a belief function onto the probability simplex [25], and studied consistent approximations of belief functions induced by classical L_p norms [26] in the space of belief functions [27]. In [28] he has shown that norm minimization can also be used to define families of geometric conditional belief functions. As to what distances are the most appropriate, Jousselme et al [29] have recently conducted a nice survey of the distance or similarity measures so far introduced in belief calculus, come out with an interesting classification, and proposed a number of generalizations of known measures. Other similarity measures between belief functions have been proposed by Shi et al [30], Jiang et al [31], and others [32], [33]. Many of these measures can be in principle employed to define conditional belief functions, or to approximate belief functions by necessity or probability measures. As the author has recently proven [18], geometrically, consonant belief functions live in a collection of simplices or "simplicial complex". Each maximal simplex of the consonant complex $\mathcal{CO}$ is associated with a maximal chain of nested non-empty (as in this paper we only consider normalized belief functions, for which $m_b(\emptyset) = 0$) focal elements: $\mathcal{C} = \{A_1 \subset A_2 \subset \cdots \subset A_{|\Theta|} = \Theta\}$. Computing the consonant belief function(s) at minimal distance from a given b.f. b involves therefore: 1) computing first a partial solution for each possible maximal chain; 2) selecting a global approximation among all the partial ones. Geometric approximation, however, can be performed in different Cartesian spaces. Indeed, a belief function can be represented either by the vector of its belief values, or the vector of its mass values. We call the set of vectors of the first kind *belief space* $\mathcal{B}$ [27], [34], and the collection of vectors of the second kind *mass space* $\mathcal{M}$ [28]. In both cases consonant b.f.s belong to a simplicial complex.

Contribution: the goal of this paper is to conduct an exhaustive, analytical study of all the consonant approximations of belief functions induced by minimizing L_1, L_2 or L_∞ distances between the consonant complex and the original belief function, in both the belief and the mass space. Even though we believe the resulting consonant approximations are likely to be useful in practical applications, our purpose at this stage is not to empirically compare them with existing approaches such as isopignistic function and outer approximations, but to initiate a theoretical study of the nature of consonant approximations induced by geometric distance minimization, starting with L_p norms as a stepping stone of a more extensive line of research. Our purpose is to point out their semantics in terms of degrees of belief and their mutual relationships, and to analytically compare them with the existing approximations. What emerges is a picture in which belief-, mass-, and pignistic-based approximations form distinct families of approximations with different semantics. As it turns out, partial approximations *in the mass space* amount to *redistributing in various ways the mass of focal elements outside the desired maximal chain to elements of the chain itself* (compare [28]). The global approximations in the L_1, L_2, L_∞ cases span the simplicial components of $\mathcal{CO}$ whose chains minimize the sum of mass, sum of square masses, and maximal mass outside the desired maximal chain, respectively. In the *belief space*, all partial L_p approximations can be considered as generalizations of the classical maximal outer approximation $m'(A_i) = b(A_i) - b(A_{i-1})$. As for the global approximations, in the L_∞ case they fall on the component(s) associated with the maximal plausibility singleton(s). In the other two cases they are, for now, of more difficult interpretation.

Limitations: in some cases, *improper* partial solutions (potentially including negative mass assignments) may be obtained: the set of approximations may fall partly outside the simplex of proper consonant belief functions, for a given desired chain of focal elements. This situation is not new, as outer approximations themselves include infinitely many improper solutions, while only the subset of acceptable solutions is retained. In the case of the present work, the set of all (admissible and not) partial solutions is typically much simpler to describe geometrically, in terms of simplices or polytopes. Computing the set of *proper* approximations in all cases requires significant further effort, which for reasons of clarity and length we reserve for the near future. However, conditions under which such partial solutions are admissible are here given. Additionally, in this paper only "normalized" belief functions, i.e., b.f.s whose mass of the empty set is nil, are considered. Unnormalized b.f.s, however, play an important role in the TBM [35] as the mass of the empty set is an indicator of conflicting evidence. The analysis of the unnormalized case is also left to future work for lack of sufficient space here.

Paper outline: we first provide the necessary background on consonant belief functions and consonant approximations (Section II), in particular on the geometric representation of belief and mass vectors (II-A) and the geometric approach to the approximation problem (II-B). We first tackle the problem in the mass space in Section III, where we: analytically compute the approximations induced by L_1, L_2 and L_∞ norms (III-A); discuss their interpretation in terms of mass re-assignment (III-B); analyze the computability and admissibility of global approximations (III-C); study the relation of the obtained approximations with classical outer consonant approximations (III-D); illustrate the results in the significant ternary case (III-E); and finally, analyze their relationships with the isopignistic approximation (III-F). In the second part of the paper we analyze the L_p approximation problem in the belief space (Section IV). Again, we compute the approximations induced by L_1 (IV-A), L_2 (IV-B) and L_∞ (IV-C) norms, respectively; propose a comprehensive view of all the approximations in $\mathcal{B}$ via lists of belief values induced by the desired maximal chain (Section IV-D); illustrate them with the help of the usual ternary example (IV-E), and draw some conclusions on the behavior of geometric approximations in the belief and in the mass space (IV-F). All proofs are collected in an Appendix.

II. Geometry of consonant belief functions

A dual mathematical representation of the evidence encoded by a belief function b is the *plausibility function* (pl.f.) $pl_b : 2^\Theta \rightarrow [0,1]$, 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. A probability function is simply a special belief function assigning non-zero masses to singletons only (*Bayesian* b.f.): $m_b(A) = 0$ if $|A| > 1$. A belief function is said, instead, to be *consonant* if its focal elements are nested, and form a totally ordered chain $A_1 \subset \cdots \subset A_m$.

Consonant b.f.s are characterized by the fact that $pl_b(A) = \max_{x\in A} pl_b(x)$ for all non-empty $A \subseteq \Theta$; the restriction $\{pl_b(x), x \in \Theta\}$ of the plausibility function to singletons only is called the *contour function*. They constitute the link between the theory of belief functions and possibility theory [2]. Each possibility measure is uniquely characterized by a *membership function* or *possibility distribution* $\pi : \Theta \rightarrow [0,1]$ s.t. $\pi(x) \doteq Pos(\{x\})$ via the formula (in the finite case) $Pos(A) = \max_{x\in A} \pi(x)$. From the fact that for consonant b.f.s $pl_b(A) = \max_{x\in A} pl_b(x)$ for all A, it follows that [1] the plausibility function pl_b associated with a b.f. b is a possibility measure (i.e., b is a necessity measure) iff b is consonant, in which case $\pi(x) = pl_b(x)$.

A. Geometric representation of uncertainty measures

1) Belief space representation: given a frame Θ, each belief function $b : 2^\Theta \rightarrow [0,1]$ is completely specified by its $N-2$ belief values $\{b(A), \emptyset \subsetneq A \subsetneq \Theta\}$, $N \doteq 2^n$ ($n \doteq |\Theta|$), (as $b(\emptyset) = 0$, $b(\Theta) = 1$ for all b.f.s), and can therefore be represented as a point of $\mathbb{R}^{N-2}$. Once introduced a set of coordinate axes $\{\vec{v}_A, \emptyset \subsetneq A \subsetneq \Theta\}$ in $\mathbb{R}^{N-2}$, a belief function b can be represented by the vector $\vec{b} = \sum_{\emptyset \subsetneq A \subsetneq \Theta} b(A)\vec{v}_A$. If we denote by b_A the *categorical* [21] belief function (also called "unanimity game" [36]) assigning all the mass to a single subset $A \subseteq \Theta$, we can prove that [27], [34] 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 ($Cl(\vec{b}_1, ..., \vec{b}_k) = \{\vec{b} \in \mathcal{B} : \vec{b} = \alpha_1\vec{b}_1 + \cdots + \alpha_k\vec{b}_k, \sum_i \alpha_i = 1,\ \alpha_i \geq 0\ \forall i\}$) of all the vectors representing categorical belief functions: $\mathcal{B} = Cl(\vec{b}_A, \emptyset \subsetneq A \subseteq \Theta)$. The belief space $\mathcal{B}$ is a simplex[1] [27], and each vector $\vec{b} \in \mathcal{B}$ representing a belief function b can be written as a convex sum as (see also [36]): $\vec{b} = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A)\vec{b}_A$.

2) Mass space representation: in the same way, each belief function is uniquely associated with the related set of mass values $\{m_b(A), \emptyset \subsetneq A \subseteq \Theta\}$ (Θ this time included). It can therefore be seen also as a point of $\mathbb{R}^{N-1}$, the vector $\vec{m}_b$ of its $N-1$ mass components: $\vec{m}_b = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A)\vec{m}_A$, where $\vec{m}_A$ is the vector of mass values of the categorical b.f. b_A: $\vec{m}_A(A) = 1$, $\vec{m}_A(B) = 0\ \forall B \neq A$. Note that in $\mathbb{R}^{N-1}$ $\vec{m}_\Theta = [0, ..., 0, 1]'$ and cannot be neglected. The collection $\mathcal{M} = Cl(\vec{m}_A, \emptyset \subsetneq A \subseteq \Theta)$ of all the mass vectors in the Cartesian space $\mathbb{R}^{N-1}$ is a simplex with $N-1$ vertices $\{\vec{m}_A, \emptyset \subsetneq A \subseteq \Theta\}$ and of dimension $N-2$.

[1] An n-dimensional *simplex* is the convex closure $Cl(x_1, ..., x_{n+1})$ of $n+1$ affinely independent [27] points $x_1, ..., x_{n+1}$ of the Euclidean space $\mathbb{R}^n$. The *faces* of an n-dimensional simplex are all the possible simplices generated by a subset of its vertices, i.e., $Cl(x_{j_1}, ..., x_{j_k})$ with $\{j_1, ..., j_k\} \subset \{1, ..., n+1\}$.

3) Binary example: in the case of a frame of discernment containing only two elements, $\Theta_2 = \{x, y\}$, each b.f. $b : 2^{\Theta_2} \rightarrow [0, 1]$ is completely determined by its belief values[2] $b(x) = m_b(x)$ and $b(y) = m_b(y)$. We can therefore collect them in a vector of $\mathbb{R}^{N-2} = \mathbb{R}^2$ (since $N = 2^2 = 4$), $\vec{b} = [m_b(x) = b(x), m_b(y) = b(y)]' \in \mathbb{R}^2$. Since $m_b(x) \geq 0$, $m_b(y) \geq 0$, and $m_b(x) + m_b(y) \leq 1$ we can easily infer that the set $\mathcal{B}_2$ of all the possible belief functions on Θ_2 can be depicted as the triangle in the Cartesian plane of Figure 1-left, whose vertices are the points $\vec{b}_\Theta = [0, 0]'$, $\vec{b}_x = [1, 0]'$, $\vec{b}_y = [0, 1]'$, which correspond respectively to the vacuous belief function b_Θ ($m_{b_\Theta}(\Theta) = 1$), the Bayesian b.f. b_x with $m_{b_x}(x) = 1$, and the Bayesian b.f. b_y with $m_{b_y}(y) = 1$.

As for the mass space, since on $\Theta_2 = \{x, y\}$ we can represent mass functions as vectors $[m_b(x), m_b(y), m_b(\Theta)]'$ of $\mathbb{R}^3$, $\mathcal{M}_2$ is a 2-dimensional simplex in $\mathbb{R}^3$ (see Figure 1-right).

On $\Theta_2 = \{x, y\}$ consonant b.f.s can have as chain of focal elements either $\{\{x\} \subset \Theta_2\}$ or $\{\{y\} \subset \Theta_2\}$. Therefore, they live in the union of two segments (see Figure 1): $\mathcal{CO}_2 = \mathcal{CO}^{\{x,\Theta\}} \cup \mathcal{CO}^{\{y,\Theta\}} = Cl(\vec{m}_x, \vec{m}_\Theta) \cup Cl(\vec{m}_y, \vec{m}_\Theta)$.

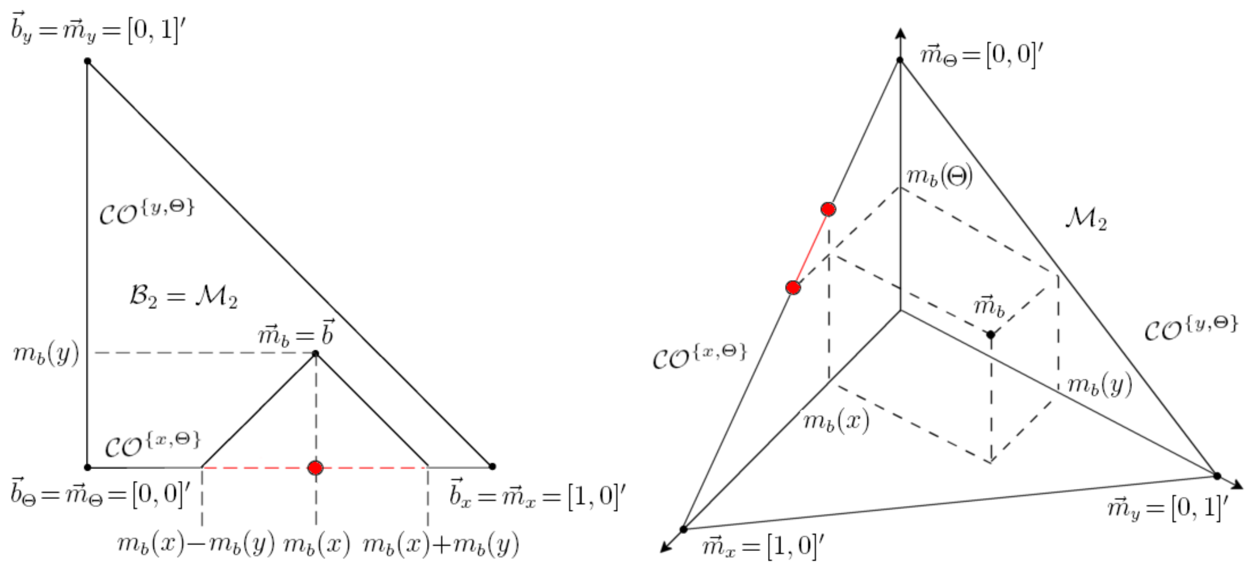

Fig. 1. Left: the belief space $\mathcal{B}_2$ for a binary frame $\Theta = \{x, y\}$ is a triangle in $\mathbb{R}^2$ whose vertices are the belief vectors $\vec{b}_x, \vec{b}_y, \vec{b}_\Theta$ associated with the categorical belief functions focused on $\{x\}, \{y\}$ and Θ, respectively. The unique L_1 consonant approximation (red circle) and the set of L_∞ consonant approximations (dashed segment) on $\mathcal{CO}^{\{x,\Theta\}}$ are shown. Right: the mass space $\mathcal{M}_2$ for the same binary frame is instead a 2-dimensional triangle embedded in $\mathbb{R}^3$, whose vertices are the mass vectors $\vec{m}_x, \vec{m}_y, \vec{m}_\Theta$. Consonant b.f.s live in the union of the segments $\mathcal{CO}^{\{x,\Theta\}} = Cl(\vec{m}_x, \vec{m}_\Theta)$ and $\mathcal{CO}^{\{y,\Theta\}} = Cl(\vec{m}_y, \vec{m}_\Theta)$. Using the L_1 norm for approximation yields in this case a whole segment of solutions (in red).

