Two New Bayesian Approximations of Belief Functions Based on Convex Geometry

Fabio Cuzzolin

***Abstract*—In this paper, we analyze from a geometric perspective the meaningful relations taking place between belief and probability functions in the framework of the geometric approach to the theory of evidence. Starting from the case of binary domains, we identify and study three major geometric entities relating a generic belief function (b.f.) to the set of probabilities $\mathcal{P}$: 1) the dual line connecting belief and plausibility functions; 2) the orthogonal complement of $\mathcal{P}$; and 3) the simplex of consistent probabilities. Each of them is in turn associated with a different probability measure that depends on the original b.f. We focus in particular on the geometry and properties of the orthogonal projection of a b.f. onto $\mathcal{P}$ and its intersection probability, provide their interpretations in terms of degrees of belief, and discuss their behavior with respect to affine combination.**

***Index Terms*—Bayesian belief functions (b.f.), commutativity, geometric approach, intersection probability, orthogonal projection, theory of evidence.**

I. Introduction

UNCERTAINTY measures play a major role in fields like artificial intelligence, where problems involving formalized reasoning are common. The theory of evidence is among the most popular such formalisms, thanks perhaps to its nature of natural extension of the classical Bayesian methodology. Indeed, the notion of *belief function* (b.f.) [1] generalizes that of finite probability, with classical probabilities forming a subclass $\mathcal{P}$ of b.f. called *Bayesian* b.f. B.F.s are defined on the power set $2^\Theta = \{A \subset \Theta\}$ of a finite domain Θ and have the form

$$b(A) = \sum_{B \subset A} m(B)$$

where $m : 2^\Theta \rightarrow [0, 1]$ is a second function called *basic probability assignment* (b.p.a.).

The interplay of belief and Bayesian functions is of course of great interest in the theory of evidence. In particular, many people worked on the problem of finding a probabilistic or possibilistic [2] approximation of an arbitrary b.f. A number of papers [3]–[6] have been published on this issue (see [7] and [8] for a review) mainly in order to find efficient implementations of the rule of combination aiming to reduce the number of focal elements. Tessem [9], for instance, incorporated only the highest-valued focal elements in his m_{klx} approximation; a similar approach inspired the *summarization* technique formulated by Lowrance *et al.* [10]. The relation between b.f.s and probabilities is as well the foundation of a popular approach to the theory of evidence, i.e., Smets' "Transferable Belief Model" [11], where beliefs are represented at credal level while decisions are made by resorting to a Bayesian b.f. called *pignistic function* [12]. On his side, Voorbraak [13] proposed to adopt the so-called *relative plausibility function* (pl.f.) $\tilde{pl}_b$, which is the unique probability that assigns to each singleton its normalized plausibility given a b.f. b with plausibility pl_b. He proved that $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities $\tilde{pl}_b \oplus p = b \oplus p \; \forall p \in \mathcal{P}$. Cobb and Shenoy [14]–[16] analyzed the properties of the relative plausibility of singletons [17] and discussed its nature of probability function that is equivalent to the original b.f.

The study of the link between b.f.s and probabilities has also been posed in a geometric setup [18]–[20]. Black in particular dedicated his doctoral thesis to the study of the geometry of b.f.s and other monotone capacities [20]. An abstract of his results can be found in [19], where he uses shapes of geometric loci to give a direct visualization of the distinct classes of monotone capacities. In particular, a number of results about lengths of edges of convex sets representing monotone capacities are given together with their "size" meant as the sum of those lengths. Another close reference is perhaps the work of Ha and Haddawy [18], who proposed an "affine operator" that can be considered a generalization of both b.f.s and interval probabilities and can be used as a tool for constructing convex sets of probability distributions. Uncertainty is modeled as sets of probabilities represented as "affine trees," while actions (modifications of the uncertain state) are defined as tree manipulators. A small number of properties of the affine operator are also presented. In a later work [21], they presented the interval generalization of the probability cross-product operator called convex closure (cc) operator. They analyzed the properties of the cc operator relative to manipulations of sets of probabilities and presented interval versions of Bayesian propagation algorithms based on it. Probability intervals were represented in a computationally efficient fashion by means of a data structure called *pcc-tree*, where branches are annotated with intervals, and nodes are annotated with convex sets of probabilities.

On our side, in a series of recent works [22]–[24], we proposed a geometric interpretation of the theory of evidence in which b.f.s are represented as points of a simplex called *belief space* [22]. As a matter of fact, as a b.f. $b : 2^\Theta \rightarrow [0, 1]$ is completely specified by its $2^{|\Theta|} - 1$ belief values $\{b(A), A \subset \Theta,$

Manuscript received August 29, 2006; revised December 14, 2006. This work was supported in part by the VISIONTRAIN project under Contract MRTN-CT-2004-005439. This paper was recommended by Associate Editor E. Santos.

The author is with the Perception Project, INRIA Rhône-Alpes, 38334 Saint Ismier Cedex, France (e-mail: Fabio.Cuzzolin@inrialpes.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2007.895991

$A \neq \emptyset\}$, it can be represented as a point of the Cartesian space $\mathbb{R}^{N-1}$, $N \doteq 2^{|\Theta|}$. In this framework, different uncertainty descriptions like upper and lower probabilities, b.f.s, and probability and possibility measures can be studied in a unified fashion.

In this paper, we use tools provided by the geometric approach (Section III) to study the interplay of belief and Bayesian functions in the framework of the belief space. We introduce two new probabilities related to a b.f., which are both derived from purely geometric considerations. We thoroughly discuss their interpretation and properties, and their relations with the other known Bayesian approximations of b.f.s, i.e., pignistic function and relative plausibility of singletons.

A. Paper Outline

More precisely, we first look for an insight by considering the simplest case in which the frame of discernment has only two elements (Section IV). It turns out that each b.f. b is associated with three different geometric entities: 1) the simplex of consistent probabilities $\mathcal{P}[b] = \{p \in \mathcal{P} : p(A) \geq b(A)\ \forall A \subset \Theta\}$; 2) the line (b, pl_b) joining b with the related pl.f. pl_b; and 3) the orthogonal complement $\mathcal{P}^{\perp}$ of the probabilistic subspace $\mathcal{P}$. These in turn determine three different probabilities associated with b: 1) the barycenter of $\mathcal{P}[b]$ or *pignistic function* $BetP[b]$; 2) the *intersection probability* $p[b]$; and 3) the *orthogonal projection* $\pi[b]$ of b onto $\mathcal{P}$. In the binary case, all those Bayesian functions coincide.

In Section V, we prove that although the line (b, pl_b) is always orthogonal to $\mathcal{P}$, it does not intersect in general the Bayesian region. However, it does intersect the region of Bayesian *normalized sum functions* (n.s.f.s), i.e., the natural generalizations of b.f.s obtained by relaxing the positivity constraint for b.p.a. This intersection yields a Bayesian n.s.f. $\varsigma[b]$.

In Section VI, we will see that $\varsigma[b]$ is in turn associated with a Bayesian b.f. $p[b]$, which we call intersection probability. We will give two different interpretations of the way this probability distributes the masses of the focal elements of b to the elements of Θ, both depending on the difference between plausibility and belief of singletons. We will also compare the combinatorial and geometric behavior of $p[b]$ with those of the pignistic function and the relative plausibility of singletons.

Section VII will instead be devoted to the study of the orthogonal projection of b onto the probability simplex $\mathcal{P}$. We will show that $\pi[b]$ always exists and is indeed a probability function. After precising the condition under which a b.f. b is orthogonal to $\mathcal{P}$, we will give two equivalent expressions of the orthogonal projection. We will see that $\pi[b]$ can be reduced to another probability signaling the distance of b from orthogonality, and that this "orthogonality flag" can in turn be interpreted as the result of a mass redistribution process analogous to that associated with the pignistic transformation. We will prove that as $BetP[b]$ does, $\pi[b]$ commutes with the affine combination operator and can therefore be expressed as a convex combination of basis pignistic functions, which confirms the strict relation between $\pi[b]$ and $BetP[b]$.

Finally, in Section VIII, we will briefly outline a comparison between the two functions introduced here by comparing their expressions as convex combinations, and formulate some conditions under which they coincide. For the sake of completeness, we will discuss the case of *unnormalized* b.f. (u.b.f.) and argue that, while $p[b]$ is not defined for a generic u.b.f. b, $\pi[b]$ exists and retains its properties.

To improve the readability of this paper, all major proofs have been moved to the Appendix.

II. Theory of Evidence

The *theory of evidence* [1] was introduced in the late 1970s by G. Shafer as a way of representing epistemic knowledge, which was inspired by the sequence of seminal works [25]–[27] of A. Dempster. In this formalism, the best representation of chance is a b.f. rather than a Bayesian mass distribution. A b.f. assigns probability values to *sets* of possibilities rather than single events.

Definition 1: A b.p.a. over a finite set or "frame of discernment" [1] Θ is a function $m : 2^{\Theta} \rightarrow [0, 1]$ on its power set $2^{\Theta} = \{A \subset \Theta\}$ such that

$$m(\emptyset) = 0 \qquad \sum_{A \subset \Theta} m(A) = 1, \quad m(A) \geq 0 \qquad \forall A \subset \Theta.$$

Subsets of Θ associated with nonzero values of m are called *focal elements*.

Definition 2: The b.f. $b : 2^{\Theta} \rightarrow [0, 1]$ associated with a b.p.a. m on Θ is defined as

$$b(A) = \sum_{B \subset A} m(B).$$

Conversely, the unique b.p.a. m_b associated with a given b.f. b can be recovered by means of the *Moebius inversion formula*

$$m_b(A) = \sum_{B \subset A} (-1)^{|A-B|} b(B) \tag{1}$$

so that there is a 1–1 correspondence between the two set functions $m_b \leftrightarrow b$. In the theory of evidence, a probability function or *Bayesian* b.f. is just a special b.f. assigning nonzero masses to singletons only: $m_b(A) = 0$, $|A| > 1$.

A dual mathematical representation of the evidence encoded by a b.f. b is the pl.f.

$$\begin{array}{rcl} pl_b : 2^{\Theta} & \rightarrow & [0, 1] \\ A & \mapsto & pl_b(A) \end{array}$$

where the plausibility $pl_b(A)$ of an event A is given by

$$\begin{aligned} pl_b(A) &\doteq 1 - b(A^c) \\ &= 1 - \sum_{B \subset A^c} m_b(B) \\ &= \sum_{B \cap A \neq \emptyset} m_b(B) \geq b(A) \end{aligned} \tag{2}$$

where A^c denotes the complement of A in Θ. For each event A, $pl_b(A)$ expresses the amount of evidence *not against* A.

III. Geometry of Belief and Pl.F.s

A. Belief Space

Motivated by the search for meaningful probabilistic approximations of b.f.s, we introduced the notion of *belief space* [22], [24], [28] as the space of all b.f.s with a given domain.[1] Consider a frame of discernment Θ and introduce in the Cartesian space $\mathbb{R}^{N-1}$, $N = 2^{|\Theta|}$ an orthonormal reference frame $\{X_A : A \subset \Theta, A \neq \emptyset\}$ (note that $\emptyset$ is not included). Each vector $v = \sum_{A\subset\Theta, A\neq\emptyset} v_A X(A)$ in $\mathbb{R}^{N-1}$ is then potentially a b.f., in which each component v_A measures the belief value of $A : v_A = b(A)$. Not every such vector $v \in \mathbb{R}^{N-1}$ however represents a valid b.f.

Definition 3: The *belief space* associated with Θ is the set of points $\mathcal{B}_\Theta$ of $\mathbb{R}^{N-1}$ that correspond to a b.f.

We will assume the domain Θ fixed and denote the belief space with $\mathcal{B}$. To determine which points "are" b.f.s, we can exploit the Moebius inversion lemma (1) by computing the corresponding b.p.a. and checking the axioms m_b must obey. It is not difficult to prove (see [29] for details) that $\mathcal{B}$ is convex. Let us call

$$b_A \doteq b \in \mathcal{B} \; s.t. \quad m_b(A) = 1 \quad m_b(B) = 0, \qquad \forall B \neq A$$

the unique b.f. assigning all the mass to a single subset A of Θ (Ath *basis* b.f.), and $\mathcal{E}_b$ the list of focal elements of b. The following theorem can then be proven [29].

Theorem 1: The set of all b.f.s with focal elements in a given collection L is closed and convex in $\mathcal{B}$, namely

$$\{b : \mathcal{E}_b \subset L\} = Cl(b_A : A \in L)$$

where Cl denotes the cc operator

$$Cl(b_1, \ldots, b_k) = \Big\{ b \in \mathcal{B} : b = \alpha_1 b_1 + \cdots + \alpha_k b_k, \quad \sum_i \alpha_i = 1, \alpha_i \geq 0 \qquad \forall i \Big\}. \tag{3}$$

The following is then just a consequence of Theorem 1.

Corollary 1: The belief space $\mathcal{B}$ is the cc of all basis b.f.s b_A

$$\mathcal{B} = Cl(b_A,\, A \subset \Theta,\, A \neq \emptyset). \tag{4}$$

The convex space delimited by a collection of (affinely independent [30]) points is called a *simplex*: Fig. 1 illustrates the simplicial form of $\mathcal{B}$. Each b.f. $b \in \mathcal{B}$ can be written as a convex sum as

$$b = \sum_{A\subset\Theta, A\neq\emptyset} m_b(A) b_A. \tag{5}$$

Geometrically, a b.p.a. m_b is nothing but the set of coordinates of b in the simplex $\mathcal{B}$. Clearly, since a probability is a b.f. assigning nonzero masses to singletons only, Theorem 1 implies the following corollary.

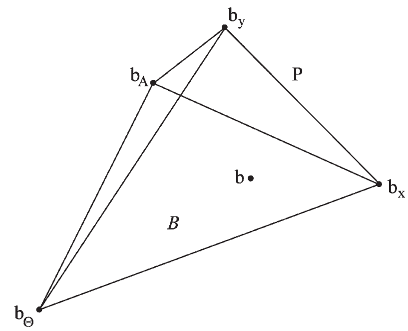

Fig. 1. Simplicial structure of the belief space $\mathcal{B}$. Its vertices are all basis b.f.s b_A represented as vectors of $\mathbb{R}^{N-1}$. The probabilistic subspace is just a subset $Cl(b_x, x \in \Theta)$ of its border.

Corollary 2: The set $\mathcal{P}$ of all Bayesian b.f.s on Θ is the simplex determined by all basis b.f.s associated with singletons[2]

$$\mathcal{P} = Cl(b_x, x \in \Theta).$$

B. Plausibility Space

As pl.f.s are also completely determined by their $N-1$ values $pl_b(A)$, $A \subset \Theta$, $A \neq \emptyset$ on the power set of Θ, they too can be seen as vectors of $\mathbb{R}^{N-1}$. We call *plausibility space* the region $\mathcal{PL}$ of $\mathbb{R}^{N-1}$ whose points correspond to pl.f.s

$$\mathcal{PL} = \Big\{ v \in \mathbb{R}^{N-1} : \exists pl_b : 2^\Theta \to [0,1] \quad s.t.\ v_A = pl_b(A), \qquad \forall A \subset \Theta, A \neq \emptyset \Big\}.$$

In [23], we proved the following proposition.

Proposition 1: $\mathcal{PL}$ is a simplex $\mathcal{PL} = Cl(pl_A, A \subset \Theta, A \neq \emptyset)$ whose vertices are

$$pl_A = -\sum_{B\subset A} (-1)^{|B|} b_B. \tag{6}$$

The vertex pl_A of the plausibility space turns out to be the plausibility vector associated with the basis b.f. b_A, $pl_A = pl_{b_A}$. Again, every plausibility vector pl_b can be uniquely expressed as a combination of the basis b.f.s b_A. We have that[3]

$$\begin{aligned} pl_b &= \sum_{B\subset\Theta} pl_b(B) X_B \\ &= \sum_{B\subset\Theta} pl_b(B) \cdot \sum_{A\supset B} b_A (-1)^{|A\setminus B|} \\ &= \sum_{A\subset\Theta} b_A \left(\sum_{B\subset A} (-1)^{|A\setminus B|} pl_b(B) \right) \end{aligned}$$

[1] Several notations in this paper have been changed with respect to other previous works in order to adopt a more standard symbology for belief and plausibility functions.

[2] With a harmless abuse of notation, we will denote the basis belief function associated with a singleton x by b_x instead of $b_{\{x\}}$. Accordingly, we will write $m_b(x), pl_b(x)$ instead of $m_b(\{x\}), pl_b(\{x\})$.

[3] Note that $pl_b(\emptyset) = 0$, so that the expression is well defined although $X_\emptyset$ does not exist.

(since by Moebius transform $X_B = \sum_{A \supset B} b_A (-1)^{|A \setminus B|}$) which yields

$$pl_b = \sum_{A \subset \Theta} \mu_b(A) b_A \tag{7}$$

where (see [23])

$$\begin{aligned} \mu_b(A) &\doteq \sum_{B \subset A} (-1)^{|A \setminus B|} pl_b(B) \\ &= (-1)^{|A|+1} \sum_{B \supset A} m_b(B), \quad A \neq \emptyset \end{aligned} \tag{8}$$

($\mu_b(\emptyset) = 0$) is the Moebius inverse of the pl.f. called *basic plausibility assignment* (b.pl.a.). The Bayesian region $\mathcal{P} = Cl(b_x, x \in \Theta)$ is part of the border of both belief and plausibility spaces.

C. N.S.F.s

It may be confusing to think of belief and pl.f.s as points of the same Cartesian space. However, this is a simple consequence of the fact that both are defined on the same domain, i.e., the power set of Θ. As Θ is finite, they can both be seen as real-valued vectors with the same number $N - 1 = 2^{|\Theta|} - 1$ of components.

Furthermore, as belief and plausibility spaces do not exhaust the whole $\mathbb{R}^{N-1}$, it is natural to wonder whether points "outside" them have any meaningful interpretation in this framework [29]. In fact, following the same principle, each vector $v = [v_1, \dots, v_A, \dots, v_\Theta]' \in \mathbb{R}^{N-1}$ can be thought of as a function $\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ s.t. $\varsigma(A) = v_A$. Each of these functions ς has a Moebius inverse $m_\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ such that

$$\varsigma(A) = \sum_{B \subset A} m_\varsigma(B)$$

i.e., *each* vector ς of $\mathbb{R}^{N-1}$ can be thought of as a *sum function* (see [31] for a brief introduction). However, m_ς does not in general meet the positivity constraint: $m_\varsigma(A) \ngeq 0 \, \forall A \subset \Theta$.

The section $\{v \in \mathbb{R}^{N-1} : v_\Theta = 1\}$ of $\mathbb{R}^{N-1}$ corresponds to the constraint $\varsigma(\Theta) = 1$, so that all points of this section are sum functions meeting the normalization axiom

$$\sum_{A \subset \Theta} m_\varsigma(A) = 1.$$

Normalized sum functions (N.S.F.s) are natural extensions of b.f.s in this geometric framework. Analogous to the case of b.f.s, we call *Bayesian n.s.f.* any n.s.f. ς such that

$$\sum_{x \in \Theta} m_\varsigma(x) = 1. \tag{9}$$

IV. Belief and Probability in the Binary Case

It may be helpful to visually render these concepts in a simple example. Fig. 2 shows the geometry of belief and plausibility spaces for a binary frame $\Theta_2 = \{x, y\}$. As $|\Theta| = 2$, b.f. and pl.f. "are" vectors $[v_x, v_y, v_\Theta]'$ of a space with $N - 1 = 2^2 - 1 = 3$ dimensions. However, since $b(\Theta) = pl_b(\Theta) = 1$ for all b, we can neglect the component $v_\Theta \equiv 1$ and represent belief and plausibility vectors as points of a plane with coordinates

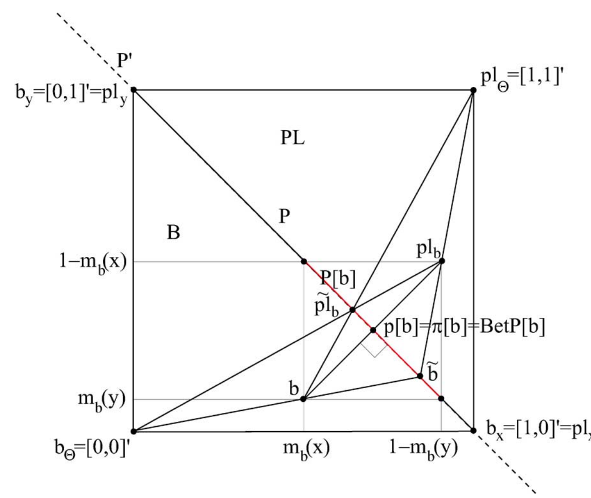

Fig. 2. In a binary frame $\Theta_2 = \{x, y\}$, both belief $\mathcal{B}$ and plausibility $\mathcal{PL}$ space are simplices with vertices $\{b_\Theta = [0,0]', b_x = [1,0]', b_y = [0,1]'\}$ and $\{pl_\Theta = [1,1]', pl_x = b_x, pl_y = b_y\}$, respectively. A b.f. b and the corresponding pl.f. pl_b are always located in symmetric positions with respect to the segment $\mathcal{P}$ of probabilities on Θ_2. The associated relative plausibility $\tilde{pl}_b$ and belief $\tilde{b}$ of singletons are shown as the intersections of the probabilistic subspace with the line joining pl_b and $b_\Theta = [0,0]'$ and the line passing through b and b_Θ, respectively. The other Bayesian functions related to b all coincide with the center of the segment of consistent probabilities $\mathcal{P}[b]$.

$$\begin{aligned} b &= [b(x) = m_b(x), b(y) = m_b(y)]' \\ pl_b &= [pl_b(x) = 1 - m_b(y), pl_b(y) = 1 - m_b(x)]'. \end{aligned}$$

In this case, the b.pl.a. of b is $\mu_b(x) = (-1)^2 \sum_{B \supset x} m_b(B) = m_b(x) + m_b(\Theta) = pl_b(x)$, $\mu_b(y) = (-1)^2 \sum_{B \supset y} m_b(B) = m_b(y) + m_b(\Theta) = pl_b(y)$, and $pl_b = pl_b(x) b_x + pl_b(y) b_y$. We can notice that both $\mathcal{B}$ and $\mathcal{PL}$ are symmetric with respect to the Bayesian region $\mathcal{P}$. Furthermore, each pair of functions (b, pl_b) determines a line *orthogonal* to $\mathcal{P}$, where b and pl_b lie on symmetric positions on the two sides of the Bayesian segment $\mathcal{P} = Cl(b_x, b_y)$.