[2]We use the notation $m_b(x), b(x), pl_b(x)$ for the values of set functions on singletons, instead of $m_b(\{x\}), b(\{x\}), pl_b(\{x\})$

B. The consonant approximation problem

1) Approximation in the consonant complex: the geometry of consonant belief functions in the general case can be described through the notion of "simplicial complex" [37]. A *simplicial complex* is a collection Σ of simplices of arbitrary dimensions such that: 1) if a simplex belongs to Σ, then all its faces of any dimension belong to Σ; 2) the intersection of any two simplices is a face of both. It can be proven that [18] the region $\mathcal{CO}_\mathcal{B}$ of consonant belief functions in the belief space is a simplicial complex, the union of a collection of (maximal) simplices, each of them associated with a maximal chain $\mathcal{C} = \{A_1 \subset \cdots \subset A_n\}$, $|A_i| = i$ of non-empty subsets of the frame Θ: $\mathcal{CO}_\mathcal{B} = \bigcup_{\mathcal{C}=\{A_1\subset\cdots\subset A_n\}} Cl(\vec{b}_{A_1}, \cdots, \vec{b}_{A_n})$. Analogously, the region $\mathcal{CO}_\mathcal{M}$ of consonant belief functions in the mass space $\mathcal{M}$ is the simplicial complex: $\mathcal{CO}_\mathcal{M} = \bigcup_{\mathcal{C}=\{A_1\subset\cdots\subset A_n\}} Cl(\vec{m}_{A_1}, \cdots, \vec{m}_{A_n})$.

Given a belief function b, we call *consonant approximation of* b *induced by a distance function* d in $\mathcal{M}/\mathcal{B}$ the b.f.(s) $co_{\mathcal{M}/\mathcal{B},d}[m_b/b]$ which minimize(s) the distance $d(\vec{m}_b, \mathcal{CO}_\mathcal{M})/d(\vec{b}, \mathcal{CO}_\mathcal{B})$ between $\vec{m}_b/\vec{b}$ and the consonant simplicial complex in $\mathcal{M}/\mathcal{B}$:

$$co_{\mathcal{M},d}[m_b] = \arg\min_{\vec{m}_{co}\in\mathcal{CO}_\mathcal{M}} d(\vec{m}_b, \vec{m}_{co}) \Big/ co_{\mathcal{B},d}[b] = \arg\min_{\vec{co}\in\mathcal{CO}_\mathcal{B}} d(\vec{b}, \vec{co}). \quad (1)$$

2) Choice of norm: consonant b.f.s are the counterparts of necessity measures in the theory of evidence, so that their plausibility functions are possibility measures. These, in turn, are related to the L_∞ (max) norm via $Pos(A) = \max_{x\in A} Pos(x)$. It is then sensible to conjecture that a consonant transformation obtained by picking as distance function d in (1) the L_∞ norm would be meaningful. Indeed, the latter (often in the space of log-likelihoods) has been continually rediscovered and extensively used in probabilistic graphical models as well, under the names of "dynamic range" [38] and "L_∞ quotient metric" [39], among others.

In the context of the approximation problem, L_p norms in general have been successfully employed to design novel transformations: for instance, the L_2 probability transformation induces the so-called "orthogonal projection" of b onto $\mathcal{P}$ [25]. The use of L_p norms to define conditional belief functions has also been brought forward [40], [28]. In the belief space, the L_p distances between two vectors of belief values $\vec{b}$ and $\vec{b}'$ are, respectively: $\|\vec{b}-\vec{b}'\|_{L_1} \doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |b(B)-b'(B)|$; $\|\vec{b}-\vec{b}'\|_{L_2} \doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta}(b(B)-b'(B))^2}$ and $\|\vec{b}-\vec{b}'\|_{L_\infty} \doteq \max_{\emptyset\subsetneq B\subseteq\Theta} |b(B)-b'(B)|$.

In the mass space, instead, they read as: $\|\vec{m}_b - \vec{m}_{b'}\|_{L_1} \doteq \sum_{\emptyset\subsetneq B\subseteq\Theta} |m_b(B) - m_{b'}(B)|$; $\|\vec{m}_b -$

$\vec{m}_{b'}\|_{L_2} \doteq \sqrt{\sum_{\emptyset\subsetneq B\subseteq\Theta}(m_b(B)-m_{b'}(B))^2}$ and $\|\vec{m}_b-\vec{m}_{b'}\|_{L_\infty} \doteq \max_{\emptyset\subsetneq B\subseteq\Theta}|m_b(B)-m_{b'}(B)|$, where $\vec{m}_b$, $\vec{m}_{b'} \in \mathcal{M}$ are vectors representing the b.p.a.s of two belief functions b, b'.

Clearly, a number of other norms can be introduced in the framework of belief functions and used to define consonant (or Bayesian) approximations. For instance, generalizations to belief functions of the classical Kullback-Leibler divergence between probability distributions or other measures based on information theory such as fidelity and entropy-based norms [41] can be studied. Many other similarity measures have indeed been proposed [30], [31], [32], [33]. The application to the approximation problem of similarity measures more specific to belief functions or inspired by classical probability is a huge task, of which this paper is just a first step.

3) Distance of a point from a simplicial complex: as the consonant complex $\mathcal{CO}$ is a *collection* of simplices which generate distinct linear spaces (in both the belief and the mass space), solving the consonant approximation problem involves finding first a number of partial solutions:

$$co^{\mathcal{C}}_{\mathcal{M},L_p}[m_b] = \arg\min_{\vec{m}_{co}\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}} \|\vec{m}_b-\vec{m}_{co}\|_{L_p} \Big/ co^{\mathcal{C}}_{\mathcal{B},L_p}[b] = \arg\min_{\vec{co}\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{B}}} \|\vec{b}-\vec{co}\|_{L_p}, \tag{2}$$

one for each maximal chain $\mathcal{C}$ of subsets of Θ. Then, the distance of b from all such partial solutions has to be assessed in order to select a global optimal approximation.

4) Moebius inversion and preservation of norms, induced orderings: given a belief function b, the corresponding basic probability assignment m_b can be obtained via *Moebius inversion* as: $m_b(A) = \sum_{B\subseteq A}(-1)^{|A\setminus B|}b(B)$. More in general, a Moebius inverse exists for any sum function $f(A) = \sum_{B\leq A} g(B)$ defined on a partially ordered set with ordering $\leq$, and is the combinatorial analogue of the derivative operator in calculus [42].

If a norm d existed for belief functions that was preserved by Moebius inversion, $d(\vec{b},\vec{b}') = d(\vec{m}_b,\vec{m}_{b'})$, then the approximation problems (2) in $\mathcal{B}$ and $\mathcal{M}$ would obviously yield the same result(s). The same would be true if Moebius inversion preserved the *ordering* induced by d:

$$d(\vec{b},\vec{b}_1) \leq d(\vec{b},\vec{b}_2) \Leftrightarrow d(\vec{m},\vec{m}_1) \leq d(\vec{m},\vec{m}_2) \quad \forall \vec{b},\vec{b}_1,\vec{b}_2 \in \mathcal{B}.$$

Unfortunately, this is not the case at least for any of the above L_p norms. Let us consider again the binary example of Section II-A, and measure the distance between the categorical belief function $\vec{b} = \vec{b}_y$ (such that $m_b(y) = 1$) and the segment $Cl(\vec{b}_x, \vec{b}_\Theta)$ of consonant b.f.s with chain of focal elements $\{x\} \subset \Theta$. When using the L_2 distance in the belief space:

$$\|\vec{b}_y - \vec{b}_\Theta\|_{L_2} = \|[0,1,1]' - [0,0,1]'\| = 1 < \|\vec{b}_y - \vec{b}_x\|_{L_2} = \|[0,1,1]' - [1,0,1]'\| = \sqrt{2},$$

and $\vec{b}_\Theta$ is closer to $\vec{b}_y$ than $\vec{b}_x$. In the mass space embedded in $\mathbb{R}^3$, instead:

$$\|\vec{m}_y - \vec{m}_\Theta\|_{L_2} = \|[0,1,0]' - [0,0,1]'\| = \sqrt{2} = \|\vec{m}_y - \vec{m}_x\|_{L_2} = \|[0,1,0]' - [1,0,0]'\| = \sqrt{2},$$

while $\|\vec{m}_y - (\vec{m}_x + \vec{m}_\Theta)/2\|_{L_2} = \sqrt{3}/2 < \sqrt{2}$. The L_2 partial consonant approximation in the first case is $\vec{b}_\Theta$, in the second $(\vec{m}_x + \vec{m}_\Theta)/2$. Similar results can be shown for L_1 and L_∞.
As a consequence, separate approximation problems (2) have to be set up in the belief and mass space, respectively. Indeed, an interesting question is whether there actually is a norm whose induced ordering is preserved by Moebius inversion. This is an extremely challenging open problem which, to the best of our knowledge, has not been studied so far and cannot be quickly addressed here, but we intend to tackle, among others, in the near future.

III. Consonant approximation in the mass space

Let us then compute first the analytical form of all L_p consonant approximations in the mass space. The mass vector associated with an arbitrary consonant b.f. co with maximal chain of non-empty focal elements $\mathcal{C}$ reads as $\vec{m}_{co} = \sum_{A\in\mathcal{C}} m_{co}(A)\vec{m}_A$, so that the difference vector is:

$$\vec{m}_b - \vec{m}_{co} = \sum_{A\in\mathcal{C}} \big(m_b(A) - m_{co}(A)\big)\vec{m}_A + \sum_{B\not\in\mathcal{C}} m_b(B)\vec{m}_B. \qquad (3)$$

We denote by $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_p}[m_b]$ (uppercase) the set of partial L_p approximations of b with maximal chain $\mathcal{C}$ calculated in the mass space $\mathcal{M}$. We drop the superscript $\mathcal{C}$ for global solutions, and use $co^{\mathcal{C}}_{\mathcal{M},L_p}[m_b]$ (lowercase) for pointwise solutions and the barycenters of sets of solutions.

A. Results of L_p consonant approximation in the mass space

Theorem 1: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, the partial L_1 consonant approximations of b with maximal chain of focal elements $\mathcal{C}$, calculated in the mass space $\mathcal{M}$, is the set of co.b.f.s co with chain $\mathcal{C}$ such that $m_{co}(A) \geq m_b(A)\ \forall A \in \mathcal{C}$. They form a simplex: $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_1}[m_b] = Cl\big(\vec{m}^{\bar{A}}_{L_1}[m_b], \bar{A}\in\mathcal{C}\big)$, whose vertices have b.p.a.:

$$\vec{m}^{\bar{A}}_{L_1}[m_b](A) = m_b(A),\ A\in\mathcal{C}, A\neq\bar{A}; \quad \vec{m}^{\bar{A}}_{L_1}[m_b](\bar{A}) = m_b(A) + \sum_{B\not\in\mathcal{C}} m_b(B), \qquad (4)$$

and whose barycenter has mass assignment:

$$co^{\mathcal{C}}_{\mathcal{M},L_1}[m_b](A) = m_b(A) + \frac{1}{n}\sum_{B\not\in\mathcal{C}} m_b(B) \quad \forall A\in\mathcal{C}. \qquad (5)$$

The set of global L_1 approximations of b in $\mathcal{M}$ is the union of the simplices of partial solutions associated with the maximal chain(s) which maximize(s) their own total original mass:

$$\mathcal{CO}_{\mathcal{M},L_1}[m_b] = \bigcup_{C \in \arg\max_{\mathcal{C}} \sum_{A\in\mathcal{C}} m_b(A)} \mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_1}[m_b].$$

In order to find the L_2 consonant approximation(s) in $\mathcal{M}$, instead, it is convenient to recall that the minimal L_2 distance between a point and a vector space is attained by the point of the vector space V such that the difference vector is orthogonal to all the generators $\vec{g}_i$ of V: $\arg\min_{\vec{q}\in V} \|\vec{p}-\vec{q}\|_{L_2} = \hat{q} \in V : \langle \vec{p}-\hat{q}, \vec{g}_i\rangle = 0 \ \forall i$ whenever $\vec{p} \in \mathbb{R}^m$, $V = span(\{\vec{g}_i, i\})$. Instead of minimizing the L_2 norm of the difference vector $\|\vec{m}_b - \vec{m}_{co}\|_{L_2}$ we impose a condition of orthogonality between the difference vector itself $\vec{m}_b - \vec{m}_{co}$ and each component $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ of the consonant complex in the mass space.

Theorem 2: Given a belief function $b : 2^\Theta \to [0,1]$ with b.p.a. m_b, the partial L_2 consonant approximation of b with maximal chain of focal elements $\mathcal{C}$, calculated in the mass space $\mathcal{M}$, has mass assignment (5): $co^{\mathcal{C}}_{\mathcal{M},L_2}[m_b] = co^{\mathcal{C}}_{\mathcal{M},L_1}[m_b]$. The set of all global such L_2 approximations is the union of the partial solutions associated with maximal chains of focal elements which minimize the sum of square masses outside the chain:

$$\mathcal{CO}_{\mathcal{M},L_2}[m_b] = \bigcup_{\mathcal{C} \in \arg\min_{\mathcal{C}} \sum_{B\notin\mathcal{C}} (m_b(B))^2} co^{\mathcal{C}}_{\mathcal{M},L_2}[m_b].$$

L_1 and L_2 global solutions fall in general onto different simplicial components of $\mathcal{CO}_{\mathcal{M}}$.

Theorem 3: Given a belief function $b : 2^\Theta \to [0,1]$ with b.p.a. m_b, the partial L_∞ consonant approximations of b with maximal chain of focal elements $\mathcal{C}$, calculated in the mass space $\mathcal{M}$, can form either a simplex:

$$\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_\infty}[m_b] = Cl\big(\vec{m}^{\bar{A}}_{L_\infty}[m_b], \bar{A} \in \mathcal{C}\big) \tag{6}$$

whose vertices have b.p.a.:

$$\vec{m}^{\bar{A}}_{L_\infty}[m_b](A) = \begin{cases} m_b(A) + \max\limits_{B\notin\mathcal{C}} m_b(B) & A \in \mathcal{C}, A \neq \bar{A}, \\ m_b(\bar{A}) + \max\limits_{B\notin\mathcal{C}} m_b(B) + \Big(\sum\limits_{B\notin\mathcal{C}} m_b(B) - n \max\limits_{B\notin\mathcal{C}} m_b(B)\Big) & A = \bar{A}, \end{cases} \tag{7}$$

when the belief function to approximate is such that:

$$\max_{B\notin\mathcal{C}} m_b(B) \geq \frac{1}{n} \sum_{B\notin\mathcal{C}} m_b(B), \tag{8}$$

or reduce, when the opposite is true, to a single consonant belief function, the barycenter of the above simplex, located on the partial L_2 approximation (and barycenter of the L_1 partial approximations) (5). When (8) holds, the global such L_∞ consonant approximations are associated with the maximal chain(s) of focal elements $\arg\min_{\mathcal{C}} \max_{B\not\in\mathcal{C}} m_b(B)$; otherwise they correspond to the maximal chains $\arg\min_{\mathcal{C}} \sum_{B\not\in\mathcal{C}} m_b(B)$.