Let us denote with $a(v_1, \dots, v_k)$ the affine subspace of a Cartesian space $\mathbb{R}^m$ generated by some points $v_1, \dots, v_k \in \mathbb{R}^m$, i.e., the set $\{v \in \mathbb{R}^m : v = \alpha_1 v_1 + \cdots + \alpha_k v_k, \sum_i \alpha_i = 1\}$.

In the binary case, the plane $\mathbb{R}^2$ in which $\mathcal{B}, \mathcal{PL}$ lie is the affine space of n.s.f.s on Θ_2. The region $\mathcal{P}'$ of all Bayesian n.s.f. is obviously (9) the line

$$\mathcal{P}' = \left\{ \varsigma \in \mathbb{R}^2 : m_\varsigma(x) + m_\varsigma(y) = 1 \right\} = a(\mathcal{P})$$

and coincides with the affine space $a(\mathcal{P}) = a(b_x, x \in \Theta)$ generated by $\mathcal{P}$.

Consider now the set of probabilities $\mathcal{P}[b]$ dominating b (*consistent* probabilities), i.e., the Bayesian b.f. such that $p(A) \geq b(A) \; \forall A \subset \Theta$. In the simple binary case, the probabilities consistent with b form a segment (1-D simplex) in $\mathcal{P}$ (see Fig. 2) whose center of mass $\overline{\mathcal{P}}$ is known [23], [32], [33] to be Smets' *pignistic function* [34], [35]

$$BetP[b] = \sum_{x\in\Theta} b_x \sum_{A\supset x} \frac{m_b(A)}{|A|} = b_x\left(m_b(x) + \frac{m_b(\Theta)}{2}\right) + b_y\left(m_b(y) + \frac{m_b(\Theta)}{2}\right). \tag{10}$$

We can notice however that it also coincides with the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$, and the intersection $p[b]$ of the line $a(b, pl_b)$ with the Bayesian simplex $\mathcal{P}$

$$p[b] = \pi[b] = BetP[b] = \overline{\mathcal{P}}[b].$$

Epistemic notions like consistency and pignistic transformation seem then to be related to geometric properties such as orthogonality. It is natural to wonder whether this is true in general or is just an artifact of the binary frame.

It is worth to notice incidentally that the *relative plausibility of singletons* $\tilde{pl}_b$ [13]

$$\tilde{pl}_b(x) \doteq \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)} \tag{11}$$

although consistent with b does *not* follow the same scheme. The same can be said of the *relative belief* of singletons, i.e., the Bayesian function

$$\tilde{b}(x) \doteq \frac{m_b(x)}{\sum_{y\in\Theta} m_b(y)}$$

assigning to each singleton x its normalized mass (see Fig. 2). We will consider their behavior separately in the near future [36].

In the following, we will instead study two other geometric loci related to b, in particular the line $a(b, pl_b)$ and the orthogonal complement $\mathcal{P}^\perp$ of $\mathcal{P}$, and introduce the two Bayesian b.f.s associated with them, i.e., orthogonal projection $\pi[b]$ and intersection probability $p[b]$. We will compare them with both pignistic function and relative plausibility of singletons, and with each other. We will provide interpretations of $\pi[b]$, $p[b]$ in terms of degrees of belief and discuss their behavior with respect to affine combination.

V. Geometry of the Dual Line

Let us then first consider the "dual line" connecting a pair of belief and plausibility measures supporting the same evidence. As a matter of fact, orthogonality turns out to be a general feature of $a(b, pl_b)$. As we just saw in the binary case, $b(\Theta) = pl_b(\Theta) = 1\ \forall b$, so that we can consider b, pl_b as points of $\mathbb{R}^{N-2}$.

A. Orthogonality

Let us consider the affine subspace $a(\mathcal{P}) = a(b_x, x\in\Theta)$ generated by the simplex of Bayesian b.f.s. This can be written as the translated version of a vector space

$$a(\mathcal{P}) = b_x + span(b_y - b_x\ \forall y\in\Theta, y\neq x)$$

where $span(b_y - b_x)$ denotes the vector space generated by the $n-1$ vectors $b_y - b_x$ ($n = |\Theta|$). After recalling that, by definition

$$b_B(A) = \begin{cases} 1, & A\supset B \\ 0, & \text{else} \end{cases} \tag{12}$$

we can point out that these vectors show a rather peculiar symmetry

$$b_y - b_x(A) = \begin{cases} 1, & A\supset\{y\}, A\not\supset\{x\} \\ 0, & A\supset\{x\},\{y\} \text{ or } A\not\supset\{x\},\{y\} \\ -1, & A\not\supset\{y\}, A\supset\{x\} \end{cases} \tag{13}$$

that can be usefully exploited.

Lemma 1: $[b_y - b_x](A^c) = -[b_y - b_x](A)\ \forall A\subset\Theta$.

Proof: By (12) $[b_y - b_x](A) = 1 \Rightarrow A\supset\{y\}, A\not\supset\{x\} \Rightarrow A^c\supset\{x\}, A^c\not\supset\{y\} \Rightarrow [b_y - b_x](A^c) = -1$ and vice-versa. On the other side, $[b_y - b_x](A) = 0 \Rightarrow A\supset\{y\}, A\supset\{x\}$ or $A\not\supset\{y\}, A\not\supset\{x\}$. In the first case, $A^c\not\supset\{x\},\{y\}$, and in the second one, $A^c\supset\{x\},\{y\}$. In both cases, $[b_y - b_x](A^c) = 0$. ■

Theorem 2: The line connecting pl_b and b in $\mathbb{R}^{N-2}$ is orthogonal to the affine space generated by the probabilistic simplex, i.e., $b - pl_b \perp a(\mathcal{P})$.

Proof [4]*:* Having denoted with X_A the Ath axis of the orthonormal reference frame $\{X_A : A\neq\Theta,\emptyset\}$ in $\mathbb{R}^{N-2}$ (see Section III), we can write their difference as

$$pl_b - b = \sum_{\emptyset\subsetneq A\subsetneq\Theta} [pl_b(A) - b(A)]\, X_A$$

where

$$\begin{aligned}[pl_b - b](A^c) &= pl_b(A^c) - b(A^c) \\ &= 1 - b(A) - b(A^c) \\ &= 1 - b(A^c) - b(A) \\ &= pl_b(A) - b(A) \\ &= [pl_b - b](A).\end{aligned} \tag{14}$$

The scalar product $\langle\cdot,\cdot\rangle$ between the vector $pl_b - b$ and the basis vectors of $a(\mathcal{P})$ is then

$$\langle pl_b - b, b_y - b_x\rangle = \sum_{\emptyset\subsetneq A\subsetneq\Theta} [pl_b - b](A)\cdot[b_y - b_x](A)$$

which by (14) becomes

$$\sum_{|A|\leq\lfloor|\Theta|/2\rfloor, A\neq\emptyset} [pl_b - b](A)\Big\{[b_y - b_x](A) + [b_y - b_x](A^c)\Big\}$$

whose addenda are all nil by Lemma 1. ■

[4] In fact, the proof is valid for $A = \Theta, \emptyset$ too.

B. Intersection With the Region of Bayesian N.S.F.s

One might be tempted to conclude that since $a(b, pl_b)$ and $\mathcal{P}$ are always orthogonal, their intersection *is* the orthogonal projection of b onto $\mathcal{P}$ as in the binary case. Unfortunately, this is not the case for in general they *do not intersect* each other.

As a matter of fact, b and pl_b belong to a (2^{n-2})-dimensional Euclidean space, while the dimension of $\mathcal{P}$ is only $n-1$. If $n=2$, $n-1=1$ and $2^n-2=2$ so that $a(\mathcal{P})$ divides the plane into two half-planes with b on one side and pl_b on the other side (see Fig. 2).

Formally, for a point on the line $a(b, pl_b)$ to be a probability, we need to find a value of α such that $b+\alpha(pl_b - b) \in \mathcal{P}$. Its components obviously are $b(A)+\alpha[pl_b(A)-b(A)]$ for any subset $A \subset \Theta, A \neq \Theta, \emptyset$ and in particular when $A=\{x\}$ is a singleton

$$b(x)+\alpha\left[pl_b(x)-b(x)\right] = b(x)+\alpha\left[1-b(x^c)-b(x)\right]. \quad (15)$$

A necessary condition for this point to belong to $\mathcal{P}$ is the normalization constraint for singletons

$$\sum_{x\in\Theta} b(x) + \alpha \sum_{x\in\Theta}\left(1-b(x^c)-b(x)\right) = 1$$

$$\Rightarrow \alpha = \frac{1-\sum_{x\in\Theta} b(x)}{\sum_{x\in\Theta}\left(1-b(x^c)-b(x)\right)} \doteq \beta[b] \quad (16)$$

which yields a single candidate value $\beta[b]$ for the line coordinate of the intersection.

Using the terminology in Section III-C, the candidate projection

$$\varsigma[b] \doteq b + \beta[b](pl_b - b) = a(b, pl_b) \cap \mathcal{P}' \quad (17)$$

(having called $\mathcal{P}'$ the set of all Bayesian n.s.f.s in $\mathbb{R}^{N-2}$) is *a Bayesian* n.s.f. but is not guaranteed to be a Bayesian b.f. For n.s.f.s, the condition $\sum_{x\in\Theta} m_\varsigma(x) = 1$ implies $\sum_{|A|>1} m_\varsigma(A) = 0$, so that $\mathcal{P}'$ can be written as

$$\mathcal{P}' = \left\{ \varsigma = \sum_{A\subset\Theta} m_\varsigma(A) b_A \in \mathbb{R}^{N-2} : \sum_{|A|=1} m_\varsigma(A) = 1, \sum_{|A|>1} m_\varsigma(A) = 0 \right\}. \quad (18)$$

Theorem 3: The coordinates of $\varsigma[b]$ with respect to the basis Bayesian b.f.s $\{b_x, x \in \Theta\}$ can be expressed in terms of the b.p.a. m_b of b as

$$m_{\varsigma[b]}(x) = m_b(x) + \beta[b] \sum_{A\supset x, A\neq x} m_b(A) \quad (19)$$

where

$$\beta[b] = \frac{1-\sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta}\left(pl_b(x)-m_b(x)\right)} = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}. \quad (20)$$

Proof: The numerator of (16) is trivially $\sum_{|B|>1} m(B)$. On the other side

$$1-b(x^c)-b(x) = \sum_{B\subset\Theta} m_b(B) - \sum_{B\subset x^c} m_b(B) - m_b(x) = \sum_{B\supset x, B\neq x} m_b(B)$$

so that the denominator of $\beta[b]$ becomes

$$\sum_{y\in\Theta}\left[pl_b(y)-b(y)\right] = \sum_{y\in\Theta}\left(1-b(y^c)-b(y)\right) = \sum_{y\in\Theta}\sum_{B\supset y, B\neq y} m_b(B) = \sum_{|B|>1} m_b(B)|B|$$

yielding (20). Equation (19) comes directly from (15) when we recall that $b(x)=m_b(x), \varsigma(x)=m_\varsigma(x)\ \forall x\in\Theta$. ∎

Equation (19) ensures that $m_{\varsigma[b]}(x)$ is positive for each $x\in\Theta$. A symmetric version can be obtained after realizing that $\left(\sum_{|B|=1} m_b(B)/\sum_{|B|=1} m_b(B)|B|\right)=1$, so that we can write

$$m_{\varsigma[b]}(x) = b(x)\frac{\sum_{|B|=1} m_b(B)}{\sum_{|B|=1} m_b(B)|B|} + [pl_b - b](x)\frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}. \quad (21)$$

It is easy to prove that the line $a(b, pl_b)$ intersects the probabilistic subspace *only for 2-additive* b.f.s (the proof can be found in the Appendix).

Theorem 4: $\varsigma[b] \in \mathcal{P}$ if and only if (iff) b is 2-additive, i.e., $m_b(A)=0 |A|>2$, and in this case, pl_b is the reflection of b through $\mathcal{P}$.

For 2-additive b.f.s, $\varsigma[b]$ is nothing but the *mean probability* function $(b+pl_b)/2$. In the general case however, the reflection of b through $\mathcal{P}$ not only does not coincide with pl_b but is also not even a pl.f. [37].

VI. Intersection Probability

We have seen that although the line $a(b, pl_b)$ is always orthogonal to $\mathcal{P}$, it does not intersect the probabilistic subspace in general, but it does intersect the region of Bayesian n.s.f.s in $\varsigma[b]$ (17). But of course (since $\sum_x m_{\varsigma[b]}(x)=1$) $\varsigma[b]$ is naturally associated with a Bayesian b.f., assigning an equal amount of mass to each singleton and 0 to each $A : |A|>1$, namely

$$p[b] \doteq \sum_{x\in\Theta} m_{\varsigma[b]}(x) b_x \quad (22)$$

where $m_{\varsigma[b]}(x)$ is given by (19). It is easy to see that $p[b]$ is a probability, since by definition $m_{p[b]}(A)=0$ for $|A|>1$, $m_{p[b]}(x) = m_{\varsigma[b]}(x) \geq 0 \quad \forall x\in\Theta$, and $\sum_{x\in\Theta} m_{p[b]}(x) = \sum_{x\in\Theta} m_{\varsigma[b]}(x) = 1$ by construction. We call $p[b]$ the

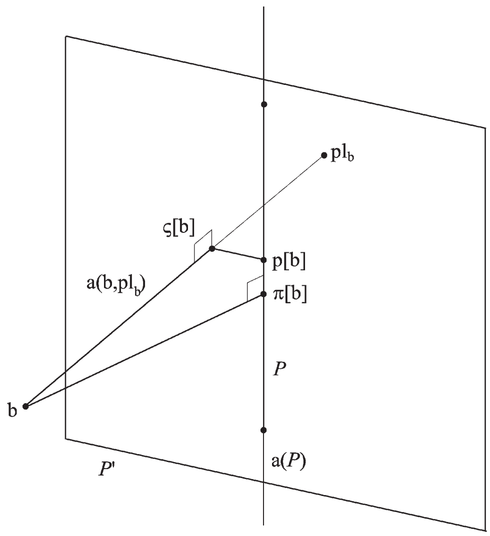

Fig. 3. Geometry of the line $a(b, pl_b)$ and relative locations of $p[b]$, $\varsigma[b]$, and $\pi[b]$. Each b.f. b and the related pl.f. pl_b lie on opposite sides of the hyperplane $\mathcal{P}'$ of the Bayesian n.s.f. that divides $\mathbb{R}^{N-2}$ into two parts. The line $a(b, pl_b)$ connecting them always intersects $\mathcal{P}'$ but not necessarily $a(\mathcal{P})$ (vertical line). This intersection $\varsigma[b]$ is naturally associated with a probability $p[b]$ (in general distinct from the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$) having the same components in the base $\{b_x, x \in \Theta\}$ of $a(\mathcal{P})$. $\mathcal{P}$ is a simplex (a segment in the figure) in $a(\mathcal{P})$: $\pi[b]$ and $p[b]$ are both "true" probabilities.

intersection probability. The geometry of $\varsigma[b]$ and $p[b]$ with respect to the regions of Bayesian b.f. and n.s.f. is sketched in Fig. 3.

A. Interpretations

1) Non-Bayesianity Flag and Relative Plausibility: A first interpretation of this new probability is immediate after noticing that

$$\beta[b] = \frac{1 - \sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta} pl_b(x) - \sum_{x\in\Theta} m_b(x)} = \frac{1 - k_{\tilde{b}}}{k_{\tilde{pl}_b} - k_{\tilde{b}}}$$

where

$$k_{\tilde{b}} = \sum_{x\in\Theta} m_b(x)$$
$$k_{\tilde{pl}_b} = \sum_{x\in\Theta} pl_b(x) = \sum_{A\subset\Theta} m_b(A)|A|$$

are the normalization factors for $\tilde{b}$ and $\tilde{pl}_b$, respectively, so that $p[b]$ can be rewritten as

$$p[b](x) = m_b(x) + (1 - k_{\tilde{b}})\frac{pl_b(x) - m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}}. \qquad (23)$$

When b is Bayesian, $pl_b(x) - m_b(x) = 0 \ \forall x \in \Theta$. If b is not Bayesian, there exists at least a singleton x such that $pl_b(x) - m_b(x) > 0$. The Bayesian b.f.

$$R[b](x) \doteq \frac{\sum_{A\supset x, A\neq x} m_b(A)}{\sum_{|A|>1} m_b(A)|A|} = \frac{pl_b(x) - m_b(x)}{\sum_{y\in\Theta}(pl_b(y) - m_b(y))}$$

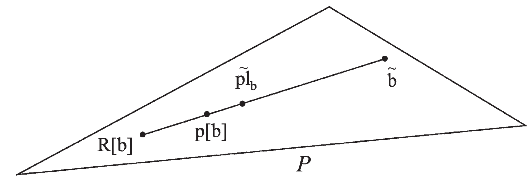

Fig. 4. Location of intersection probability $p[b]$ and relative plausibility of singletons $\tilde{pl}_b$ with respect to the non-Bayesianity flag $R[b]$. They both lie on the segment joining $R[b]$ and the relative belief of singletons $\tilde{b}$, but $\tilde{pl}_b$ is closer to $\tilde{b}$ than $p[b]$.

then measures the relative contribution of each singleton x to the non-Bayesianity of b. Equation (23) shows in fact that the non-Bayesian mass $1 - k_{\tilde{b}}$ is assigned by $p[b]$ to each singleton according to its relative contribution $R[b](x)$ to the non-Bayesianity of b.

The flag probability $R[b]$ also relates the intersection probability $p[b]$ to other two classical Bayesian approximations, i.e., the relative plausibility $\tilde{pl}_b$ and belief $\tilde{b}$ of singletons, as (23) reads as

$$p[b] = k_{\tilde{b}}\tilde{b} + (1 - k_{\tilde{b}})R[b]. \qquad (24)$$

Geometrically, since $k_{\tilde{b}} = \sum_{x\in\Theta} m_b(x) \leq 1$, $p[b]$ belongs to the segment linking $R[b]$ with the relative belief of singletons $\tilde{b}$ with convex coordinate the total mass of singletons $k_{\tilde{b}}$. But now, the relative pl.f. can also be written in terms of $\tilde{b}$ and $R[b]$ as by definition

$$\begin{aligned} R[b](x) &= \frac{pl_b(x) - m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \\ &= \frac{pl_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} - \frac{m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \\ &= \tilde{pl}_b(x)\frac{k_{\tilde{pl}_b}}{k_{\tilde{pl}_b} - k_{\tilde{b}}} - \tilde{b}(x)\frac{k_{\tilde{b}}}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \end{aligned}$$

since $\tilde{pl}_b(x) = pl_b(x)/k_{\tilde{pl}_b}$, and $\tilde{b}(x) = m_b(x)/k_{\tilde{b}}$, so that

$$\tilde{pl}_b = \left(\frac{k_{\tilde{b}}}{k_{\tilde{pl}_b}}\right)\tilde{b} + \left(1 - \frac{k_{\tilde{b}}}{k_{\tilde{pl}_b}}\right)R[b].$$

Both $\tilde{pl}_b$ and $p[b]$ belong to $Cl(R[b], \tilde{b})$. However, as $k_{\tilde{pl}_b} = \sum_{A\subset\Theta} m_b(A)|A| \geq 1$, $k_{\tilde{b}}/k_{\tilde{pl}_b} \leq k_{\tilde{b}}$, which in turn implies that $p[b]$ is closer to $R[b]$ than the relative pl.f. $\tilde{pl}_b$ (see Fig. 4). The convex coordinate of $\tilde{pl}_b$ in $Cl(R[b], \tilde{b})$ measures the ratio between total mass and plausibility of singletons. Obviously, when $k_{\tilde{b}} = 0$ ($\tilde{b}$ does not exists), $p[b] = \tilde{pl}_b = R[b]$ by (23).

2) Meaning of the Ratio $\beta[b]$ and Pignistic Function: To shed more light on $p[b]$ and get an alternative interpretation of the intersection probability, it is useful to compare $p[b]$ as expressed in (23) with another classical Bayesian approximation of b, i.e., the pignistic function

$$BetP[b](x) \doteq \sum_{A\supset x} \frac{m_b(A)}{|A|} = m_b(x) + \sum_{A\supset x, A\neq x} \frac{m_b(A)}{|A|}.$$

We can notice that in $BetP[b]$, the mass of each event A, $|A| > 1$ is considered *separately*, and its mass $m_b(A)$ is *equally* shared among the elements of A. In $p[b]$, instead, it is the *total* mass $\sum_{|A|>1} m_b(A) = 1 - k_{\tilde{b}}$ of nonsingletons that is considered, and this total mass is distributed *proportionally* to their non-Bayesian contribution to each element of Θ.

How should $\beta[b]$ be interpreted then? If we write $p[b](x)$ as

$$p[b](x) = m_b(x) + \beta[b]\left(pl_b(x) - m_b(x)\right) \tag{25}$$

we can observe that a fraction measured by $\beta[b]$ of its non-Bayesian contribution $pl_b(x) - m_b(x)$ is *uniformly* assigned to each singleton. This leads to another parallelism between $p[b]$ and $BetP[b]$. It suffices to note that if $|A| > 1$

$$\beta[b_A] = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} = \frac{1}{|A|}$$

so that both $p[b](x)$ and $BetP[b](x)$ assume the form

$$m_b(x) + \sum_{A \supset x, A \neq x} m_b(A)\beta_A$$

where $\beta_A = const = \beta[b]$ for $p[b]$, while $\beta_A = \beta[b_A]$ in case of the pignistic function.

Under which condition $p[b]$ and pignistic function coincide? A sufficient condition can be achieved by decomposing $\beta[b]$ as

$$\begin{aligned}\beta[b] &= \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}\\ &= \frac{\sum_{k=2}^{n}\sum_{|B|=k} m_b(B)}{\sum_{k=2}^{n}\left(k\sum_{|B|=k} m_b(B)\right)}\\ &= \frac{\sigma^2 + \cdots + \sigma^n}{2\sigma^2 + \cdots + n\sigma^n}\end{aligned} \tag{26}$$

after defining $\sigma^k \doteq \sum_{|B|=k} m_b(B)$.

Theorem 5: Intersection probability and pignistic function coincide if $\exists k \in [2, \ldots, n]$ such that $\sigma^i = 0\ \forall i \neq k$, i.e., the focal elements of b have size 1 or k only.