B. Semantics of partial consonant approximations in $\mathcal{M}$

Summarizing, the partial L_p approximations of a mass function m_b calculated in $\mathcal{M}$ are:

$$\begin{array}{ll} \mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_1} & [m_b] = Cl\big(\vec{m}^{\bar{A}}_{L_1}[m_b], \bar{A}\in\mathcal{C}\big) = \Big\{co\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}} : m_{co}(A)\geq m_b(A)\ \forall A\in\mathcal{C}\Big\}; \\ co^{\mathcal{C}}_{\mathcal{M},L_2} & [m_b] = co^{\mathcal{C}}_{\mathcal{M},L_1}[m_b] = co\in\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}\ :\ m_{co}(A) = m_b(A) + \frac{1}{n}\sum_{B\not\in\mathcal{C}} m_b(B); \\ \mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_\infty} & [m_b] = Cl(\vec{m}^{L_\infty}_{\bar{A}}, \bar{A}\in\mathcal{C}) \end{array} \quad (9)$$

if (8) holds, otherwise simply $co^{\mathcal{C}}_{\mathcal{M},L_\infty}[m_b] = co^{\mathcal{C}}_{\mathcal{M},L_2}[m_b] = co^{\mathcal{C}}_{\mathcal{M},L_1}[m_b]$. We can observe that, for each desired maximal chain of focal elements $\mathcal{C}$:

1) the L_1 partial approximations of b are those co.b.f.s whose basic probabilities (not beliefs) *dominates that of b over all the elements of the chain*: $m_{co}(A)\geq m_b(A)\ \forall A\in\mathcal{C}$;
2) this set is a fully admissible simplex, whose vertices are obtained by *re-assigning all the mass outside the desired chain to a single focal element of the chain* itself (see (4));
3) its barycenter coincides with the L_2 partial approximation with the same chain, which redistributes the original mass of focal elements outside the chain *to all the elements of the chain on an equal basis* (5);
4) when the partial L_∞ approximation is unique, it coincides with the L_2 approximation and the barycenter of the L_1 approximations;
5) otherwise, it is a simplex whose vertices assign to each element of the chain (but one) *the maximal mass outside the chain*, and whose barycenter is still the L_2 approximation.

Note that the simplex of L_∞ partial solutions (point 5)) may fall outside the simplex of consonant b.f.s with the same chain, therefore some of those approximations can be non-admissible.

As a general trait, approximations in the mass space amount to some redistribution of the original mass to focal elements of the desired maximal chain: their geometric relationship is depicted in Figure 2.

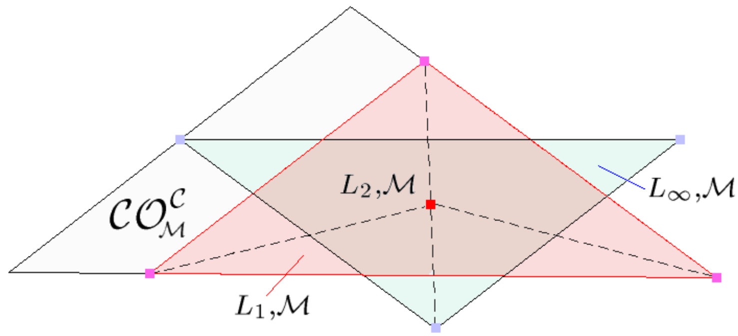

Fig. 2. Graphical representation of the relationships between the different (partial) L_p consonant approximations with desired maximal chain $\mathcal{C}$, in the related simplex $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ of the consonant complex $\mathcal{CO}$.

C. *Interpretation, computability and admissibility of global solutions*

As far as *global* solutions are concerned, we can observe the following facts:

- in the L_1 case, the optimal chain(s) are $\arg\min_{\mathcal{C}} \sum_{B\not\in\mathcal{C}} m_b(B) = \arg\max_{\mathcal{C}} \sum_{A\in\mathcal{C}} m_b(A)$;
- in the L_2 case, these are $\arg\min_{\mathcal{C}} \sum_{B\not\in\mathcal{C}} (m_b(B))^2$;
- in the L_∞ case, the optimal chain(s) are $\arg\min_{\mathcal{C}} \max_{B\not\in\mathcal{C}} m_b(B)$ unless the approximation is unique, in which case the optimal chains are as in the L_1 case.

While the L_2 global approximation is of difficult interpretation, both the L_1 and L_∞ solutions are supported by an intuitive rationale, as they are associated with the chains which minimize the total/maximal mass originally outside the desired maximal chain.

1) Admissibility of partial and global solutions: we know that all L_1/L_2 partial solutions are always admissible. As for the L_∞ case, not even global solutions are guaranteed to have all admissible vertices (7): indeed, $\Delta = \sum_{B\not\in\mathcal{C}} m_b(B) - n \cdot \max_{B\not\in\mathcal{C}} m_b(B) \leq 0$ as we are under condition (8), therefore $\vec{m}^{\bar{A}}_{L_\infty}[m_b](\bar{A})$ can be negative. The computation of the admissible part of this set of solutions is not trivial, and is left to future work.

2) Computational complexity of global solutions: in terms of computability, finding the global L_1/L_2 approximations involves therefore finding the maximal mass/square mass chain(s). This is expensive, as we have to examine all $n!$ of them. The most favorable case in terms of complexity is the L_∞ one (unless (8) does not hold), as all the chains not containing the maximal mass element(s) are optimal. Looking for the maximal mass focal elements requires a single pass of the list of focal elements, with complexity $O(2^n)$ rather than $O(n!)$. On the other hand, in this case the global consonant approximations are spread over a potentially large number of simplicial components of $\mathcal{CO}$, and are therefore less informative.

3) Comparison with generalized isopignistic, contour-based approximations: this behavior compares rather unfavorably with that of (the generalization of) two classical approximations.

Definition 1: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, its *isopignistic* consonant approximation [16] is defined as the unique consonant b.f. $co_{iso}[b]$ such that $BetP[co_{iso}[b]] = BetP[b]$, where $BetP[.]$ is the pignistic transform mapping each belief function b to the pignistic probability function: $BetP[b](x) = \sum_{A \supseteq x} \frac{m_b(A)}{|A|}$. Its contour function is:

$$pl_{co_{iso}[b]}(x) = \sum_{x' \in \Theta} \min \Big\{ BetP[b](x), BetP[b](x') \Big\}. \tag{10}$$

It is well known that, given the contour function pl_b of a *consistent* belief function $b : 2^\Theta \rightarrow [0,1]$ (a b.f. such that $\max_x pl_b(x) = 1$) we can obtain the unique consonant b.f. which has pl_b as contour function via the following formulae:

$$m_{co}(A_i) = pl_b(x_i) - pl_b(x_{i+1}) \quad \forall i = 1, ..., n-1; \quad m_{co}(A_n) = pl_b(x_n), \tag{11}$$

where $x_1, ..., x_n$ are the singletons of Θ sorted by plausibility value, and $A_i = \{x_1, ..., x_i\}$ for all $i = 1, ..., n$. Such a unique transformation is not in general feasible for *arbitrary* belief functions. The isopignistic transform builds a contour function (possibility distribution) from the pignistic values of the singletons. Given the list of singletons $x_1, ..., x_n$ ordered by pignistic value, (10) reads as: $pl_{co_{iso}[b]}(x_i) = 1 - \sum_{j=1}^{i-1} \big(BetP[b](x_j) - BetP[b](x_i) \big) = \sum_{j=i}^{n} BetP[b](x_j) + (i-1) BetP[b](x_i)$. By applying (11) we obtain $m_{co_{iso}[b]}(A_n) = n \cdot BetP[b](x_n)$, and:

$$m_{co_{iso}[b]}(A_i) = i \cdot \big(BetP[b](x_i) - BetP[b](x_{i+1}) \big), \quad i = 1, ..., n-1. \tag{12}$$

If we apply (12) to an arbitrary ordering of the singletons, we obtain what is in general not guaranteed to be a proper consonant belief function (as it can have negative masses): we call it the *generalized isopignistic function* of b.

The mapping (11) can be used to define another interesting approximation, as follows.

Definition 2: Given a b.f. $b : 2^\Theta \rightarrow [0,1]$, its *generalized contour-based* consonant approximation with maximal chain of non-empty f.e.s $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ has b.p.a.:

$$m_{co^{\mathcal{C}}_{con}[b]}(A_i) = \begin{cases} 1 - pl_b(x_2) & i = 1, \\ pl_b(x_i) - pl_b(x_{i+1}) & i = 2, ..., n-1, \\ pl_b(x_n) & i = n, \end{cases} \tag{13}$$

where $\{x_1\} = A_1$, $\{x_i\} \doteq A_i \setminus A_{i-1}$ for all $i = 2, ..., n$.

We call such approximation "generalized" as it uses the (unnormalized) contour function of an arbitrary b.f. b *as if* it was a possibility distribution, by replacing the plausibility of the maximal element with 1, and applies the mapping (11) to an arbitrary ordering of the singletons (instead of the one induced by plausibility), represented by an arbitrary chain of focal elements $\mathcal{C}$.

To be admissible, (13) requires sorting the plausibility values of the singletons (complexity $O(n \log n)$), while the isopignistic one requires $n \cdot (n-1)$ comparisons as we need to compare $BetP[b](x)$ with $BetP[b](x')$, $\forall\, x, x' \in \Theta$ via (10). One must add also the complexity of actually computing the value of $BetP[b](x)$ ($pl_b(x)$) from a mass vector, which requires n scans (one for each singleton x) with an overall complexity of $n \cdot 2^n$.

D. Relation with outer consonant approximations

Proposition 1: [18] Given a b.f. $b : 2^\Theta \rightarrow [0,1]$ with b.p.a. m_b, the set of partial outer consonant approximation of b with maximal chain of non-empty focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ is the convex closure $\mathcal{OC}^{\mathcal{C}}[b] = Cl(co^{\mathcal{C}}_{\vec{B}}[b], \forall \vec{B})$ of the co.b.f.s with b.p.a.:

$$m_{co^{\mathcal{C}}_{\vec{B}}[b]}(A_i) = \sum_{A \subseteq \Theta : \vec{B}(A) = A_i} m_b(A) \quad \forall i = 1, ..., n, \tag{14}$$

each associated with an "assignment function" $\vec{B} : 2^\Theta \rightarrow \mathcal{C}$, $A \mapsto \vec{B}(A) \supseteq A$ mapping each focal element A to one of the elements of the chain containing it. Note that the points (14) are not all guaranteed to be proper vertices of the polytope $\mathcal{OC}^{\mathcal{C}}[b]$, as some of them can be obtained via convex combinations of others.

The outer approximation produced by the permutation $\rho = \{x_{\rho(1)}, \cdots, x_{\rho(n)}\}$ of singletons of Θ which generates the desired maximal chain of focal elements via $A_i = \{x_{\rho(1)}, \cdots, x_{\rho(i)}\}$, i.e.,

$$m_{co^{\mathcal{C}}_{max}[b]}(A_i) = \sum_{B \subseteq A_i, B \not\subseteq A_{i-1}} m_b(B) = b(A_i) - b(A_{i-1}) \quad \forall i = 1, ..., n, \tag{15}$$

is an actual vertex of $\mathcal{OC}^{\mathcal{C}}[b]$, and the *maximal* outer consonant approximation of b with maximal chain $\mathcal{C}$. Indeed, an interesting relationship between outer consonant and L_1 consonant approximation in the mass space can be proven.

Theorem 4: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, the set of partial L_1 consonant approximations $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_1}[m_b]$ with maximal chain of focal elements $\mathcal{C}$, calculated in the mass space $\mathcal{M}$, and the set $\mathcal{OC}^{\mathcal{C}}[b]$ of its partial outer consonant approximations with the same chain have non-empty intersection. This intersection contains at least the convex closure of the candidate vertices (14) of $\mathcal{OC}^{\mathcal{C}}[b]$ whose assignment functions are such that $\vec{B}(A_i) = A_i$ for all $i = 1, ..., n$.

Proof. Clearly if $\vec{B}(A_i) = A_i$ for all $i = 1, ..., n$, then the mass $m_b(A_i)$ is re-assigned to A_i itself for each element A_i of the chain. Hence $m_{co}(A_i) \geq m_b(A_i)$, and the co.b.f. belongs to $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_1}[m_b]$ (see Equation (9)). □

In particular, $co^{\mathcal{C}}_{max}[b]$ (15) and the outer approximation which reassigns all the mass to Θ:

$$m_{co}(A) = m_b(A) \; \forall A \in \mathcal{C}, A \neq \Theta; \quad m_{co}(\Theta) = m_b(\Theta) + \sum_{B \not\in \mathcal{C}} m_b(B) \tag{16}$$

both belong to both (partial) outer and $L_1, \mathcal{M}$ consonant approximations. Approximation (16) is generated by the trivial assignment function: $\vec{B}(B) = \Theta$ for all $B \not\in \mathcal{C}$.

A negative result can, on the other hand, be proven for L_∞ approximations: given a belief function b and a maximal chain $\mathcal{C}$, (partial) outer consonant approximations $\mathcal{OC}^{\mathcal{C}}[b]$ and partial L_∞ approximations $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M},L_\infty}[m_b]$ in $\mathcal{M}$ are not guaranteed to have non-empty intersection. Let us rewrite the set of constraints for L_∞ approximations in $\mathcal{M}$ under condition (8) as:

$$\begin{cases} m_{co}(A) - m_b(A) \leq \max_{B \not\in \mathcal{C}} m_b(B) & A \in \mathcal{C}, A \neq \Theta, \\ \sum_{A \in \mathcal{C}, A \neq \Theta} \big(m_{co}(A) - m_b(A) \big) \geq \big(\sum_{B \not\in \mathcal{C}} m_b(B) - \max_{B \not\in \mathcal{C}} m_b(B) \big). \end{cases} \tag{17}$$

Indeed, when (8) does not hold, $co^{\mathcal{C}}_{\mathcal{M},L_\infty}[m_b] = co^{\mathcal{C}}_{\mathcal{M},L_2}[m_b]$ which is in general outside $\mathcal{OC}^{\mathcal{C}}[b]$.

To be a pseudo vertex of the set of partial outer approximations, a co.b.f. co must be the result of re-assigning the mass of each focal element to an element of the chain which contains it. Imagine that all the focal elements not in the desired chain $\mathcal{C}$ have the same mass: $m_b(B) = const$ for all $B \not\in \mathcal{C}$. Only up to $n-1$ of them can be reassigned to elements of the chain different from Θ. As a matter of fact, if you reassigned n outside focal elements to such elements of the chain, in absence of mass redistribution internal to the chain, some element $A \in \mathcal{C}$ of the chain would surely violate the first constraint in (17), as it should receive mass from at least two outside f.e.s, yielding $m_{co}(A) - m_b(A) \geq 2 \max_{B \not\in \mathcal{C}} m_b(B) > \max_{B \not\in \mathcal{C}} m_b(B)$.
Indeed, this is true even if mass redistribution happens within the chain. Imagine that some mass $m_b(A)$, $A \in \mathcal{C}$ is reassigned to some other $A' \in \mathcal{C}$. By the first constraint in (17), this is allowed only if $m_b(A) \leq \max_{B \not\in \mathcal{C}} m_b(B)$. Therefore the mass of just one outside focal element can still be reassigned to A, while now none can be reassigned to A'. In both cases, since the number of elements outside the chain $m = 2^n - 1 - n$ is greater than n (unless $n \leq 2$) the second equation of (17) implies $(n-1) \max_{B \not\in \mathcal{C}} m_b(B) \geq (m-1) \max_{B \not\in \mathcal{C}} m_b(B)$ which cannot hold under (8).