Proof: $p[b] = BetP[b]$ is equivalent to

$$\begin{aligned}m_b(x) + \sum_{A \supset x, A \neq x} m_b(A)\beta[b] &= m_b(x) + \sum_{A \supset x, A \neq x} \frac{m_b(A)}{|A|}\\ &\equiv \sum_{A \supset x, A \neq x} m_b(A)\beta[b]\\ &= \sum_{A \supset x, A \neq x} \frac{m_b(A)}{|A|}.\end{aligned}$$

If $\exists k : m_b(A) = 0$ for $|A| \neq k$, then $\beta[b] = 1/k$, and the equality is met. ■

In particular, this is true when $\Sigma^i = 0$, $i > 2$, i.e., when b is 2-additive. The condition of Theorem 5 is in fact a rather straightforward generalization of the concept of 2-additivity.

3) Example: Let us see a simple example to briefly discuss the two interpretations of $p[b]$ introduced above. Consider a ternary frame $\Theta = \{x, y, z\}$, and a b.f. b with b.p.a. given by

$$\begin{aligned}m_b(x) &= 0.1 & m_b(y) &= 0\\ m_b(z) &= 0.2 & m_b(\{x,y\}) &= 0.3\\ m_b(\{x,z\}) &= 0.1 & m_b(\{y,z\}) &= 0\\ m_b(\Theta) &= 0.3.\end{aligned}$$

Recalling (23), the total mass of singletons is $k_{\tilde{b}} = 0.1 + 0 + 0.2 = 0.3$, while the non-Bayesian contributions of x, y, z are respectively

$$\begin{aligned}pl_b(x) - m_b(x) &= m_b(\Theta) + m_b(\{x,y\}) + m_b(\{x,z\}) = 0.7\\ pl_b(y) - m_b(y) &= m_b(\{x,y\}) + m_b(\Theta) = 0.6\\ pl_b(z) - m_b(z) &= m_b(\{x,z\}) + m_b(\Theta) = 0.4\end{aligned}$$

so that the non-Bayesian flag has values $R(x) = 0.7/1.7$, $R(y) = 0.6/1.7$, $R(z) = 0.4/1.7$.

For each singleton then, the original b.p.a. $m_b(x)$ is increased by a share of the mass of nonsingletons $1 - k_{\tilde{b}} = 0.7$ proportional to the value of $R(x)$, i.e.,

$$\begin{aligned}p[b](x) &= m_b(x) + (1 - k_{\tilde{b}})R(x)\\ &= 0.1 + 0.7 * 0.7/1.7\\ &= 0.388\\ p[b](y) &= m_b(y) + (1 - k_{\tilde{b}})R(y)\\ &= 0 + 0.7 * 0.6/1.7\\ &= 0.247\\ p[b](z) &= m_b(z) + (1 - k_{\tilde{b}})R(z)\\ &= 0.2 + 0.7 * 0.4/1.7\\ &= 0.365.\end{aligned}$$

Equivalently, the line coordinate $\beta[b]$ of $p[b]$ is equal to

$$\begin{aligned}&\frac{1 - k_{\tilde{b}}}{m_b(\{x,y\})\,|\{x,y\}| + m_b(\{x,z\})\,|\{x,z\}| + m_b(\Theta)|\Theta|}\\ &= \frac{0.7}{0.3 * 2 + 0.1 * 2 + 0.3 * 3} = \frac{0.7}{1.7}\end{aligned}$$

and measures the share of $pl_b(x) - m_b(x)$ assigned to each singleton

$$\begin{aligned}p[b](x) &= m_b(x) + \beta[b]\left(pl_b(x) - m_b(x)\right)\\ &= 0.1 + 0.7/1.7 * 0.7\\ p[b](y) &= m_b(y) + \beta[b]\left(pl_b(y) - m_b(y)\right)\\ &= 0 + 0.7/1.7 * 0.6\\ p[b](z) &= m_b(z) + \beta[b]\left(pl_b(z) - m_b(z)\right)\\ &= 0.2 + 0.7/1.7 * 0.4.\end{aligned}$$

VII. Orthogonal Projection

Although the intersection of the line $a(b, pl_b)$ with the region $\mathcal{P}'$ of the Bayesian n.s.f. is not always in $\mathcal{P}$, an orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ is obviously guaranteed to exist as $a(\mathcal{P})$ is nothing but a linear subspace in the space of n.s.f.s (such as b). An explicit calculation of $\pi[b]$, however, requires a description of the orthogonal complement of $a(\mathcal{P})$ in $\mathbb{R}^{N-2}$. Let us denote with $n = |\Theta|$ the cardinality of Θ.

A. Orthogonality Condition

We need to find a necessary and sufficient condition for an arbitrary vector $v = \sum_{A\subset\Theta} v_A X_A$ to be orthogonal[5] to the probabilistic subspace $a(\mathcal{P})$. If we compute the scalar product $\langle v, b_y - b_x\rangle$ between v and the generators $b_y - b_x$ of $a(P)$, we get

$$\left\langle \sum_{A\subset\Theta} v_A X_A, b_y - b_x \right\rangle = \sum_{A\subset\Theta} v_A [b_y - b_x](A)$$

which remembering (13) becomes

$$\langle v, b_y - b_x\rangle = \sum_{A\supset y, A\not\supset x} v_A - \sum_{A\supset x, A\not\supset y} v_A.$$

The orthogonal complement $a(\mathcal{P})^\perp$ of $a(\mathcal{P})$ can then be expressed as

$$v(\mathcal{P})^\perp = \left\{ v : \sum_{A\supset y, A\not\supset x} v_A = \sum_{A\supset x, A\not\supset y} v_A \;\forall y \neq x \right\}.$$

If the vector v in particular is a b.f. ($v_A = b(A)$)

$$\sum_{A\supset y, A\not\supset x} b(A) = \sum_{A\supset y, A\not\supset x} \sum_{B\subset A} m_b(B) = \sum_{B\subset\{x\}^c} m_b(B) 2^{n-1-|B\cup\{y\}|}$$

since $2^{n-1-|B\cup\{y\}|}$ is the number of subsets A of $\{x\}^c$ containing both B and y, and the orthogonality condition becomes

$$\sum_{B\subset\{x\}^c} m_b(B) 2^{n-1-|B\cup\{y\}|} = \sum_{B\subset\{y\}^c} m_b(B) 2^{n-1-|B\cup\{x\}|}, \quad \forall y \neq x.$$

Now, sets $B \subset \{x, y\}^c$ appear in both summations with the same coefficient (since in that case $|B \cup \{x\}| = |B \cup \{y\}| = |B| + 1$), and the equation, after erasing the common factor 2^{n-2}, reduces to

$$\sum_{B\supset y, B\not\supset x} m_b(B) 2^{1-|B|} = \sum_{B\supset x, B\not\supset y} m_b(B) 2^{1-|B|}, \qquad \forall y \neq x \tag{27}$$

which expresses the desired orthogonality condition.

Theorem 6: The orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ can be expressed in terms of the b.p.a. m_b of b as

$$\pi[b](x) = \sum_{A\supset x} m_b(A) 2^{1-|A|} + \sum_{A\subset\Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right) \tag{28}$$

$$\pi[b](x) = \sum_{A\supset x} m_b(A) \left(\frac{1 + |A^c| 2^{1-|A|}}{n} \right) + \sum_{A\not\supset x} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right). \tag{29}$$

Equation (29) shows that $\pi[b]$ *is indeed a probability*, since both $1 + |A^c| 2^{1-|A|} \geq 0$ and $1 - |A| 2^{1-|A|} \geq 0 \;\; \forall |A| = 1, \ldots, n$. This is not at all trivial, as $\pi[b]$ is the projection of b onto the *affine* space $a(\mathcal{P})$ and could have in principle assigned negative masses to one or more singletons. $\pi[b]$ is hence another valid candidate to the role of the probabilistic approximation of b.f. b.

B. Orthogonality Flag

Theorem 6 does not apparently provide any intuition about the meaning of $\pi[b]$ in terms of degrees of belief. In fact, if we process (29), we can reduce π to a new Bayesian function strictly related to the pignistic function.

Theorem 7: $\pi[b] = \overline{\mathcal{P}}(1 - k_O[b]) + k_O[b] O[b]$, where $\overline{\mathcal{P}}$ is the uniform probability, and

$$O[b](x) = \frac{\bar{O}[b](x)}{k_O[b]} = \frac{\sum_{A\supset x} m_b(A) 2^{1-|A|}}{\sum_{A\subset\Theta} m_b(A) |A| 2^{1-|A|}} = \frac{\sum_{A\supset x} \frac{m_b(A)}{2^{|A|}}}{\sum_{A\subset\Theta} \frac{m_b(A)|A|}{2^{|A|}}} \tag{30}$$

is a Bayesian b.f.

As $0 \leq |A| 2^{1-|A|} \leq 1$ for all $A \subset \Theta$, $k_O[b]$ assumes values in the interval [0, 1]. Theorem 7 then implies that the orthogonal projection is always located on the line segment $Cl(\overline{\mathcal{P}}, O[b])$ joining the uniform, noninformative probability, and the Bayesian function $O[b]$.

By (30), it turns out that $\pi[b] = \overline{\mathcal{P}}$ iff $O[b] = \overline{\mathcal{P}}$ (since $k_O[b] > 0$). The meaning of $O[b]$ becomes clear when noticing that condition (27) (under which a b.f. b is orthogonal to $a(\mathcal{P})$) can be rewritten as

$$\begin{aligned}
&\sum_{B\supset y, B\not\supset x} m_b(B) 2^{1-|B|} + \sum_{B\supset y, x} m_b(B) 2^{1-|B|} \\
&\quad = \sum_{B\supset x, B\not\supset y} m_b(B) 2^{1-|B|} + \sum_{B\supset y, x} m_b(B) 2^{1-|B|} \\
&\quad \equiv \sum_{B\supset y} m_b(B) 2^{1-|B|} = \sum_{B\supset x} m_b(B) 2^{1-|B|} \\
&\quad \equiv \bar{O}[b](x) = const \\
&\quad \equiv O[b](x) = const = \overline{\mathcal{P}} \qquad \forall x \in \Theta.
\end{aligned}$$

[5]The proof is again valid for $A = \Theta, \emptyset$ too. See Section VIII-A.

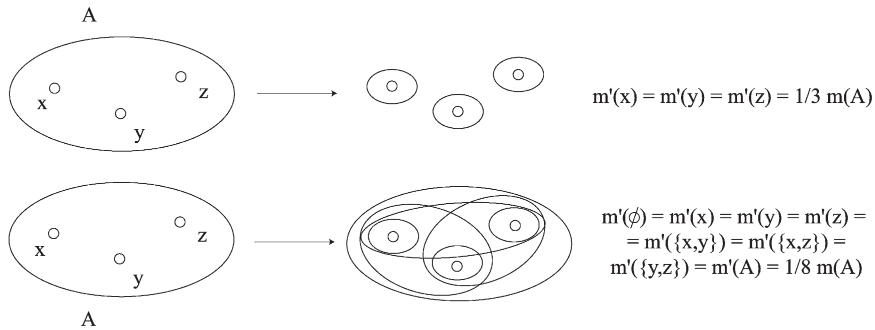

Fig. 5. Redistribution processes associated with pignistic transformation and orthogonal projection. (Top) In the pignistic transformation, the mass of each focal element A is distributed among its elements. (Bottom) In the orthogonal projection instead (through the orthogonality flag), the mass of each f.e. A is divided among all its subsets $B \subset A$. In both cases, the related relative plausibility of singletons yields a Bayesian b.f.

Therefore, $\pi[b] = \overline{\mathcal{P}}$ iff $b \perp a(\mathcal{P})$, and $O - \overline{\mathcal{P}}$ measures the nonorthogonality of b with respect to $\mathcal{P}$. $O[b]$ then deserves the name of *orthogonality flag*.

C. Interpretation in Terms of Plausibilities and Redistribution Processes

A compelling link can be drawn between orthogonal projection and pignistic function by means of the orthogonality flag $O[b]$. Let us define the two b.f.s

$$b_{\|} \doteq \frac{1}{k_{\|}} \sum_{A \subset \Theta} \frac{m_b(A)}{|A|} b_A$$

$$b_{2\|} \doteq \frac{1}{k_{2\|}} \sum_{A \subset \Theta} \frac{m_b(A)}{2^{|A|}} b_A$$

where $k_{\|}$ and $k_{2\|}$ are the normalization factors needed to make them two admissible b.f.

Theorem 8: $O[b]$ is the relative plausibility of singletons of $b_{2\|}$, and $BetP[b]$ is the relative plausibility of singletons of $b_{\|}$.

Proof: By definition of pl.f.

$$pl_{b_{2\|}}(x) = \sum_{A \supset x} m_{b_{2\|}}(A) = \frac{1}{k_{2\|}} \sum_{A \supset x} \frac{m_b(A)}{2^{|A|}} = \frac{\bar{O}[b]}{2k_{2\|}}$$

$$\sum_{x \in \Theta} pl_{b_{2\|}}(x) = \frac{1}{k_{2\|}} \sum_{x \in \Theta} \sum_{A \supset x} \frac{m_b(A)}{2^{|A|}} = \frac{k_O[b]}{2k_{2\|}}$$

by (39). Hence, $\tilde{pl}_{b_{2\|}}(x) = \bar{O}[b]/k_O[b] = O[b]$. Equivalently

$$pl_{b_{\|}}(x) = \sum_{A \supset x} m_{b_{\|}}(A) = \frac{1}{k_{\|}} \sum_{A \supset x} \frac{m_b(A)}{|A|} = \frac{1}{k_{\|}} BetP[b](x)$$

and since $\sum_x BetP[b](x) = 1$, $\tilde{pl}_{b_{\|}}(x) = BetP[b](x)$. ■

The two functions $b_{\|}$ and $b_{2\|}$ represent two different processes acting on b (see Fig. 5). The first one redistributes the mass of each focal element among its *singletons* (yielding directly a Bayesian b.f. $BetP[b]$). The second one distributes the b.p.a. of each event A among its *subsets* $B \subset A$ ($\emptyset, A$ included). In this second case, we get a u.b.f. [38] b^U

$$m_{b^U}(A) = \sum_{B \supset A} \frac{m_b(B)}{2^{|B|}}$$

whose relative belief of singletons $\tilde{b}^U$ is in fact the orthogonality flag $O[b]$.

1) Example: Let us consider again as an example the b.f. b on the ternary frame seen in Section VI-A3. To get the orthogonality flag $O[b]$, we need to apply the redistribution process of Fig. 5 (bottom) to each focal element of b. In this case, their masses are divided among their subsets as

$$\begin{aligned}
m(x) = 0.1 &\mapsto m'(x) = m'(\emptyset) = 0.1/2 = 0.05 \\
m(z) = 0.2 &\mapsto m'(z) = m'(\emptyset) = 0.2/2 = 0.1 \\
m(\{x,y\}) = 0.3 &\mapsto m'(\{x,y\}) = m'(x) = m'(y) \\
&= m'(\emptyset) = 0.3/4 = 0.075 \\
m(\{x,z\}) = 0.1 &\mapsto m'(\{x,z\}) = m'(x) = m'(z) \\
&= m'(\emptyset) = 0.1/4 = 0.025 \\
m(\Theta) = 0.3 &\mapsto m'(\Theta) = m'(\{x,y\}) = m'(\{x,z\}) \\
&= m'(\{y,z\}) = m'(x) = m'(y) \\
&= m'(z) = m'(\emptyset) = 0.3/8 = 0.0375.
\end{aligned}$$

By summing all contributions related to singletons on the right-hand side, we get

$$m_{b^U}(x) = 0.05 + 0.075 + 0.025 + 0.0375 = 0.1875$$

$$m_{b^U}(y) = 0.075 + 0.0375 = 0.1125$$

$$m_{b^U}(z) = 0.1 + 0.025 + 0.0375 = 0.1625$$

whose sum is the normalization factor

$$k_O[b] = m_{b^U}(x) + m_{b^U}(y) + m_{b^U}(z) = 0.4625$$

so that by normalizing, we get $O[b] = [0.405, 0.243, 0.351]'$. The orthogonal projection $\pi[b]$ is finally the convex combination of $O[b]$ and $\overline{\mathcal{P}} = [1/3, 1/3, 1/3]'$ with coordinate $k_O[b]$

$$\begin{aligned}\pi[b] &= \overline{\mathcal{P}}(1 - k_O[b]) + k_O[b]O[b]\\ &= [1/3,1/3,1/3]'(1-0.4625)+0.4625[0.405, 0.243, 0.351]'\\ &= [0.366, 0.291, 0.342]'.\end{aligned}$$

D. Orthogonal Projection and Affine Combination

As a confirmation of this relationship, orthogonal projection and pignistic function both commute with affine combination.

Theorem 9: Orthogonal projection and affine combination commute, i.e., if $\alpha_1 + \alpha_2 = 1$

$$\pi[\alpha_1 b_1 + \alpha_2 b_2] = \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2].$$

Proof: By Theorem 7, $\pi[b] = (1 - k_O[b])\overline{\mathcal{P}} + \bar{O}[b]$, where $k_O[b] = \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}$, and $\bar{O}[b](x) = \sum_{A\supset x} m_b(A)2^{1-|A|}$. Hence

$$\begin{aligned}k_O[\alpha_1 b_1 + \alpha_2 b_2] &= \sum_{A\subset\Theta} (\alpha_1 m_{b_1}(A) + \alpha_2 m_{b_2}(A))\,|A|2^{1-|A|}\\ &= \alpha_1 k_O[b_1] + \alpha_2 k_O[b_2],\\ \bar{O}[\alpha_1 b_1 + \alpha_2 b_2](x) &= \sum_{A\supset x} (\alpha_1 m_{b_1}(A) + \alpha_2 m_{b_2}(A))\,2^{1-|A|}\\ &= \alpha_1 \bar{O}[b_1] + \alpha_2 \bar{O}[b_2]\end{aligned}$$

which in turn implies (since $\alpha_1 + \alpha_2 = 1$)

$$\begin{aligned}\pi[\alpha_1 b_1 + \alpha_2 b_2] =\ & (1 - \alpha_1 k_O[b_1] - \alpha_2 k_O[b_2])\,\overline{\mathcal{P}}\\ & + \alpha_1 \bar{O}[b_1] + \alpha_2 \bar{O}[b_2]\\ =\ & \alpha_1 \left[(1 - k_O[b_1])\,\overline{\mathcal{P}} + \bar{O}[b_1]\right]\\ & + \alpha_2 \left[(1 - k_O[b_2])\,\overline{\mathcal{P}} + \bar{O}[b_2]\right]\\ =\ & \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2].\end{aligned}$$

■

This property can be used to find an alternative expression of the orthogonal projection as the *convex combination of the pignistic functions associated with all basis b.f.s*.

Lemma 2: The orthogonal projection of a basis b.f. b_A is given by $\pi[b_A] = (1 - |A|2^{1-|A|})\overline{\mathcal{P}} + |A|2^{1-|A|}\overline{\mathcal{P}}_A$, where $\overline{\mathcal{P}}_A = (1/|A|)\sum_{x\in A} b_x$ is the center of mass of all the probabilities with support in A.

Proof: By (30), $k_O[b_A] = |A|2^{1-|A|}$, so that

$$\bar{O}[b_A](x) = \begin{cases} 2^{1-|A|}, & x \in A\\ 0, & x \notin A\end{cases} \Rightarrow O[b_A](x) = \begin{cases} \frac{1}{|A|}, & x \in A\\ 0, & x \notin A\end{cases}$$

i.e., $O[b_A] = (1/|A|)\sum_{x\in A} b_x = \overline{\mathcal{P}}_A$. ■

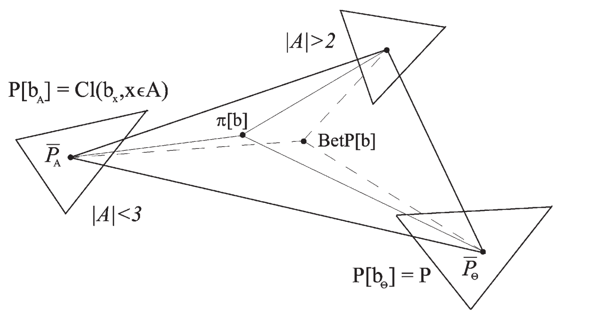

Fig. 6. Orthogonal projection $\pi[b]$ and pignistic function $BetP[b]$ are both located on the simplex whose vertices are all the basis pignistic functions, i.e., the uniform probabilities associated with each single event A. However, the convex coordinates of $\pi[b]$ are weighted by a factor $k_O[b_A] = |A|2^{1-|A|}$, yielding a point that is closer to vertices related to lower size events.

Theorem 10: The orthogonal projection can be expressed as a convex combination of all noninformative probabilities with support on a single event A as

$$\begin{aligned}\pi[b] &= \overline{\mathcal{P}}\left(1 - \sum_{A\neq\Theta}\alpha_A\right) + \sum_{A\neq\Theta}\alpha_A\overline{\mathcal{P}}_A\\ \alpha_A &\doteq m_b(A)|A|2^{1-|A|}.\end{aligned} \tag{31}$$

Proof:

$$\pi[b] = \pi\left[\sum_{A\subset\Theta} m_b(A) b_A\right] = \sum_{A\subset\Theta} m_b(A)\pi[b_A]$$

by Theorem 9, which by Lemma 2 becomes

$$\begin{aligned}&\sum_{A\subset\Theta} m_b(A)\left[(1 - |A|2^{1-|A|})\overline{\mathcal{P}} + |A|2^{1-|A|}\overline{\mathcal{P}}_A\right]\\ &= \left(1 - \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\right)\overline{\mathcal{P}} + \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\overline{\mathcal{P}}_A\\ &= \left(1 - \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\right)\overline{\mathcal{P}} + \sum_{A\neq\Theta} m_b(A)|A|2^{1-|A|}\overline{\mathcal{P}}_A\\ &\quad + m_b(\Theta)|\Theta|2^{1-|\Theta|}\overline{\mathcal{P}}\end{aligned}$$

i.e., (31). ■

As $\overline{\mathcal{P}}_A = BetP[b_A]$, we recognize that

$$\begin{aligned}BetP[b] &= \sum_{A\subset\Theta} m_b(A) BetP[b_A]\\ \pi[b] &= \sum_{A\neq\Theta}\alpha_A BetP[b_A] + \left(1 - \sum_{A\neq\Theta}\alpha_A\right) BetP[b_\Theta]\end{aligned} \tag{32}$$

with $\alpha_A = m_b(A)k_O[b_A]$. Both orthogonal projection and pignistic function are convex combinations of all basis pignistic functions. However, as $k_O[b_A] = |A|2^{1-|A|} < 1$ for $|A| > 2$, the orthogonal projection turns out to be closer to the vertices associated with events of lower cardinality (see Fig. 6).