In particular, $co^{\mathcal{C}}_{max}[b]$ is not necessarily an $L_\infty, \mathcal{M}$ approximation of b.

E. Ternary example

It can be useful to compare the different approximations in the toy case of a ternary frame, $\Theta = \{x, y, z\}$. Let the desired consonant approximation(s) have maximal chain $\mathcal{C} = \{\{x\} \subset \{x, y\} \subset \Theta\}$. Figure 3 illustrates the different partial L_p consonant approximations in $\mathcal{M}$ in the simplex of consonant belief functions with chain $\mathcal{C}$, for a b.f. b with masses:

$$m_b(x) = 0.2, \quad m_b(y) = 0.3, \quad m_b(x, z) = 0.5. \tag{18}$$

According to the formulae at page 8 of [43], the set of partial outer consonant approximations of (18) with chain $\{\{x\} \subset \{x, y\} \subset \Theta\}$ is the convex closure of the candidate vertices:

$$\begin{array}{lll lll} \vec{m}_{B_1/B_2} & = & [m_b(x), m_b(y), 1 - m_b(x) - m_b(y)]', & \vec{m}_{B_7/B_8} & = & [0, m_b(x), 1 - m_b(x)]', \\ \vec{m}_{B_3/B_4} & = & [m_b(x), 0, 1 - m_b(x)]', & \vec{m}_{B_9/B_{10}} & = & [0, m_b(y), 1 - m_b(y)]', \\ \vec{m}_{B_5/B_6} & = & [0, m_b(x) + m_b(y), 1 - m_b(x) - m_b(y)]', & \vec{m}_{B_{11}/B_{12}} & = & [0, 0, 1]', \end{array} \tag{19}$$

plotted in Figure 3 as empty squares. Note that, by Theorem 4, both $co^{\mathcal{C}}_{max}[b]$ (15) (yellow square) and (16) belong to the intersection of (partial) outer and $L_1, \mathcal{M}$ consonant approximations, and that the $L_2, \mathcal{M}$ partial approximation is not a (partial) outer consonant approximation.

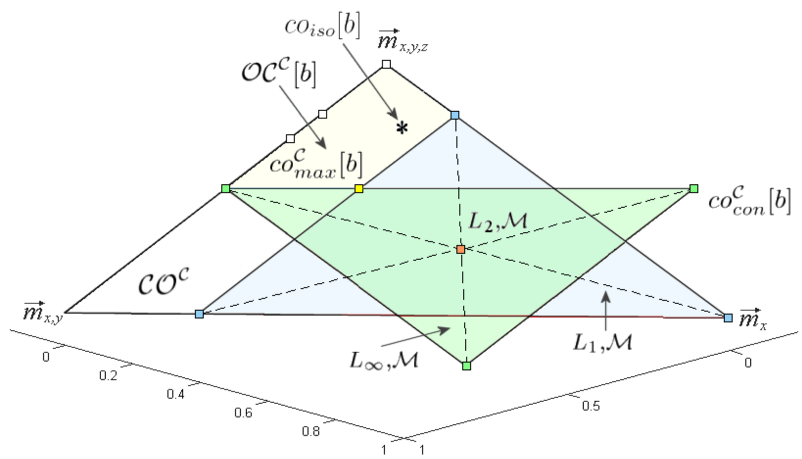

Fig. 3. The simplex $\mathcal{CO}^{\mathcal{C}}$ in the mass space of consonant belief functions with maximal chain $\mathcal{C} = \{\{x\} \subset \{x, y\} \subset \Theta\}$ defined on $\Theta = \{x, y, z\}$, and the L_p partial consonant approximations in $\mathcal{M}$ of the belief function with basic probabilities (18). The $L_2, \mathcal{M}$ approximation is plotted as a red square, as the barycenter of both the sets of $L_1, \mathcal{M}$ (blue triangle) and $L_\infty, \mathcal{M}$ (green triangle) approximations. The maximal outer approximation is denoted by a yellow square, the contour-based approximation is a vertex of the triangle $L_\infty, \mathcal{M}$. The related set $\mathcal{OC}^{\mathcal{C}}[b]$ of partial outer consonant approximations (19) is also shown for comparison (light yellow), while the isopignistic function is represented by a star.

As for the isopignistic and the contour-based approximations, they are in this case

$$\vec{m}_{iso} = [0.15, 0.1, 0.75]', \quad \vec{m}_{con} = [1 - pl_b(y), pl_b(y) - pl_b(z), pl_b(z)]' = [0.7, -0.2, 0.5]'.$$

The pignistic values of the elements in this example are $BetP[b](x) = 0.45$, $BetP[b](y) = 0.3$, $BetP[b](z) = 0.25$ so that the chain associated with the isopignistic approximation is indeed $\mathcal{C} = \{\{x\}, \{x,y\}, \Theta\}$. Notice, though, that generalized isopignistic approximations (i.e., pseudo consonant belief functions computed via (12) whose mass values, though, are not guaranteed to be positive) can be computed for all the chains via Equation (12). The generalized contour-based approximation $co^{\mathcal{C}}_{con}[b]$ is not admissible in this case, as the orderings of the singletons induced by plausibility and pignistic values are different.

F. Relation with contour and isopignistic approximations

It is worth to formally study the relationships of the $L_p, \mathcal{M}$ approximations with the contour-based and the isopignistic approximations as well.

1) Generalized contour-based approximation: From the (counter)-example of Figure 3 it follows that there exist belief functions for which the generalized contour-based approximation (13) is neither an outer consonant approximation ($co^{\mathcal{C}}_{con}[b] \notin \mathcal{OC}^{\mathcal{C}}[b]$) nor an $L_1, \mathcal{M}$ approximation, and that in general it is distinct from the unique $L_2, \mathcal{M}$ approximation.
On the other hand, Figure 3 seems to suggest that $co^{\mathcal{C}}_{con}[b]$ could be one of the vertices of the simplex of $L_\infty, \mathcal{M}$ approximations. Indeed, we can show that there exist belief functions for which the generalized contour-based approximation even falls outside this simplex.
The latter is determined by the system of constraints (17). On the other hand, by (13):

$$m_{co^{\mathcal{C}}_{con}[b]}(A_1) = \sum_{A \not\ni x_2} m_b(A) = m_b(A_1) + \sum_{\emptyset \subsetneq B \subseteq A_2^c} m_b(A_1 \cup B),$$

so that $m_{co^{\mathcal{C}}_{con}[b]}(A_1) - m_b(A_1) = \sum_{A \supsetneq \{x_1\}, A \not\ni x_2} m_b(A)$. Now, if b is such that $\arg\max_{B \in \mathcal{C}} m_b(B)$ also contains x_1 but not x_2, and there are other such subsets with non-zero mass, the first constraint in (17) is not met. Hence, $co^{\mathcal{C}}_{con}[b]$ is not a partial L_∞ approximation in $\mathcal{M}$.

2) Generalized isopignistic approximation: as for the (generalized) isopignistic approximation, the example shows that there are belief functions (such as (18)) for which such approximation (12) does not belong to either the set of $L_1, \mathcal{M}$ partial approximations nor the set of $L_\infty, \mathcal{M}$ partial approximations, and it is distinct from the unique $L_2, \mathcal{M}$ approximation, for any choice of the coordinate chart. On the other hand, the example suggests that $co_{iso}[b]$ could be always an outer consonant approximation. On the contrary, a simple counterexample shows that this is not so. Let $\Theta = \{x, y, z\}$ and b a belief function such that $BetP[b](x) \geq BetP[b](y) \geq BetP[b](z)$.

The isopignistic approximation will then have as chain of focal elements: $\{x\} \subset \{x, y\} \subset \Theta$. By (12) the isopignistic function has mass of $A_1 = \{x\}$ equal to:

$$m_{co_{iso}[b]}(A_1 = x) = BetP[b](x) - BetP[b](y) = m_b(x) - m_b(y) + \frac{m_b(x,z) - m_b(y,z)}{2}.$$

But if b is s.t. $m_b(y) = m_b(y,z) = 0$, $m_b(x,z) \neq 0$ we have that $co_{iso}[b](x) = m_{co_{iso}[b]}(x) = m_b(x) + \frac{m_b(x,z)}{2} > m_b(x) = b(x)$, i.e., the isopignistic is not an outer approximation.

IV. CONSONANT APPROXIMATION IN THE BELIEF SPACE

Consonant approximations in the mass space have quite natural semantics in terms of mass redistributions. As we see in this Section, instead, (partial) L_p approximations in the belief space are closely associated with lists of belief values determined by the desired maximal chain, and through them to the maximal outer approximation (15), as we will understand in Section IV-D.

A. L_1 approximation

Theorem 5: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, and a maximal chain of non-empty focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ in Θ, the partial L_1 consonant approximations $\mathcal{CO}^{\mathcal{C}}_{\mathcal{B},L_1}[b]$ in $\mathcal{B}$ with maximal chain $\mathcal{C}$ have mass vectors forming the convex closure:

$$Cl\Big(\big[b^1, b^2 - b^1, \cdots, b^i - b^{i-1}, \cdots, 1 - b^{n-1}\big]', b^i \in \big\{\lambda^i_{int1}, \lambda^i_{int2}\big\}\ \forall i = 1, ..., n-1\Big), \qquad (20)$$

where, $\forall\ i = 1, ..., n-1$, $\lambda^i_{int1}, \lambda^i_{int2}$ are the median elements of the list of belief values:

$$\mathcal{L}_i = \Big\{b(A), A \supseteq A_i, A \not\supseteq A_{i+1}\Big\}. \qquad (21)$$

In particular, $b^{n-1} = \lambda^{n-1}_{int1} = \lambda^{n-1}_{int2} = b(A_{n-1})$. As a result, (20) is a polytope with 2^{n-2} vertices.

Note that, even though the approximation is computed in $\mathcal{B}$, we present the result in terms of mass assignments as they are easier to interpret. The same holds for the other L_p approximations in $\mathcal{B}$. Due to the nature of partially ordered set of 2^Θ, the median values of the above lists (21) cannot be analytically identified in full generality (even though they can be easily computed numerically), but in some special cases (see Section IV-E).

By (20), the barycenter of the set of partial L_1 consonant approximations in $\mathcal{B}$ has mass vector:

$$m_{co^{\mathcal{C}}_{\mathcal{B},L_1}[b]} = \left[\frac{\lambda^1_{int1} + \lambda^1_{int2}}{2}, \frac{\lambda^2_{int1} + \lambda^2_{int2}}{2} - \frac{\lambda^1_{int1} + \lambda^1_{int2}}{2}, \cdots, 1 - b(A_{n-1})\right]'. \qquad (22)$$

Theorem 6: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, its global L_1 consonant approximations $\mathcal{CO}_{\mathcal{B},L_1}[b]$ in $\mathcal{B}$ live in the collection of partial such approximations associated with maximal chain(s) which maximize the cumulative lower halves of the $n-1$ lists of belief values $\mathcal{L}_i$ (21): $\arg\max_{\mathcal{C}} \sum_{i}^{n-1} \sum_{b(A)\in\mathcal{L}_i, b(A)\leq\lambda^i_{int_1}} b(A)$.

B. (Partial) L_2 approximation

To find the partial consonant approximation(s) at minimal L_2 distance from b in $\mathcal{B}$ we need to impose the orthogonality of the difference vector $\vec{b}-\vec{co}$ with respect to any given simplicial component $\mathcal{CO}^{\mathcal{C}}_{\mathcal{B}}$ of the complex $\mathcal{CO}_{\mathcal{B}}$:

$$\langle \vec{b}-\vec{co}, \vec{b}_{A_i}-\vec{b}_\Theta\rangle = \langle \vec{b}-\vec{co}, \vec{b}_{A_i}\rangle = 0 \quad \forall A_i \in \mathcal{C}, i=1,...,n-1, \tag{23}$$

as $\vec{b}_\Theta = \vec{0}$ is the origin of the Cartesian space in $\mathcal{B}$, and $\vec{b}_{A_i}-\vec{b}_\Theta$ for $i=1,...,n-1$ are the generators of the component $\mathcal{CO}^{\mathcal{C}}_{\mathcal{B}}$. Plugging the expression

$$\vec{b}-\vec{co} = \sum_{A\notin\mathcal{C}} m_b(A)\vec{b}_A + \sum_{A\in\mathcal{C},A\neq\Theta} \big(m_b(A)-m_{co}(A)\big)\vec{b}_A$$

of the difference vector in the base $\{\vec{b}_A\}$ (rather than the base $\{\vec{v}_A\}$) into the orthogonality condition (23) yields the following linear system of equations:

$$\left\{ \sum_{A\notin\mathcal{C}} m_b(A)\langle\vec{b}_A,\vec{b}_{A_i}\rangle + \sum_{A\in\mathcal{C},A\neq\Theta} \big(m_b(A)-m_{co}(A)\big)\langle\vec{b}_A,\vec{b}_{A_i}\rangle = 0, \quad \forall i=1,...,n-1. \right. \tag{24}$$

This is a linear system in $n-1$ unknowns $m_{co}(A_i)$, $i=1,...,n-1$ and $n-1$ equations.
Let us extend the definition of $\mathcal{L}_i$ by setting $\mathcal{L}_0 \doteq \{b(\emptyset)=0\}$, $\mathcal{L}_n \doteq \{b(\Theta)=1\}$: once again, the L_2 partial approximation of b is a function of the list of belief values (21).

Theorem 7: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, its partial L_2 consonant approximation $co^{\mathcal{C}}_{\mathcal{B},L_2}[b]$ in $\mathcal{B}$ with maximal chain $\mathcal{C}=\{A_1\subset\cdots\subset A_n, |A_i|=i\}$ is unique, and has b.p.a.:

$$m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i) = ave(\mathcal{L}_i) - ave(\mathcal{L}_{i-1}) \quad \forall i=1,...,n, \tag{25}$$

where $ave(\mathcal{L}_i)$ is the average of the list of belief values $\mathcal{L}_i$, which for (21) reads as:

$$ave(\mathcal{L}_i) = \frac{1}{2^{|A^c_{i+1}|}} \sum_{A\supseteq A_i, A\not\supseteq A_{i+1}} b(A) \quad \forall i=1,...,n-1. \tag{26}$$

The computation of the global L_2 approximation(s) is rather involved, and not addressed here.

C. L_∞ approximation

Partial L_∞ approximations in $\mathcal{B}$ also form a polytope, with this time 2^{n-1} vertices.

Theorem 8: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, its partial L_∞ consonant approximations in the belief space $\mathcal{CO}^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]$ with maximal chain of non-empty focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$ have mass vectors which live in the following convex closure of 2^{n-1} vertices:

$$Cl\left(\begin{array}{c|l} [b^1, b^2 - b^1, \cdots, b^i - b^{i-1}, \cdots, 1 - b^{n-1}]' & \forall i = 1, ..., n-1 \\ b^i \in \left\{ -b(A_1^c) + \dfrac{b(A_i) + b(\{x_{i+1}\}^c)}{2}, b(A_1^c) + \dfrac{b(A_i) + b(\{x_{i+1}\}^c)}{2} \right\} & \end{array} \right). \tag{27}$$

The barycenter $co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]$ of this set has mass assignment $m_{co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]}(A_n) = 1 - b(A_{n-1})$, and:

$$m_{co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]}(A_i) = \begin{cases} \dfrac{b(A_1) + b(\{x_2\}^c)}{2} & i = 1, \\ \dfrac{b(A_i) - b(A_{i-1})}{2} + \dfrac{pl_b(\{x_i\}) - pl_b(\{x_{i+1}\})}{2} & i = 2, ..., n-1, \end{cases} \tag{28}$$

Note that, since $b(A_1^c) = 1 - pl_b(A_1) = 1 - pl_b(x_1)$, the size of the polytope (27) of partial L_1 approximations of b is a function of the plausibility of the smallest focal element A_1 of the desired maximal chain only. As expected, it reduces to zero only when the b is consistent (the intersection of all its focal elements is non-empty [26]) and $A_1 = \{x_1\}$ has plausibility 1.

A straightforward interpretation of the barycenter of the partial L_∞ approximations in $\mathcal{B}$ in terms of degrees of belief is possible when we notice that, for all $i = 1, ..., n$

$$m_{co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]}(A_i) = (m_{co^{\mathcal{C}}_{max}[b]}(A_i) + m_{co^{\mathcal{C}}_{con}[b]}(A_i))/2$$

(recall Equations (15) and (13)), i.e., the barycenter is the average of the maximal outer consonant approximation and what we called "contour-based" consonant approximation.

To compute the *global* L_∞ approximation of the original belief function b in $\mathcal{B}$, we need to locate as usual the partial solution whose L_∞ distance from b is the smallest. Given the expression (43) of the L_∞ norm of the difference vector (see the proof of Theorem 8), such partial distance is (for each maximal chain $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$) equal to $b(A_1^c)$. Therefore the global L_∞ consonant approximations of b in the belief space are associated with the chains of focal elements: $\arg\min_{\mathcal{C}} b(A_1^c) = \arg\min_{\mathcal{C}}(1 - pl_b(A_1)) = \arg\max_{\mathcal{C}} pl_b(A_1)$.

Theorem 9: Given a belief function $b : 2^\Theta \rightarrow [0,1]$, the set of global L_∞ consonant approximations of b in the belief space is the collection of partial L_∞ approximations (27) associated with maximal chains whose non-empty smallest focal element is associated with the maximal plausibility singleton: $\mathcal{CO}_{\mathcal{B},L_\infty}[b] = \bigcup_{\mathcal{C}: A_1 = \{\arg\max_x pl_b(x)\}} \mathcal{CO}^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]$.

D. Approximations in $\mathcal{B}$ as generalized maximal outer approximations

From Theorems 5, 7 and 8, a comprehensive view of the results of this Section can be given in terms of the lists of belief values $\mathcal{L}_0 \doteq \{b(\emptyset) = 0\}$, $\mathcal{L}_i \doteq \Big\{b(A), A \supseteq A_i, A \not\supseteq A_{i+1}\Big\}$ $\forall i = 1, ..., n-1$, and $\mathcal{L}_n \doteq \{b(\Theta) = 1\}$. The b.p.a.s of *all the L_p partial approximations in the belief space are differences of simple functions of belief values taken from these lists* (which are uniquely determined by the desired chain of non-empty focal elements $A_1 \subset \cdots \subset A_n$), as

$$
\begin{array}{ll}
m_{co^{\mathcal{C}}_{max}[b]}(A_i) = \min(\mathcal{L}_i) - \min(\mathcal{L}_{i-1}); & m_{co^{\mathcal{C}}_{con}[b]}(A_i) = \max(\mathcal{L}_i) - \max(\mathcal{L}_{i-1}); \\
m_{co^{\mathcal{C}}_{\mathcal{B},L_1}[b]}(A_i) = \frac{int_1(\mathcal{L}_i)+int_2(\mathcal{L}_i)}{2} - \frac{int_1(\mathcal{L}_{i-1})+int_2(\mathcal{L}_{i-1})}{2}; & m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i) = ave(\mathcal{L}_i) - ave(\mathcal{L}_{i-1}); \\
m_{co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]}(A_i) = \frac{\max(\mathcal{L}_i)+\min(\mathcal{L}_i)}{2} - \frac{\max(\mathcal{L}_{i-1})+\min(\mathcal{L}_{i-1})}{2}, &
\end{array}
\tag{29}
$$

$\forall i = 1, ..., n$, where the b.p.a. of $co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]$ comes from (28). For each vertex of the L_1 polytope, either $int_1(\mathcal{L}_i)$ or $int_2(\mathcal{L}_i)$ is picked from the list $\mathcal{L}_i$ for each element A_i of the chain: $m_{co}(A_i) = int_1(\mathcal{L}_i)/int_2(\mathcal{L}_i) - int_1(\mathcal{L}_{i-1})/int_2(\mathcal{L}_{i-1})$. For each vertex of the L_∞ polytope, either $\max(\mathcal{L}_i)$ or $\min(\mathcal{L}_i)$ is picked: $m_{co}(A_i) = \max(\mathcal{L}_i)/\min(\mathcal{L}_i) - \max(\mathcal{L}_{i-1})/\min(\mathcal{L}_{i-1})$.

The different approximations in $\mathcal{B}$ (29) correspond therefore to different choices of a representative for the lists $\mathcal{L}_i$. The maximal outer approximation $co^{\mathcal{C}}_{max}[b]$ is obtained by picking as representative $\min(\mathcal{L}_i)$, $co^{\mathcal{C}}_{con}[b]$ amounts to picking $\max(\mathcal{L}_i)$, the barycenter of the L_1 approximations to choosing the average innermost (median) value, the barycenter of the L_∞ approximations to the average outermost value, the L_2 solution to picking the overall average value of the list. Each vertex of the L_1 solution set amounts to selecting, for each component, either one of the innermost values; each vertex of the L_∞ polytope, either one of the outermost values.

1) Interpretation of the list $\mathcal{L}_i$: belief functions are defined on a partially ordered set, the power set $2^\Theta = \{A \subseteq \Theta\}$, of which a maximal chain is a maximal totally ordered subset. Therefore, given two elements of the chain $A_i \subset A_{i+1}$, there is a number of "intermediate" focal elements A which contain the former but not the latter. If 2^Θ was to be a totally ordered set, the list $\mathcal{L}_i$ would contain a single element $b(A_i)$ and all the L_p approximations (29) would reduce to the maximal outer consonant approximation $co^{\mathcal{C}}_{max}[b]$, with b.p.a. $m_{co^{\mathcal{C}}_{max}[b]}(A_i) = b(A_i) - b(A_{i-1})$. The diversity of L_p approximations in $\mathcal{B}$ is therefore a consequence of belief functions being defined on partially ordered sets: together with the contour-based approximation (13), they can all be seen as member of a coherent family of approximations.

2) Admissibility: as it is clear from the table of Equation (29), the b.p.a.s of all the L_p approximations in the belief space are differences of vectors of all positive values; indeed, differences of shifted version of the same positive vector. As such vectors $\left[\frac{int_1(\mathcal{L}_i)+int_2(\mathcal{L}_i)}{2}, i = 1,...,n\right]'$, $\left[\frac{\max(\mathcal{L}_i)+\min(\mathcal{L}_i)}{2}, i = 1,...,n\right]'$, $[ave(\mathcal{L}_i), i = 1,...,n]'$ are not guaranteed to be monotonically increasing for any arbitrary maximal chain $\mathcal{C}$, none of the related partial approximations are guaranteed to be entirely admissible. However, sufficient conditions under which they are admissible can be worked out by studying the structure of the list of belief values (21). Let us first consider $co^{\mathcal{C}}_{max}$ and $co^{\mathcal{C}}_{con}$. As $\min(\mathcal{L}_{i-1}) = b(A_{i-1}) \leq b(A_i) = \min(\mathcal{L}_i)$ $\forall i = 2,...,n$, the maximal partial outer approximation is admissible for all maximal chains $\mathcal{C}$. As for the contour-based approximation, $\max(\mathcal{L}_i) = b(A_i \cup A^c_{i+1}) = b(x^c_{i+1}) = 1 - pl_b(x_{i+1})$ $\forall i = 1,...,n-1$, $\max(\mathcal{L}_{i-1}) = 1 - pl_b(x_i)$ $\forall i = 2,...,n$, so that $\max(\mathcal{L}_i) - \max(\mathcal{L}_{i-1}) = pl_b(x_i) - pl_b(x_{i+1})$ $\forall i = 2,...,n-1$. This difference is guaranteed non-negative *if the chain $\mathcal{C}$ is generated by singletons sorted by their plausibility values*. As a consequence, as

$$m_{co^{\mathcal{C}}_{\mathcal{B},L_\infty}[b]}(A_i) = \frac{\max(\mathcal{L}_i) - \max(\mathcal{L}_{i-1})}{2} + \frac{\min(\mathcal{L}_i) - \min(\mathcal{L}_{i-1})}{2},$$

the barycenter of the set of $L_\infty, \mathcal{B}$ approximations is also admissible on the same chain(s).

E. Graphical comparison in a ternary example

As we did in the mass space analysis, it can be helpful to visualize the different L_p consonant approximations in the belief space when $\Theta = \{x, y, z\}$, and compare them with the approximations in the mass space on the same example of Section III-E (Figure 4).

To obtain a homogeneous comparison, we plot both sets of approximations (in the belief and in the mass space) as vectors of mass values. When $\Theta = \{x, y, z\}$ and $A_1 = \{x\}$, $A_2 = \{x, y\}$, $A_3 = \{x, y, z\}$ the relevant lists of belief values are $\mathcal{L}_1 = \{b(x), b(x, z)\}$ and $\mathcal{L}_2 = \{b(x, y)\}$, so that $\min(\mathcal{L}_1) = int_1(\mathcal{L}_1) = b(x)$, $\max(\mathcal{L}_1) = int_2(\mathcal{L}_1) = b(x, z)$, $ave(\mathcal{L}_1) = \frac{b(x)+b(x,z)}{2}$; $\min(\mathcal{L}_2) = int_1(\mathcal{L}_2) = \max(\mathcal{L}_2) = int_2(\mathcal{L}_2) = ave(\mathcal{L}_2) = b(x, y)$. Therefore, the set of L_1 partial consonant approximations is, by Equation (20), the segment $Cl(\vec{m}^1_{L_1}, \vec{m}^2_{L_1})$ with vertices:

$$\vec{m}^1_{L_1} = \big[b(x), b(x, y) - b(x), 1 - b(x, y)\big]', \;\; \vec{m}^2_{L_1} = \big[b(x, z), b(x, y) - b(x, z), 1 - b(x, y)\big]' \quad (30)$$

(see Figure 4). Note that this set is not entirely admissible, not even in the ternary case. We also know that the maximal partial outer approximation (15) is not in general a vertex of the polygon

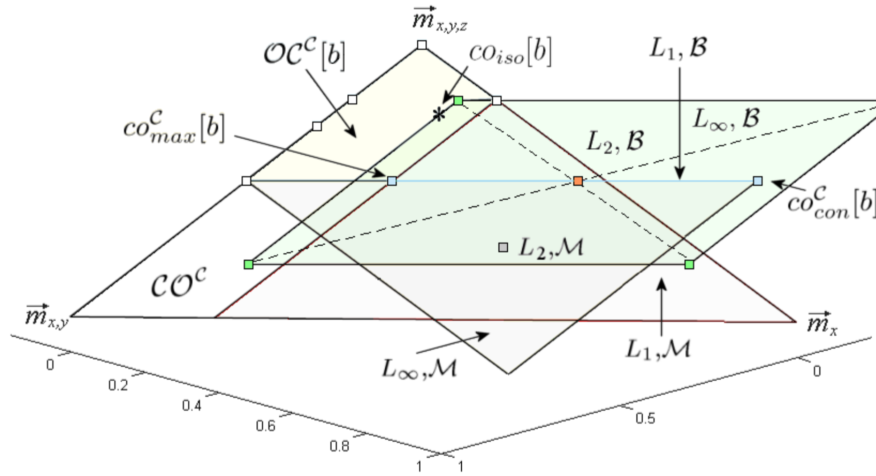

Fig. 4. Comparison between L_p partial consonant approximations in the mass $\mathcal{M}$ and belief $\mathcal{B}$ spaces for the belief function with basic probabilities (18) on $\Theta = \{x, y, z\}$. The $L_2, \mathcal{B}$ approximation is plotted as a red square, as the barycenter of both the sets of $L_1, \mathcal{B}$ (blue segment) and $L_\infty, \mathcal{B}$ (green quadrangle) approximations. Contour-based and maximal outer approximations are in this example the extreme of the segment $L_1, \mathcal{B}$ (blue squares). The polytope of partial outer consonant approximations (yellow), the isopignistic approximation (star) and the various L_p partial approximations in $\mathcal{M}$ (in gray levels) are also drawn.

of L_1 partial approximations in $\mathcal{B}$, unlike what the ternary example (for which $int_1(\mathcal{L}_1) = b(x)$) suggests. The partial L_2 approximation in $\mathcal{B}$ is, by Equation (29), unique, with mass vector:

$$\vec{m}_{co_{\mathcal{B},L_2}[b]} = \vec{m}_{co_{\mathcal{B},L_\infty}[b]} = \left[\frac{b(x) + b(x,z)}{2}, b(x,y) - \frac{b(x) + b(x,z)}{2}, 1 - b(x,y)\right]', \qquad (31)$$

and coincides with the barycenter of the set of partial L_∞ approximations (note that this is not so in the general case). As for the full set of partial L_∞ approximations, this has vertices (27):

$$\begin{array}{lll}
\vec{m}^1_{L_\infty} & = & \left[\frac{b(x)+b(x,z)}{2} - b(y,z), \quad b(x,y) - \frac{b(x)+b(x,z)}{2}, \quad 1 - b(x,y) + b(y,z)\right]'; \\
\vec{m}^2_{L_\infty} & = & \left[\frac{b(x)+b(x,z)}{2} - b(y,z), \quad b(x,y) - \frac{b(x)+b(x,z)}{2} + 2b(y,z), \quad 1 - b(x,y) - b(y,z)\right]'; \\
\vec{m}^3_{L_\infty} & = & \left[\frac{b(x)+b(x,z)}{2} + b(y,z), \quad b(x,y) - \frac{b(x)+b(x,z)}{2} - 2b(y,z), \quad 1 - b(x,y) + b(y,z)\right]'; \\
\vec{m}^4_{L_\infty} & = & \left[\frac{b(x)+b(x,z)}{2} + b(y,z), \quad b(x,y) - \frac{b(x)+b(x,z)}{2}, \quad 1 - b(x,y) - b(y,z)\right]',
\end{array}$$

which as expected are not all admissible (see Figure 4 again).