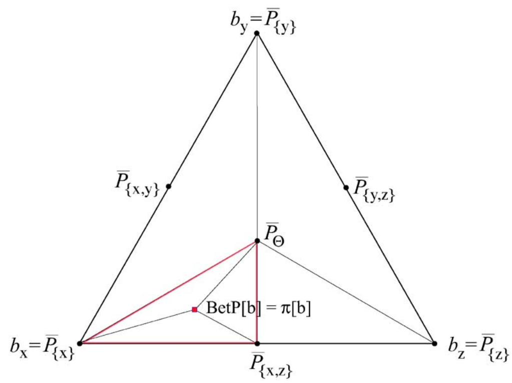

Fig. 7. Orthogonal projection and pignistic function for the b.f. (33) on the ternary frame $\Theta_3 = \{x, y, z\}$.

1) Example—Ternary Case: Let us consider as an example a ternary frame $\Theta_3 = \{x, y, x\}$ and a b.f. on Θ_3 with b.p.a.

$$\begin{aligned} m_b(x) &= 1/3 \\ m_b(\{x, z\}) &= 1/3 \\ m_b(\Theta_3) &= 1/3 \\ m_b(A) &= 0, \quad A \neq \{x\}, \{x, z\}, \Theta_3. \end{aligned} \tag{33}$$

According to (31)

$$\begin{aligned} \pi[b] &= 1/3\overline{\mathcal{P}}_{\{x\}} + 1/3\overline{\mathcal{P}}_{\{x,z\}} + (1 - 1/3 - 1/3)\overline{\mathcal{P}} \\ &= \frac{1}{3}b_x + \frac{1}{3}\frac{b_x + b_z}{2} + \frac{1}{3}\frac{b_x + b_y + b_z}{3} \\ &= b_x\left(\frac{1}{3} + \frac{1}{6} + \frac{1}{9}\right) + b_z\left(\frac{1}{6} + \frac{1}{9}\right) + b_y\frac{1}{9} \\ &= \frac{11}{18}b_x + \frac{1}{9}b_y + \frac{5}{18}b_z \end{aligned}$$

and the orthogonal projection is the barycenter of the simplex $Cl(\overline{\mathcal{P}}_{\{x\}}, \overline{\mathcal{P}}_{\{x,z\}}, \overline{\mathcal{P}})$ (see Fig. 7). On the other side

$$\begin{aligned} BetP[b](x) &= \frac{m_b(x)}{1} + \frac{m_b(x, z)}{2} + \frac{m_b(\Theta_3)}{3} = \frac{11}{18} \\ BetP[b](y) &= \frac{1}{9} \\ BetP[b](z) &= \frac{1}{6} + \frac{1}{9} = \frac{5}{18} \end{aligned}$$

i.e., $BetP[b] = \pi[b]$. This is true for each b.f. $b \in \mathcal{B}_3$, since for (32) if $|\Theta| = 3$ then $\alpha_A = m_b(A)$ for $|A| \leq 2$, and $1 - \sum_A \alpha_A = 1 - \sum_{A \neq \Theta} m_b(A) = m_b(\Theta)$.

2) Distance Between $BetP$ and π in the Quaternary Case: To get a hint of the relationship between orthogonal projection and pignistic function in the general case, let us compare their expressions in the simplest case in which they are distinct: a frame $\Theta = \{x, y, z, w\}$ of size 4. Their analytic expressions for the generic element $x \in \Theta$ are

$$\begin{aligned} BetP[b](x) = m_b(x) &+ \frac{1}{2}\left(m_b(\{x, y\}) + m_b(\{x, z\}) + m_b(\{x, w\})\right) \\ &+ \frac{1}{3}\left(m_b(\{x, y, z\}) + m_b(\{x, y, w\}) + m_b(\{x, z, w\})\right) \\ &+ \frac{1}{4}m_b(\Theta) \\ \pi[b](x) = m_b(x) &+ \frac{1}{2}\left(m_b(\{x, y\}) + m_b(\{x, z\}) + m_b(\{x, w\})\right) \\ &+ \frac{5}{16}\left(m_b(\{x, y, z\}) + m_b(\{x, y, w\}) + m_b(\{x, z, w\})\right) \\ &+ \frac{1}{16}m_b(\{y, z, w\}) + \frac{1}{4}m_b(\Theta). \end{aligned} \tag{34}$$

They are very similar to each other. Basically, the difference is that $\pi[b]$ also counts the masses of focal elements in $\{x\}^c$ (with a small contribution), while $BetP[b]$ by definition does not.

After computing their difference

$$\begin{aligned} BetP[b](x) - \pi[b](x) = \frac{1}{48}\big[&m_b(\{x, y, z\}) + m_b(\{x, y, w\}) \\ &+ m_b(\{x, z, w\}) - 3m_b(\{y, z, w\})\big] \end{aligned}$$

we can study their L_2 distance as b varies. After introducing the notation

$$\begin{aligned} y_1 &\doteq m_b(\{x, y, z\}) \quad y_2 \doteq m_b(\{x, y, w\}) \\ y_3 &\doteq m_b(\{x, z, w\}) \quad y_4 \doteq m_b(\{y, z, w\}) \end{aligned}$$

we can maximize (minimize) the norm

$$\begin{aligned} \|BetP[b] - \pi[b]\|^2 &\doteq \sum_x |BetP[b](x) - \pi[b](x)|^2 \\ &= (y_1 + y_2 + y_3 - 3y_4)^2 \\ &\quad + (y_1 + y_2 + y_4 - 3y_3)^2 \\ &\quad + (y_1 + y_3 + y_4 - 3y_2)^2 \\ &\quad + (y_2 + y_3 + y_4 - 3y_1)^2 \end{aligned}$$

by imposing $(\partial/\partial y_i)\|BetP[b](\mathbf{y}) - \pi[b](\mathbf{y})\|^2 = 0$ subject to $y_1 + y_2 + y_3 + y_4 = 1$. The unique solution turns out to be

$$\mathbf{y} = [1/4, 1/4, 1/4, 1/4]'$$

which corresponds to [after replacing this solution into (34)] $BetP[b] = \pi[b] = \overline{\mathcal{P}}$, where $\overline{\mathcal{P}} = [1/4, 1/4, 1/4, 1/4]'$ is the uniform probability on Θ. In other words, the distance between pignistic function and orthogonal projection is minimal (zero) when all size 3 subsets have the same mass.

It is then natural to suppose that their difference must be maximal when all the mass is concentrated on a single size-3 event. This is in fact correct: $\|BetP[b] - \pi[b]\|^2$ is maximal and equal to $1^2 + 1^2 + 1^2 + (-3)^2 = 12$ when $y_i = 1$, $y_j = 0$ $\forall j \neq i$, i.e., the mass of one among $\{x, y, z\}, \{x, y, w\}, \{x, z, w\}$, $\{y, z, w\}$ is one.

VIII. Brief Discussion

The intuition for both the novel probabilistic approximations of a b.f. we introduced in this paper is provided by the analysis of the interplay between belief and probability spaces in the context of the geometric approach to the theory of evidence. Both intersection probability and orthogonal projection are related to the notion of orthogonality: the orthogonality of the dual line and that of $\pi[b] - b$ with respect to $\mathcal{P}$. Nevertheless, they possess different interpretations in terms of mass assignment, and relate in significant but distinct ways with the pignistic transformation.

An interesting parallel between $p[b]$ and $\pi[b]$ comes from their geometric description as points of a segment. Theorem 7 and (24)

$$\pi[b] = k_O[b]O[b] + \overline{\mathcal{P}}(1 - k_O[b])$$
$$p[b] = k_{\tilde{b}}\tilde{b} + (1 - k_{\tilde{b}})R[b]$$

state that they can both be written as convex combinations that depend on some flag probabilities associated with them, namely orthogonality and non-Bayesianity flag, respectively

$$\pi[b] \leftrightarrow O[b]$$
$$p[b] \leftrightarrow R[b].$$

It is then worth to study the condition under which $p[b]$ and orthogonal projection $\pi[b]$ are the same probability.

A trivial consequence of Theorem 4 is that when b is 2-additive, $\pi[b] = p[b] = \varsigma[b]$. This though gives us just "pointwise" information on the relationship between intersection probability and orthogonal projection. It would definitively be worth conducting a study of the distance between all Bayesian approximations of b.f.s, $BetP$, π, p, $\tilde{pl}_b$, $\tilde{b}$ as b varies in $\mathcal{B}$, in order to understand how they depend on the b.p.a. of b. We started doing this for the pair $BetP[b]$, $\pi[b]$ in the case of quaternary frames (Section VII-D2), getting some interesting results. We reserve to explore this direction thoroughly in the near future.

A. U.B.F.s

We also wish to add a remark on the validity of the results presented in this paper. They have been in fact obtained for "classical" b.f.s for which the mass assigned to the empty set is $0 : b(\emptyset) = m_b(\emptyset) = 0$. However, it makes sense in certain situations to work with u.b.f.s [38], i.e., b.f.s admitting nonzero support $m_b(\emptyset) \neq 0$ for the empty set [39]. $m_b(\emptyset)$ is an indicator of the amount of conflict in the evidence carried by a b.f. b but can also be interpreted as the possibility that the existing frame of discernment does not exhaust all the possible outcomes of the problem. U.B.F.s are naturally associated with vectors with $N = 2^{|\Theta|}$ coordinates. A new set of basis u.b.f. can then be defined

$$\{b_A \in \mathbb{R}^N, \emptyset \subseteq A \subseteq \Theta\}$$

this time including a vector $b_\emptyset \doteq [1\ 0\ \cdots 0]'$. Note also that in this case $b_\Theta = [0\ \cdots 0\ 1]'$.

It is natural to wonder whether the above discussion, and in particular definition and properties of $p[b]$ and $\pi[b]$, retains its validity. Let us consider again the binary case. We now have to use four coordinates associated with all events in Θ: $\emptyset$, $\{x\}$, $\{y\}$, and Θ. Remember that in the case of u.b.f.

$$b(A) = \sum_{\emptyset \subsetneq B \subseteq A} m_b(B), \qquad A \neq \emptyset$$

i.e., the contribution of the empty set is not considered when computing the belief value of an event $A \neq \emptyset$.[6] The corresponding basis belief and pl.f.s are then

$$\begin{aligned} b_\emptyset &= [1,0,0,0]' & pl_\emptyset &= [0,0,0,0]' \\ b_x &= [0,1,0,1]' & pl_x &= [0,1,0,1]' = b_x \\ b_y &= [0,0,1,1]' & pl_y &= [0,0,1,1]' = b_y \\ b_\Theta &= [0,0,0,1]' & pl_\Theta &= [0,1,1,1]'. \end{aligned}$$

A striking difference with the "classical" case is that $b(\Theta) = 1 - m_b(\emptyset) = pl_b(\Theta)$, which implies that both belief and plausibility spaces are *not* in general subsets of the section $v_\Theta = 1$ of $\mathbb{R}^N$. In other words, u.b.f. and u.pl.f. are not n.s.f.s (Section III-C).

More precisely, b, pl_b are n.s.f. iff $b(\emptyset) \neq 0$. As a consequence, *the line* $a(b, pl_b)$ *is not guaranteed to intersect the affine space* $\mathcal{P}'$ *of the Bayesian* n.s.f.