F. *Belief versus mass space approximations*

We can draw some conclusions by comparing the results of Section III and Section IV:

- L_p consonant approximations in the *mass* space are basically associated with different but related *mass redistribution* processes: the mass outside the desired chain of focal elements is re-assigned in some way to the elements of the chain;

- their relationships with classical outer approximations (on one hand) and approximations based on the pignistic transform (on the other) are rather weak;
- the different L_p approximations in $\mathcal{M}$ are characterized by natural geometric relations;
- consonant approximation in the *belief* space is inherently linked to the lists of belief values of focal elements "intermediate" between each pair of elements of the desired chain;
- the classical outer consonant approximations and the contour-based approximation are also approximations of the same type: indeed, the latter and the L_p approximations in the belief space can be seen as different generalizations of the maximal outer approximation, induced by the nature of partially ordered set of the power set;
- in the mass space, some partial approximations are always entirely admissible and should be preferred (this is the case for the L_1 and L_2 approximations in $\mathcal{M}$), some others are not;
- as for the belief case, even though all partial L_p approximations are differences between shifted versions of the same positive vector, admissibility is not guaranteed for all maximal chains; however, sufficient conditions exist.

Table (I) illustrates the behavior of the different geometric consonant approximations explored in this paper, in terms of multiplicity/admissibility/global solutions.

	multiplicity	**admissibility**	**global solution(s)**
$L_1, \mathcal{M}$	simplex	entirely	$\arg\min_{\mathcal{C}} \sum_{B \not\in \mathcal{C}} m_b(B)$
$L_2, \mathcal{M}$	point, barycenter of $L_1, \mathcal{M}$	yes	$\arg\min_{\mathcal{C}} \sum_{B \not\in \mathcal{C}} (m_b(B))^2$
$L_\infty, \mathcal{M}$	point / simplex	yes / not entirely	$\arg\min_{\mathcal{C}} \sum_{B \not\in \mathcal{C}} m_b(B)$ / $\arg\min_{\mathcal{C}} \max_{B \not\in \mathcal{C}} m_b(B)$
$L_1, \mathcal{B}$	polytope	not guaranteed	not easy to interpret
$L_2, \mathcal{B}$	point	not guaranteed	not known
$L_\infty, \mathcal{B}$	polytope	depending on plausibilities of singletons	$\arg\max_{\mathcal{C}} pl(A_1)$

TABLE I

PROPERTIES OF THE GEOMETRIC CONSONANT APPROXIMATIONS STUDIED HERE, IN TERMS OF MULTIPLICITY AND ADMISSIBILITY OF PARTIAL SOLUTIONS, AND THE RELATED GLOBAL SOLUTIONS.

1) On the links between approximations in $\mathcal{M}$ and $\mathcal{B}$: approximations in $\mathcal{B}$ and approximations in $\mathcal{M}$ do not coincide. This is a direct consequence of the fact that Moebius inversion does not preserve either L_p norms or the ordering induced by them, as it was clear from the counterexamples discussed in Section II-B.4. Though they are distinct, there can in principle still be links of some sort between L_p approximations in the two spaces. Let us consider, in particular, partial approximations. The ternary example of Figure 4 suggests the following conjectures:

1) the L_2 partial approximation in $\mathcal{B}$ is one of the L_1 partial approximations in $\mathcal{M}$;
2) the L_2 partial approximation in $\mathcal{B}$ is one of the L_∞ partial approximations in $\mathcal{M}$, and possibly belongs to the border of the simplex (6);
3) the L_1 partial approximation in $\mathcal{B}$ is also an element of (6).

Unfortunately, counterexamples can be provided in all these cases. Let us express $m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i)$ (25) as a function of the masses of the original b.f. b. For all $i = 1, ..., n-1$:

$$
\begin{array}{rcl}
ave(\mathcal{L}_i) & = & \displaystyle\frac{1}{2^{|A^c_{i+1}|}} \sum_{A\supseteq A_i, A\not\supseteq A_{i+1}} b(A) = \frac{1}{2^{|A^c_{i+1}|}} \sum_{A\supseteq A_i, A\not\supseteq A_{i+1}} \sum_{\emptyset\neq B\subseteq A} m_b(B) \\
& = & \displaystyle\frac{1}{2^{|A^c_{i+1}|}} \sum_{B\not\supseteq\{x_{i+1}\}} m_b(B) \cdot \Big|\{A = A_i \cup C : \emptyset \subseteq C \subseteq A^c_{i+1}, A \supseteq B\}\Big| \\
& = & \displaystyle\frac{1}{2^{|A^c_{i+1}|}} \sum_{B\not\supseteq\{x_{i+1}\}} m_b(B) \cdot 2^{|A^c_{i+1}\setminus(B\cap A^c_{i+1})|} = \sum_{B\not\supseteq\{x_{i+1}\}} \frac{m_b(B)}{2^{|B\cap A^c_{i+1}|}}
\end{array}
$$

so that, for all $i = 2, ..., n-1$: $m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i) = ave(\mathcal{L}_i) - ave(\mathcal{L}_{i-1}) =$

$$
\sum_{B\not\supseteq\{x_{i+1}\}} \frac{m_b(B)}{2^{|B\cap A^c_{i+1}|}} - \sum_{B\not\supseteq\{x_i\}} \frac{m_b(B)}{2^{|B\cap A^c_i|}} = \sum_{B\supseteq\{x_i\}, B\not\supseteq\{x_{i+1}\}} \frac{m_b(B)}{2^{|B\cap A^c_{i+1}|}} - \sum_{B\supseteq\{x_{i+1}\}, B\not\supseteq\{x_i\}} \frac{m_b(B)}{2^{|B\cap A^c_i|}}. \quad (32)
$$

Note that $m_b(A_i)$ is one of the terms of the first summation. Now, conjecture (1) requires the above mass to be greater than or equal to $m_b(A_i)$ for all i (Theorem 1): clearly though, by (32), the difference $m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i) - m_b(A_i)$ is not guaranteed to be positive.

As for conjecture 2), the set of $L_\infty, \mathcal{M}$ partial approximations is determined, once again, by the constraints (17). Now, suppose that b is such that $m_b(B) = 0$ for all $B \supseteq \{x_{i+1}\}$, $B \not\supseteq \{x_i\}$. Then $m_{co^{\mathcal{C}}_{\mathcal{B},L_2}[b]}(A_i) = \sum_{B\supseteq\{x_i\}, B\not\supseteq\{x_{i+1}\}} \frac{m_b(B)}{2^{|B\cap A^c_{i+1}|}}$ which contains, among other addenda, $\sum_{B\subseteq A_i, B\supseteq\{x_i\}} m_b(B)$ (for if $B \subseteq A_i$ we have $2^{|B\cap A^c_{i+1}|} = 2^{|\emptyset|} = 1$). Clearly, if $\arg\max_{B\not\in\mathcal{C}} m_b(B)$ is a proper subset of A_i containing $\{x_i\}$, and other subsets $B \supseteq \{x_i\}, B \not\supseteq \{x_{i+1}\}$ distinct from the latter and A_i have non-zero mass, the first constraint of (17) is not met.

Finally, consider conjecture 3). By Theorem 5, all the vertices of the set of $L_1, \mathcal{B}$ partial approximations have as mass $m_{co}(A_1)$ either one of the median elements of the list $\mathcal{L}_1 =$

$\{b(A) : A \supseteq A_1, \not\supseteq A_2\}$. These median elements are of the form $b(A_1 \cup C)$, for some $C \subseteq A_2^c$

$$b(A_1 \cup C) = m_b(A_1) + \sum_{B \subseteq A_1 \cup C, C \subseteq A_2^c, B \neq A_1} m_b(B),$$

so that $m_{co}(A_1) - m_b(A_1) = \sum_{B \subseteq A_1 \cup C, C \subseteq A_2^c, B \neq A_1} m_b(B)$.

Once again, if b is such that $\arg\max_{B \notin \mathcal{C}} m_b(A)$ is one of these subsets $B \subseteq A_1 \cup C$, $C \subseteq A_2^c$ and it is not the only one with non-zero mass, the first constrain in (17) is not met for $A = A_1$ by any of the vertices of the set of $L_1, \mathcal{B}$ approximations. Hence, the latter has empty intersection with the set of $L_\infty, \mathcal{M}$ partial approximations.

In conclusion, not only approximations in $\mathcal{M}$ and $\mathcal{B}$ are distinct, due to the properties of Moebius inversion, but they are not related in a straightforward way either.

2) Three families of consonant approximations: indeed, approximations in the mass and the belief space turn out to be inherently related to completely different philosophies to the consonant approximation problem: mass redistribution versus generalized maximal outer approximation. While mass space proxies correspond to different mass redistribution processes, L_p consonant approximation in the belief space amounts to generalizing in different but related ways the classical approach incarnated by the maximal outer approximation (15). The latter, together with the contour-based approximation (13) form therefore a different, coherent family of consonant approximations. As for the isopignistic approximation, it seems to be completely unrelated to approximations in both the mass and the belief space, as it naturally fits in the context of the Transferable Belief Model and the use of the pignistic function. Isopignistic, mass-space and belief-space consonant approximations form three distinct families of approximations, with fundamentally different rationales: which approach to use will therefore vary according to the chosen framework, and the problem at hand.

V. CONCLUSIONS

In this paper we studied all the consonant approximations of belief functions induced by minimizing L_p distances to the consonant complex, in both the mass space of basic probability vectors and the belief space of belief vectors. While interpretations for such approximations in the mass space are rather natural in terms of mass redistribution, approximations in the belief space are generalizations of the maximal outer and the contour-based approximations. We compared all these L_p approximations with each other and with other classical consonant approximations, and illustrated them with the help of a running example. The rationale for this work comes from the potential utility of possibility transforms as a tool to reduce the inherent exponential complexity of belief calculus. The analysis conducted here is just a first step in a wider programme of work, whose prosecution will likely involve the analysis of other types of distances between belief functions, the existence of norms preserved under Moebius inversion, and the empirical testing of these and other approximations in challenging, real-world setups.

APPENDIX

Proof of Theorem 1: the L_1 norm of the difference vector (3) is:

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

expressed as a function of the variables $\{\beta(A) \doteq m_b(A) - m_{co}(A), A \in \mathcal{C}, A \neq \Theta\}$. As

$$\sum_{A\in\mathcal{C}} \beta(A) = \sum_{A\in\mathcal{C}} \big(m_b(A) - m_{co}(A)\big) = \sum_{A\in\mathcal{C}} m_b(A) - 1 = -\sum_{B\notin\mathcal{C}} m_b(B),$$

we have that: $\beta(\Theta) = -\sum_{B\notin\mathcal{C}} m_b(B) - \sum_{A\in\mathcal{C}, A\neq\Theta} \beta(A)$. Therefore, the above norm reads as:

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_1} = \Bigg| -\sum_{B\notin\mathcal{C}} m_b(B) - \sum_{A\in\mathcal{C},A\neq\Theta} \beta(A) \Bigg| + \sum_{A\in\mathcal{C},A\neq\Theta} |\beta(A)| + \sum_{B\notin\mathcal{C}} m_b(B). \tag{33}$$

The norm (33) is a function of the form $\sum_i |x_i| + \Big| -\sum_i x_i - k \Big|$, $k \geq 0$, which has an entire simplex of minima, namely: $x_i \leq 0 \ \forall i$, $\sum_i x_i \geq -k$. The minima of the L_1 norm (33) are therefore the solutions to the following system of constraints:

$$\begin{cases} \beta(A) \leq 0 \ \forall A \in \mathcal{C}, A \neq \Theta; \quad \sum_{A\in\mathcal{C},A\neq\Theta} \beta(A) \geq -\sum_{B\notin\mathcal{C}} m_b(B). \end{cases} \tag{34}$$

This reads, in terms of the mass assignment m_{co} of the desired consonant approximation, as:

$$\begin{cases} m_{co}(A) \geq m_b(A) & \forall A \in \mathcal{C}, A \neq \Theta, \\ \sum_{A\in\mathcal{C},A\neq\Theta} \big(m_b(A) - m_{co}(A)\big) \geq -\sum_{B\notin\mathcal{C}} m_b(B). \end{cases} \tag{35}$$

Note that the last constraint reduces to

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

i.e., $m_{co}(\Theta) \geq m_b(\Theta)$. Therefore the partial L_1 approximations in $\mathcal{M}$ are those consonant b.f.s *co* s.t. $m_{co}(A) \geq m_b(A) \ \forall A \in \mathcal{C}$. The vertices of the set of partial approximations (34) are given by the vectors of variables $\{\vec{\beta}_{\bar{A}}, \bar{A} \in \mathcal{C}\}$ such that: $\vec{\beta}_{\bar{A}}(\bar{A}) = m_b(B)$, for $\vec{\beta}_{\bar{A}}(A) = 0$ for $A \neq \bar{A}$ whenever $\bar{A} \neq \Theta$, while $\vec{\beta}_\Theta = \vec{0}$. Immediately, in terms of masses the vertices of the set of partial L_1 approximations have b.p.a. (4) and barycenter (5). To find the *global* L_1 consonant approximation(s) over the whole consonant complex, we need to locate the component $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ at minimal L_1 distance from $\vec{m}_b$. All the partial approximations (35) onto $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ have L_1 distance from $\vec{m}_b$ equal to $2\sum_{B\not\in\mathcal{C}} m_b(B)$. Therefore, the minimal distance component(s) of the complex are those associated with maximal chains that originally have maximal mass with respect to m_b.

Proof of Theorem 2: we can pick as generators of $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ the vectors $\{\vec{m}_A - \vec{m}_\Theta, A \in \mathcal{C}, A \neq \Theta\}$. The orthogonality condition translates as: $\langle \vec{m}_b - \vec{m}_{co}, \vec{m}_A - \vec{m}_\Theta \rangle = 0$ for all $A \in \mathcal{C}$, $A \neq \Theta$. The vector $\vec{m}_A - \vec{m}_\Theta$ is such that: $\vec{m}_A - \vec{m}_\Theta(B) = 1$ if $B = A$, -1 if $B = \Theta$, 0 if $B \neq A, \Theta$. Hence, the orthogonality condition becomes $\beta(A) - \beta(\Theta) = 0$ for all $A \in \mathcal{C}, A \neq \Theta$, where again $\beta(A) = m_b(A) - m_{co}(A)$. As $\beta(\Theta) = -\sum_{B\not\in\mathcal{C}} m_b(B) - \sum_{A\in\mathcal{C},A\neq\Theta} \beta(A)$ (see the proof of Theorem 1), the orthogonality condition becomes $2\beta(A) + \sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq A,\Theta} \beta(B) = 0$ for all $A \in \mathcal{C}, A \neq \Theta$. Its solution is $\beta(A) = \frac{-\sum_{B\not\in\mathcal{C}} m_b(B)}{n}$ $\forall A \in \mathcal{C}$, $A \neq \Theta$, as by substitution $-\frac{2}{n}\sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\not\in\mathcal{C}} m_b(B) - \frac{n-2}{n}\sum_{B\not\in\mathcal{C}} m_b(B) = 0$, i.e., (5).

To find the global L_2 approximation(s), we need to compute the L_2 distance of $\vec{m}_b$ from the closest such partial solution. We have: $\|\vec{m}_b - \vec{m}_{co}\|^2_{L_2} = \sum_{A\subseteq\Theta}(m_b(A) - m_{co}(A))^2 =$

$$= \sum_{A\in\mathcal{C}} \left(\frac{\sum_{B\not\in\mathcal{C}} m_b(B)}{n} \right)^2 + \sum_{B\not\in\mathcal{C}} (m_b(B))^2 = \frac{\left(\sum_{B\not\in\mathcal{C}} m_b(B)\right)^2}{n} + \sum_{B\not\in\mathcal{C}} (m_b(B))^2,$$

which is minimized by the component $\mathcal{CO}^{\mathcal{C}}_{\mathcal{M}}$ that minimizes $\sum_{B\not\in\mathcal{C}}(m_b(B))^2$.

Proof of Theorem 3: the L_∞ norm of the difference vector is equal to: $\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} = \max\left\{ \max_{A\in\mathcal{C}} |\beta(A)|, \max_{B\not\in\mathcal{C}} m_b(B) \right\}$. As $\beta(\Theta) = \sum_{B\in\mathcal{C}} m_b(B) - 1 - \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B)$, we have that $|\beta(\Theta)| = |\sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B)|$ and the norm to minimize becomes:

$$\|\vec{m}_b - \vec{m}_{co}\|_{L_\infty} = \max\left\{ \max_{A\in\mathcal{C},A\neq\Theta} |\beta(A)|, \left| \sum_{B\not\in\mathcal{C}} m_b(B) + \sum_{B\in\mathcal{C},B\neq\Theta} \beta(B) \right|, \max_{B\not\in\mathcal{C}} m_b(B) \right\}. \quad (36)$$

This is a function of the form:

$$\max\left\{ |x_1|, |x_2|, |x_1 + x_2 + k_1|, k_2 \right\}, \quad 0 \leq k_2 \leq k_1 \leq 1. \quad (37)$$

Such a function has two possible behaviors in terms of its minimal points in the plane x_1, x_2.

Case 1. If $k_1 \leq 3k_2$ its contour function has the form rendered in Figure 5-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 general case of an arbitrary number

$m-1$ of variables $x_1, ..., x_{m-1}$ such that $x_i \geq -k_2$, $\sum_i x_i \leq k_2 - k_1$, the set of minimal points is a simplex with m vertices: each vertex v^i is such that $v^i(j) = -k_2 \; \forall j \neq i$; $v^i(i) = -k_1 + (m-1)k_2$ (obviously $v^m = [-k_2, \cdots, -k_2]$). For (36), in the first case $\Big(\max_{B \not\in \mathcal{C}} m_b(B) \geq \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B) \Big)$ the set of partial L_∞ approximations is given by the following system of inequalities:

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

This determines a simplex of solutions $Cl(\vec{m}^{\bar{A}}_{L_\infty}[m_b], \bar{A} \in \mathcal{C})$ with vertices:

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

or, in terms of their basic probability assignments, (7). Its barycenter is given by:

$$\frac{1}{n} \sum_{\bar{A} \in \mathcal{C}} \vec{m}^{\bar{A}}_{L_\infty}[m_b](A) = \frac{1}{n} \Big(n \cdot m_b(A) + \sum_{B \not\in \mathcal{C}} m_b(B) \Big) = m_b(A) + \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B),$$

i.e., the L_2 partial approximation (5). The corresponding minimal L_∞ norm of the difference vector is, according to (36), equal to $\max_{B \not\in \mathcal{C}} m_b(B)$.

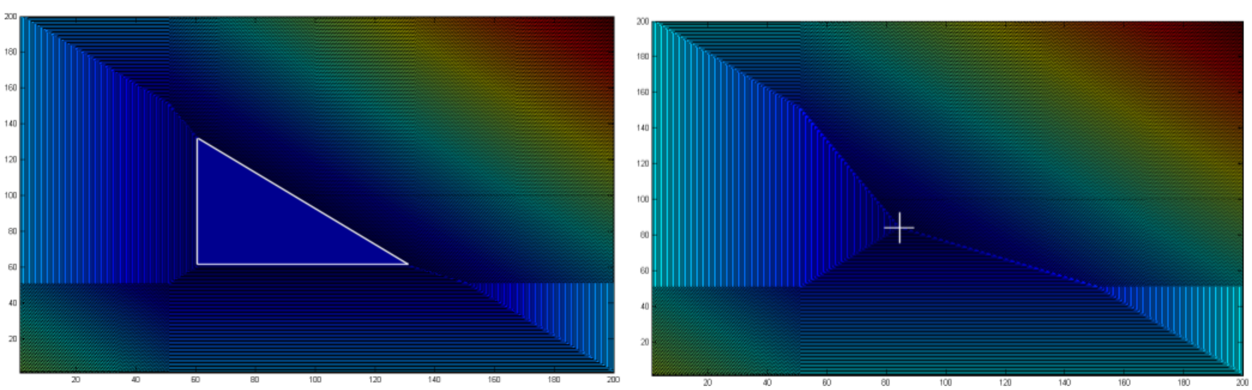

Fig. 5. Left: contour function (level sets) and minimal points (white triangle) of a function of the form (37), when $k_1 \leq 3k_2$. In the example $k_2 = 0.4$ and $k_1 = 0.5$. Right: contour function and minimal point of a function of the form (37), when $k_1 \geq 3k_2$. In this example $k_2 = 0.1$ and $k_1 = 0.5$.

Case 2. In the second case $k_1 > 3k_2$, i.e., for the norm (36), $\max_{B \not\in \mathcal{C}} m_b(B) < \frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B)$, the contour function of (37) is as in Figure 5-right. There is a single minimal point, located in $[-1/3k_1, -1/3k_1]$. For an arbitrary number $m-1$ of variables the minimal point is $[(-1/m)k_1, \cdots, (-1/m)k_1]'$, i.e., for system (36), $\beta(A) = -\frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B)$ for all $A \in \mathcal{C}, A \neq \Theta$ or, in terms of basic probability assignments, (5) (the mass of Θ is obtained by normalization). The corresponding minimal L_∞ norm of the difference vector is $\frac{1}{n} \sum_{B \not\in \mathcal{C}} m_b(B)$.

Lemma 1: Given a belief function $b : 2^\Theta \rightarrow [0,1]$ and an arbitrary consonant b.f. co defined on the same frame with maximal chain of non-empty focal elements $\mathcal{C} = \{A_1 \subset \cdots \subset A_n, |A_i| = i\}$, the difference between the corresponding vectors in the belief space is:

$$\vec{b} - \vec{co} = \sum_{A \not\supset A_1} b(A)\vec{v}_A + \sum_{i=1}^{n-1} \sum_{A \supseteq A_i, A \not\supset A_{i+1}} \vec{v}_A \Big[\gamma(A_i) + b(A) - \sum_{j=1}^{i} m_b(A_j) \Big], \tag{38}$$

where $\gamma(A) = \sum_{B \subseteq A, B \in \mathcal{C}} (m_b(B) - m_{co}(B))$.

Proof of Lemma 1: in the belief space the original b.f. and the desired consonant approximation are written as $\vec{b} = \sum_{\emptyset \subsetneq A \subsetneq \Theta} b(A)\vec{v}_A$, $\vec{co} = \sum_{A \supseteq A_1} \Big(\sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big) \vec{v}_A$. Their difference vector is:

$$\begin{aligned} \vec{b} - \vec{co} &= \sum_{A \not\supseteq A_1} b(A)\vec{v}_A + \sum_{A \supseteq A_1} \vec{v}_A \Big[b(A) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] \\ &= \sum_{A \not\supseteq A_1} b(A)\vec{v}_A + \sum_{A \supseteq A_1} \vec{v}_A \Big[\sum_{\emptyset \subsetneq B \subseteq A} m_b(B) - \sum_{B \subseteq A, B \in \mathcal{C}} m_{co}(B) \Big] \\ &= \sum_{A \not\supseteq A_1} b(A)\vec{v}_A + \sum_{A \supseteq A_1} \vec{v}_A \Big[\sum_{B \subseteq A, B \in \mathcal{C}} (m_b(B) - m_{co}(B)) + \sum_{B \subseteq A, B \notin \mathcal{C}} m_b(B) \Big] \\ &= \sum_{A \not\supseteq A_1} b(A)\vec{v}_A + \sum_{A \supseteq A_1} \vec{v}_A \Big[\gamma(A) + b(A) - \sum_{A_j \in \mathcal{C}, A_j \subseteq A} m_b(A_j) \Big], \end{aligned} \tag{39}$$

after introducing the auxiliary variables $\gamma(A) = \sum_{B \subseteq A, B \in \mathcal{C}} (m_b(B) - m_{co}(B))$. All the terms in (39) associated with subsets $A \supseteq A_i, A \not\supseteq A_{i+1}$ depend on the same auxiliary variable $\gamma(A_i)$, while the difference in the component $\vec{v}_\Theta$ is trivially $1 - 1 = 0$. Therefore, we obtain (38).

Proof of Theorem 5: after recalling the expression (38) of the difference vector $\vec{b} - \vec{co}$ in the belief space, the latter's L_1 norm reads as:

$$\|\vec{b} - \vec{co}\|_{L_1} = \sum_{i=1}^{n-1} \sum_{A \supseteq A_i, A \not\supseteq A_{i+1}} \Big| \gamma(A_i) + b(A) - \sum_{j=1}^{i} m_b(A_j) \Big| + \sum_{A \not\supseteq A_1} |b(A)|. \tag{40}$$

The norm (40) can be decomposed into a number of summations which depend on a single auxiliary variable $\gamma(A_i)$. Such components are of the form $|x + x_1| + ... + |x + x_n|$, with an even number of "nodes" $-x_i$. Let us consider the simple function of Figure 6-left: it is easy to see that similar functions are minimized by the interval of values comprised between their two innermost nodes, i.e., in the case of norm (40): $\sum_{j=1}^{i} m_b(A_j) - \lambda^i_{int1} \leq \gamma(A_i) \leq \sum_{j=1}^{i} m_b(A_j) - \lambda^i_{int2}$ $\forall i = 1, ..., n-1$, i.e., $\sum_{j=1}^{i} m_b(A_j) - \lambda^i_{int2} \leq \sum_{j=1}^{i} (m_b(A_j) - m_{co}(A_j)) \leq \sum_{j=1}^{i} m_b(A_j) - \lambda^i_{int2}$

$\forall i = 1, ..., n-2$. This is equivalent to $\lambda^i_{int1} \leq \sum_{j=1}^{i} m_{co}(A_j) \leq \lambda^i_{int2}$ $\forall i = 1, ..., n-2$, while $m_{co}(A_{n-1}) = b(A_{n-1})$, as by definition (21) $\lambda^{n-1}_{int1} = \lambda^{n-1}_{int2} = b(A_{n-1})$.
This is a set of constraints of the form $l_1 \leq x \leq u_1$, $l_2 \leq x + y \leq u_2$, $l_3 \leq x + y + z \leq u_3$, also expressed as $l_1 \leq x \leq u_1$, $l_2 - x \leq y \leq u_2 - x$, $l_3 - (x + y) \leq z \leq u_3 - (x + y)$. This is a polytope whose 2^{n-2} vertices are obtained by assigning to x, $x + y$, $x + y + z$ etcetera either their lower or their upper bound. For the specific set above this yields exactly (20).

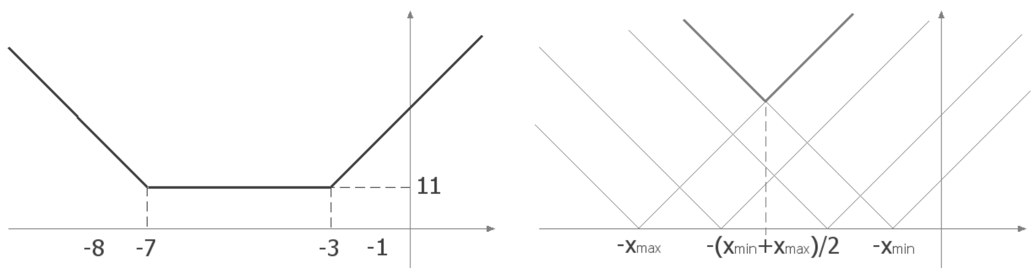

Fig. 6. Left: the minimization of the L_1 distance from the consonant subspace involves minimizing functions such as the one depicted above, $|x + 1| + |x + 3| + |x + 7| + |x + 8|$, which is minimized by $3 \leq x \leq 7$. Right: the minimization of the L_∞ distance from the consonant subspace involves minimizing functions of the form $\max\{|x + x_1|, ..., |x + x_n|\}$ (in bold).

Proof of Theorem 6: the minimal value of a function of the form $|x + x_1| + ... + |x + x_n|$ is $\sum_{i \geq int_2} x_i - \sum_{i \leq int_1} x_i$. In the case of the L_1 norm (40), such minimal attained value is:

$$\sum_{A:A \supseteq A_i, A \not\supseteq A_{i+1}, b(A) \geq \lambda^i_{int_2}} b(A) - \sum_{A:A \supseteq A_i, A \not\supseteq A_{i+1}, b(A) \leq \lambda^i_{int_1}} b(A),$$

since in the difference the addenda $\sum_{j=1}^{i} m_b(A_j)$ disappear. Overall the minimal L_1 norm is: $\sum_{i=1}^{n-2} \Big(\sum_{A:A \supseteq A_i, A \not\supseteq A_{i+1}, b(A) \geq \lambda^i_{int_2}} b(A) - \sum_{A:A \supseteq A_i, A \not\supseteq A_{i+1}, b(A) \leq \lambda^i_{int_1}} b(A) \Big) + \sum_{A \not\supseteq A_1} b(A) = \sum_{\emptyset \subsetneq A \subsetneq \Theta, A \neq A_{n-1}} b(A) - 2 \sum_{i=1}^{n-2} \sum_{A:A \supseteq A_i, A \not\supseteq A_{i+1}, b(A) \leq \lambda^i_{int_1}} b(A)$, which is minimized by $\arg\max_{\mathcal{C}} \sum_i^{n-1} \sum_{b(A) \in \mathcal{L}_i, b(A) \leq \lambda^i_{int_1}} b(A)$.

Proof of Theorem 7: by replacing the hypothesized solution (25) for the L_2 approximation in $\mathcal{B}$ in the system of constraints (24) we get, for all $j = 1, ..., n-1$:

$$\left\{ \sum_{A \subsetneq \Theta} m_b(A) \langle \vec{b}_A, \vec{b}_{A_j} \rangle - ave(\mathcal{L}_{n-1}) \langle \vec{b}_{A_{n-1}}, \vec{b}_{A_{n-1}} \rangle - \sum_{i=1}^{n-2} ave(\mathcal{L}_i) \Big(\langle \vec{b}_{A_i}, \vec{b}_{A_j} \rangle - \langle \vec{b}_{A_{i+1}}, \vec{b}_{A_j} \rangle \Big) = 0, \right.$$

where $\langle \vec{b}_{A_{n-1}}, \vec{b}_{A_{n-1}} \rangle = 1$ for all j, while (since $\langle \vec{b}_A, \vec{b}_B \rangle = |\{C \subsetneq \Theta : C \supseteq A, B\}| = 2^{|(A \cup B)^c|} - 1$): $\langle \vec{b}_{A_i}, \vec{b}_{A_j} \rangle - \langle \vec{b}_{A_{i+1}}, \vec{b}_{A_j} \rangle = \langle \vec{b}_{A_j}, \vec{b}_{A_j} \rangle - \langle \vec{b}_{A_j}, \vec{b}_{A_j} \rangle = 0$ whenever $i < j$, and $\langle \vec{b}_{A_i}, \vec{b}_{A_j} \rangle - \langle \vec{b}_{A_{i+1}}, \vec{b}_{A_j} \rangle = \big(|\{A \supseteq A_i, A_j\}| - 1\big) - \big(|\{A \supseteq A_{i+1}, A_j\}| - 1\big) = |\{A \supseteq A_i\}| - |\{A \supseteq A_{i+1}, \}| =$

$2^{|A^c_{i+1}|}$ whenever $i \geq j$. Therefore the system of constraints becomes (as $2^{A^c_n} = 2^{|\emptyset|} = 1$):

$$\left\{ \sum_{A \subsetneq \Theta} m_b(A) \langle \vec{b}_A, \vec{b}_{A_j} \rangle - \sum_{i=j}^{n-1} ave(\mathcal{L}_i) 2^{|A^c_{i+1}|} = 0 \qquad j = 1, ..., n-1, \right.$$

which, given the expression (26) for $ave(\mathcal{L}_i)$, reads as:

$$\left\{ \sum_{A \subsetneq \Theta} m_b(A) \langle \vec{b}_A, \vec{b}_{A_j} \rangle - \sum_{i=j}^{n-1} \sum_{A \supseteq A_i, A \not\supseteq A_{i+1}} b(A) = 0. \qquad j = 1, ..., n-1. \right. \tag{41}$$

Let us study the second addenda of each equation above. We get: $\sum_{i=j}^{n-1} \sum_{A \supseteq A_i, A \not\supseteq \{x_{i+1}\}} b(A) = \sum_{A_j \subseteq A \subsetneq \Theta} b(A)$, as any $A \supseteq A_j$, $A \neq \Theta$ is such that $A \supseteq A_i$ and $A \not\supseteq A_{i+1}$ for some A_i in the desired maximal chain which contains A_j. Indeed, let us define x_{i+1} as the lowest index element (according to the ordering associated with the desired focal chain $A_1 \subset \cdots \subset A_n$, i.e., $x_j \doteq A_j \setminus A_{j-1}$) among those singletons in A^c. But then, by construction, $A \supseteq A_i$ and $A \not\supseteq \{x_{i+1}\}$. Finally: $\sum_{A_j \subseteq A \subsetneq \Theta} b(A) = \sum_{A_j \subseteq A \subsetneq \Theta} \sum_{C \subseteq A} m_b(C) = \sum_{C \subsetneq \Theta} m_b(C) \big| \{A : C \subseteq A \subsetneq \Theta, A \supseteq A_j\} \big|$ where $|\{A : C \subseteq A \subsetneq \Theta, A \supseteq A_j\}| = |\{A : A \supseteq (C \cup A_j), A \neq \Theta\}| = 2^{|(C \cup A_j)^c|} - 1 = \langle \vec{b}_C, \vec{b}_{A_j} \rangle$ so that, summarizing, $\sum_{i=j}^{n-1} \sum_{A \supseteq A_i, A \not\supseteq \{x_{i+1}\}} b(A) = \sum_{C \subsetneq \Theta} m_b(C) \langle \vec{b}_C, \vec{b}_{A_j} \rangle$.

By replacing the latter into (41) we obtain the trivial identity $0 = 0$.

Proof of Theorem 8: given the expression (38) for the difference vector of interest in the belief space, we can compute the explicit form of its L_∞ norm as:

$$\begin{aligned} \|\vec{b} - \vec{co}\|_\infty &= \max \left\{ \max_i \max_{A \supseteq A_i, A \not\supseteq A_{i+1}} \left| \gamma(A_i) + b(A) - \sum_{j=1}^{i} m_b(A_j) \right|, \max_{A \not\supseteq A_1} \left| \sum_{B \subseteq A} m_b(B) \right| \right\} \\ &= \max \left\{ \max_i \max_{A \supseteq A_i, A \not\supseteq A_{i+1}} \left| \gamma(A_i) + b(A) - \sum_{j=1}^{i} m_b(A_j) \right|, b(A_1^c) \right\}, \end{aligned} \tag{42}$$

as $\max_{A \not\supseteq A_1} \left| \sum_{B \subseteq A} m_b(B) \right| = b(A_1^c)$. Now, (42) can be minimized separately for each $i = 1, ..., n-1$. Clearly, the minimum is attained when the variable elements in (42) are not greater than the constant element $b(A_1^c)$:

$$\max_{A \supseteq A_i, A \not\supseteq A_{i+1}} \left| \gamma(A_i) + b(A) - \sum_{j=1}^{i} m_b(A_j) \right| \leq b(A_1^c). \tag{43}$$

The left hand side of (43) is a function of the form $\max\big\{|x + x_1|, ..., |x + x_n|\big\}$ (see Figure 6-right). Such functions are minimized by $x = -\frac{x_{min} + x_{max}}{2}$ (see Figure 6-right again). In the case of (43), such minimum and maximum offset values are, respectively,

$$\lambda^i_{min} = b(A_i) - \sum_{j=1}^{i} m_b(A_j), \;\; \lambda^i_{max} = b(\{x_{i+1}\}^c) - \sum_{j=1}^{i} m_b(A_j) = b(A_i \cup A^c_{i+1}) - \sum_{j=1}^{i} m_b(A_j),$$

once defined $\{x_{i+1}\} = A_{i+1} \setminus A_i$. As for each value of γ, $|\gamma(A_i) + \gamma|$ is dominated by either $|\gamma(A_i) + \lambda^i_{min}|$ or $|\gamma(A_i) + \lambda^i_{max}|$, the norm of the difference vector is minimized by the values of $\gamma(A_i)$ such that: $\max\left\{|\gamma(A_i) + \lambda^i_{min}|, |\gamma(A_i) + \lambda^i_{max}|\right\} \leq b(A^c_1) \quad \forall i = 1, ..., n-1$, i.e., $-\frac{\lambda^i_{min} + \lambda^i_{max}}{2} - b(A^c_1) \leq \gamma(A_i) \leq -\frac{\lambda^i_{min} + \lambda^i_{max}}{2} + b(A^c_1)$ for $i = 1, ..., n-1$.

In terms of mass assignments, this is equivalent to:

$$-b(A^c_1) + \frac{b(A_i) + b(\{x_{i+1}\}^c)}{2} \leq \sum_{j=1}^{i} m_{co}(A_i) \leq b(A^c_1) + \frac{b(A_i) + b(\{x_{i+1}\}^c)}{2}. \quad (44)$$

Once again this is a set of constraints of the form $l_1 \leq x \leq u_1$, $l_2 \leq x+y \leq u_2$, $l_3 \leq x+y+z \leq u_3$, also expressed as $l_1 \leq x \leq u_1$, $l_2 - x \leq y \leq u_2 - x$, $l_3 - (x+y) \leq z \leq u_3 - (x+y)$, which is a polytope with vertices obtained by assigning to x, $x+y$, $x+y+z$ etcetera either their lower or their upper bound. This generates 2^{n-1} possible combinations, which for the specific set (44) yields (see the proof of Theorem 5) Equation (27). As for the barycenter of (27), we have that:

$$\begin{array}{rcl} m_{co}(A_1) & = & \frac{b(A_1)+b(\{x_2\}^c)}{2} \\ m_{co}(A_i) & = & \frac{b(A_i)+b(\{x_{i+1}\}^c)}{2} - \frac{b(A_{i-1})+b(\{x_i\}^c)}{2} = \frac{b(A_i)-b(A_{i-1})}{2} + \frac{pl_b(\{x_i\})-pl_b(\{x_{i+1}\})}{2}, \\ m_{co}(A_n) & = & 1 - \sum_{i=2}^{n-1} \left[\frac{b(A_i)+b(\{x_{i+1}\}^c)}{2} - \frac{b(A_{i-1})+b(\{x_i\}^c)}{2}\right] - \frac{b(A_1)+b(\{x_2\}^c)}{2} = 1 - b(A_{n-1}). \end{array}$$

REFERENCES

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

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

[3] N. Wilson, "A Monte-Carlo algorithm for Dempster-Shafer belief", in *Proc. of UAI*, pp. 414-417, 1991.

[4] I. Kramosil, "Monte-Carlo estimations for belief functions", in *Proc. of iFUZZY*, vol. 16, pp. 339–357, 1998.

[5] S. Moral and A. Salmeron, "A Monte-Carlo algorithm for combining Dempster-Shafer belief based on approximate pre-computation", in *Proc. of ECSQARU*, vol. 1638 of LNAI, Springer, pp. 305315, 1999.

[6] B. Tessem, "Approximations for efficient computation in the theory of evidence", *Artif. Intell.*, vol. 61, no. 2, pp. 315–329, 1993.

[7] T. Denoeux, "Inner and outer approximation of belief structures using a hierarchical clustering approach", *Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 4, pp. 437–460, 2001.

[8] B. R. Cobb and P. P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models", *Int. J. of Approximate Reasoning*, vol. 41, no. 3, pp. 314–330, 2006.

[9] M. Daniel, "On transformations of belief functions to probabilities", *Int. J. of Intelligent Systems*, vol. 21, no. 3, pp. 261–282, 2006.

[10] F. Voorbraak, “A computationally efficient approximation of Dempster-Shafer theory”, *Int. J. on Man-Machine Studies*, vol. 30, pp. 525–536, 1989.

[11] J. J. Sudano, “Equivalence between belief theories and naive Bayesian fusion for systems with independent evidential data”, in *Proc. of ISIF*, 2003.

[12] D. Dubois and H. Prade, “Consonant approximations of belief functions,” *Int. J. of Approximate Reasoning*, vol. 4, pp. 419–449, 1990.

[13] C. Joslyn and G. Klir, “Minimal information loss possibilistic approximations of random sets”, in *Proc. of the FUZZ-IEEE Conference*, J. Bezdek, Ed., pp. 1081–1088, 1992.

[14] C. Joslyn, “Possibilistic normalization of inconsistent random intervals”, in *Proc. of the Conference of the International Institute for General Systems Studies*, pp. 44–51, 1997.

[15] P. Baroni, “Extending consonant approximations to capacities”, in *Proc. of IPMU*, pp. 1127–1134, 2004.

[16] D. Dubois, H. Prade, and S. Sandri, “On possibility-probability transformations”, in *Fuzzy Logic: State of the Art*, R. Lowen and M. Lowen, Eds. Kluwer Academic Publisher, p. 103-112, 1993.

[17] J. F. Geer and G. J. Klir, “A mathematical analysis of information-preserving transformations between probabilistic and possibilistic formulations of uncertainty”, *Int. J. of General Systems*, vol. 20, no. 2, pp. 143–176, 1992.

[18] F. Cuzzolin, “The geometry of consonant belief functions: simplicial complexes of necessity measures”, *Fuzzy Sets and Systems*, vol. 161, no. 10, pp. 1459–1479, 2010.

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

[20] E. Miranda, I. Couso, and P. Gil, “Extreme points of credal sets generated by 2-alternating capacities”, *Int. J. of Approximate Reasoning*, vol. 33, pp. 95–115, 2003.

[21] P. Smets and R. Kennes, “The transferable belief model”, *Artif. Intell.*, vol. 66, pp. 191–234, 1994.

[22] D. Dubois and H. Prade, “Unfair coins and necessity measures: towards a possibilistic interpretation of histograms”, *Fuzzy Sets and Systems*, vol. 10, no. 1, pp. 15–20, 1983.

[23] P. Smets, “Belief functions on real numbers”, *Int. J. of Approximate Reasoning*, vol. 40, no. 3, pp. 181–223, 2005.

[24] A. Aregui and T. Denoeux, “Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities”, *Int. J. of Approximate Reasoning*, vol. 49, no. 3, pp. 575–594, 2008.

[25] F. Cuzzolin, “Two new Bayesian approximations of belief functions based on convex geometry”, *IEEE Tr. SMC-B*, vol. 37, no. 4, pp. 993–1008, 2007.

[26] ——, “Consistent approximation of belief functions”, in *Proc. of ISIPTA*, 2009.

[27] ——, “A geometric approach to the theory of evidence”, *IEEE Tr. SMC-C*, vol. 38, no. 4, pp. 522–534, 2008.

[28] ——, “Geometric conditioning of belief functions”, in *Proc. of BELIEF*, 2010.

[29] A.-L. Jousselme and P. Maupin, “On some properties of distances in evidence theory”, in *Proc. of BELIEF*, 2010.

[30] C. Shi, Y. Cheng, Q. Pan, and Y. Lu, “A new method to determine evidence distance”, in *Proc. of CiSE*, pp. 1–4, 2010.

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

[32] V. Khatibi and G. Montazer, “A new evidential distance measure based on belief intervals”, *Scientia Iranica - Transactions D: Computer Science and Engineering and Electrical Engineering*, vol. 17, no. 2, pp. 119–132, 2010.

[33] J. Diaz, M. Rifqi, and B. Bouchon-Meunier, “A similarity measure between basic belief assignments”, in *FUSION’06*.

[34] F. Cuzzolin and R. Frezza, “Geometric analysis of belief space and conditional subspaces”, in *Proc. of ISIPTA*, 2001.

[35] P. Smets, “The nature of the unnormalized beliefs encountered in the transferable belief model”, in *Proc. of UAI*, pp. 292–297, 1992.

[36] I. Gilboa and D. Schmeidler, “Additive representations of non-additive measures and the Choquet integral”, *Annals of Operations Research*, vol. 52, no. 1, pp. 43–65, 1994.

[37] B. Dubrovin, S. Novikov, and A. Fomenko, *Modern Geometry Methods and Applications*. Nauka, Moscow, 1979.

[38] E. T. Ihler, J. W. Fisher III and A. S. Willsky, “Message errors in belief propagation”, in *Advances in Neural Information Processing Systems*, MIT Press, 2004.

[39] J. M. Mooij and H. J. Kappen, “Sufficient Conditions for Convergence of the Sum-Product Algorithm”, *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4422–4437, 2007.

[40] F. Cuzzolin, “Geometry of Dempster’s rule of combination”, *IEEE Tr. SMC-B*, vol. 34, no. 2, pp. 961–977, 2004.

[41] A.-L. Jousselme and P. Maupin, “Distances in evidence theory: Comprehensive survey and generalizations”, *Int. J. of Approximate Reasoning*, vol. 53, no. 2, pp. 118-145, 2012.

[42] M. Aigner, *Combinatorial theory*. Classics in Mathematics, Springer, New York, 1979.

[43] F. Cuzzolin, “Complexes of outer consonant approximations”, in *Proc. of ECSQARU*, 2009.