Consider for instance the line connecting $b_\emptyset$ and $pl_\emptyset$ in the binary case

$$\alpha\, b_\emptyset + (1 - \alpha)\, pl_\emptyset = \alpha\, [1,0,0,0]', \qquad \alpha \in \mathbb{R}.$$

As $\mathcal{P}' = \{[a, b, (1 - b), -a]', a, b \in \mathbb{R}\}$, there clearly is no value $\alpha \in \mathbb{R}$ s.t. $\alpha \cdot [1,0,0,0]' \in \mathcal{P}'$.

Simple calculations show that in fact $a(b, pl_b) \cap \mathcal{P}' \neq \emptyset$ iff $b(\emptyset) = 0$ (i.e., b is "classical") or (trivially) $b \in \mathcal{P}$. This is true in the general case.

Proposition 2: $p[b]$ and $\beta[b]$ are well defined for classical b.f.s only.

It is interesting to note that however the orthogonality results of Section V-A *are still valid* since Lemma 1 does not involve the empty set, while the proof of Theorem 2 is valid for the components $A = \emptyset, \Theta$ too (as $b_y - b_x(A) = 0$ for $A = \emptyset, \Theta$).

Proposition 3: $a(b, pl_b)$ is orthogonal to $\mathcal{P}$ for each u.b.f. b, although $\varsigma[b] = a(b, pl_b) \cap \mathcal{P}' \neq \emptyset$ iff b is a b.f.

Analogously, the orthogonality condition (27) does not concern the mass of the empty set. The orthogonal projection $\pi[b]$ of a u.b.f. b is then well defined (check Theorem 6's proof), and

[6] In the unnormalized case, the notation b is usually reserved for *implicability* functions, while belief functions are denoted by Bel [12].

it is still given by (28) and (29), where this time the summations on the right-hand side include the empty set too

$$\pi[b](x) = \sum_{A \supset x} m_b(A) 2^{1-|A|} + \sum_{\emptyset \subseteq A \subset \Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right)$$

$$\pi[b](x) = \sum_{A \supset x} m_b(A) \left(\frac{1 + |A^c| 2^{1-|A|}}{n} \right) + \sum_{\emptyset \subseteq A \not\supset x} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right).$$

IX. Conclusion

In this paper, we introduced two new probabilistic approximations of b.f.s, which are both derived from purely geometric considerations. They are indeed associated with two different geometric loci: the dual line passing through b and pl_b in the belief space; and the orthogonal complement of the probability subspace.

After proving that the line $a(b, pl_b)$ is always orthogonal to $\mathcal{P}$ and intersects the region of the Bayesian n.s.f. $\mathcal{P}'$, we introduced the probability $p[b]$ associated with this intersection and discussed two interpretations of $p[b]$ in terms of non-Bayesian contributions of singletons.

On the other side, after precising the condition under which a b.f. b is orthogonal to $\mathcal{P}$, we gave two equivalent expressions of the orthogonal projection of b onto $\mathcal{P}$. We saw that $\pi[b]$ can be reduced to another probability signaling the distance of b from orthogonality, and that this "orthogonality flag" can in turn be interpreted as the result of a mass redistribution process analogous to that associated with the pignistic transformation. We proved that $\pi[b]$ commutes with the affine combination operator and can therefore be expressed as a convex combination of basis pignistic functions, which confirms the strict relation between $\pi[b]$ and $BetP[b]$.

We finally studied the difference between intersection probability and orthogonal projection, and discussed which results retain their validity in the case of u.b.f.s.

We have seen when discussing the binary case that, while $BetP[b]$, $p[b]$, and $\pi[b]$ belong to the same "family" of Bayesian approximations of b (as they coincide under 2-additivity), the relative plausibility $\tilde{p}[b]$ and belief $\tilde{b}$ of singletons [13] do not fit in the same scheme. In the near future, we will show that $\tilde{p}[b]$ turns out to be the best Bayesian approximation of a b.f. in the framework of Dempster's combination rule, and investigate the dual geometry of relative plausibility and belief of singletons [36]. Naturally enough, the geometric approach can also be exploited to study the problem of approximating a b.f. with a possibility measure or "consistent" b.f. [2]. Last but not least, it will be definitively worth to seek for a complete picture of the conditions under which all different Bayesian approximations of b coincide as a crucial contribution to a full understanding their semantics.

Appendix
Proofs

Proof of Theorem 4

By definition (17), $\varsigma[b]$ can be written in terms of the reference frame $\{b_A, A \subset \Theta\}$ as

$$\sum_{A \subset \Theta} m_b(A) b_A + \beta[b] \left(\sum_{A \subset \Theta} \mu_b(A) b_A - \sum_{A \subset \Theta} m_b(A) b_A \right) = \sum_{A \subset \Theta} b_A \left[m_b(A) + \beta[b] \left(\mu_b(A) - m_b(A) \right) \right]$$

since $\mu_b(.)$ is the Moebius inverse of $pl_b(.)$. For $\varsigma[b]$ to be a Bayesian b.f., accordingly, all the components related to nonsingleton subsets need to be zero

$$m_b(A) + \beta[b] \left(\mu_b(A) - m_b(A) \right) = 0, \qquad \forall A : |A| > 1.$$

This condition in turn reduces to (recalling expression (20) of $\beta[b]$)

$$\mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \left[\sum_{|B|>1} m_b(B)|B| - \sum_{|B|>1} m_b(B) \right] = 0$$
$$\equiv \mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>1} m_b(B)(|B| - 1) = 0 \tag{35}$$

$\forall A : |A| > 1$. But now, $\sum_{|B|>1} m_b(B)(|B|-1) = \sum_{|B|>1} m_b(B) + \sum_{|B|>2} m_b(B)(|B| - 2)$, so that (35) reads as

$$[\mu_b(A) + m_b(A)] \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>2} m_b(B)(|B|-2) = 0$$
$$\equiv [m_b(A) + \mu_b(A)]\, M_1[b] + m_b(A) M_2[b] = 0 \tag{36}$$

$\forall A : |A| > 1$, after defining $M_1[b] \doteq \sum_{|B|>1} m_b(B)$, and $M_2[b] \doteq \sum_{|B|>2} m_b(B)(|B| - 2)$, respectively.

Now, it is easy to note that

$$M_1[b] = 0 \Leftrightarrow m_b(B) = 0 \qquad \forall B : |B| > 1 \Leftrightarrow b \in \mathcal{P}$$
$$M_2[b] = 0 \Leftrightarrow m_b(B) = 0 \qquad \forall B : |B| > 2$$

as all the terms inside the summations are nonnegative by definition of b.p.a.. We can distinguish three cases: 1) $M_1 = 0 = M_2$ ($b \in \mathcal{P}$); 2) $M_1 \neq 0$ but $M_2 = 0$, and finally 3) $M_1 \neq 0 \neq M_2$. If $M_1 = M_2 = 0$, then b is a probability (trivially), while if $M_1 \neq 0 \neq M_2$, then (36) implies $m_b(A) = \mu_b(A) = 0$, $|A| > 1$ i.e., $b \in \mathcal{P}$, which is a contradiction.

The only nontrivial case is then $M_2 = 0$, where condition (36) becomes

$$M_1[b]\, [m_b(A) + \mu_b(A)] = 0, \qquad \forall A : |A| > 1.$$

For all $|A| > 2$, we have that $m_b(A) = \mu_b(A) = 0$ (since $M_2 = 0$), and the constraint is met. If $|A| = 2$, instead $\mu_b(A) = (-1)^{|A|+1} \sum_{B \supset A} m_b(B) = (-1)^{2+1} m_b(A) = -m_b(A)$ (since $m_b(B) = 0 \, \forall B \supset A, |B| > 2$) so that $\mu_b(A) + m_b(A) = 0$, and the constraint is again met. Finally, as the coordinate $\beta[b]$ of $\varsigma[b]$ on the line $a(b, pl_b)$ can then be rewritten as

$$\beta[b] = \frac{M_1[b]}{M_2[b] + 2M_1[b]} \tag{37}$$

if $M_2 = 0$, then $\beta[b] = 1/2$, and $\varsigma[b] = (b + pl_b)/2$.

Proof of Theorem 6

Finding the orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ is equivalent to imposing the condition $\langle \pi[b] - b, b_y - b_x \rangle = 0 \; \forall \; y \neq x$. Replacing the masses of $\pi - b$

$$\begin{cases} \pi(x) - m_b(x), & x \in \Theta \\ -m_b(A), & |A| > 1 \end{cases}$$

into (27) yields, after extracting the singletons x from the summation, the system

$$\begin{cases} \pi(y) = \pi(x) + \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} + m_b(y) \\ \qquad - m_b(x) - \sum_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \qquad \forall y \neq x \\ \sum_{y \in \Theta} \pi(y) = 1. \end{cases} \tag{38}$$

After replacing the first $n-1$ equations of (38) into the normalization constraint, we get

$$\pi(x) + \sum_{y \neq x} \Bigg[\pi(x) + m_b(y) - m_b(x) + \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} - \sum_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \Bigg] = 1$$

which is equivalent to

$$\begin{aligned} n\pi(x) = & \, 1 + (n-1) m_b(x) - \sum_{y \neq x} m_b(y) \\ & + \sum_{y \neq x} \sum_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \\ & - \sum_{y \neq x} \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|}. \end{aligned}$$

But now

$$\sum_{y \neq x} \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} = \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A|$$

as all events A not containing x do contain some $y \neq x$, and they are counted $|A|$ times (i.e., once for each element they contain). Instead

$$\begin{aligned} \sum_{y \neq x} \sum_{A \supsetneq x, A \not\supset y} m_b(A) 2^{1-|A|} & = \sum_{A \supset x, 1<|A|<n} m_b(A) 2^{1-|A|} (n - |A|) \\ & = \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} (n - |A|) \end{aligned}$$

for $n - |A| = 0$ when $A = \Theta$. Hence, $\pi(x)$ is equal to

$$\begin{aligned} & \frac{1}{n} \Bigg[1 + (n-1) m_b(x) - \sum_{y \neq x} m_b(y) - \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A| \\ & \qquad + \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} (n - |A|) \Bigg] \\ & = \frac{1}{n} \Bigg[n \, m_b(x) + 1 - \sum_{y \in \Theta} m_b(y) + n \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad - \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} |A| - \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A| \Bigg]. \end{aligned}$$

We then just need to note that $-\sum_{y \in \Theta} m_b(y) = -\sum_{|A|=1} m_b(A) |A| 2^{1-|A|}$, so that the orthogonal projection can be finally expressed as

$$\begin{aligned} \pi(x) & = \frac{1}{n} \Bigg[n \, m_b(x) + n \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad + 1 - \sum_{A \subset \Theta} m_b(A) |A| 2^{1-|A|} \Bigg] \\ & = m_b(x) + \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad + \sum_{A \subset \Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right) \end{aligned}$$

i.e., (28), and since

$$\begin{aligned} 2^{1-|A|} + \frac{1}{n} - \frac{|A|}{n} 2^{1-|A|} & = \frac{1 + 2^{1-|A|} (n - |A|)}{n} \\ & = \frac{1 + 2^{1-|A|} |A^c|}{n} \end{aligned}$$

we get the second form (29).

Proof of Theorem 7

By (28), we can write

$$\begin{aligned} \pi[b](x) & = \bar{O}[b](x) + \frac{1}{n} \Bigg(\sum_{A \subset \Theta} m_b(A) \\ & \qquad - \sum_{A \subset \Theta} m_b(A) |A| 2^{1-|A|} \Bigg) \\ & = \bar{O}[b](x) + \frac{1}{n} \left(1 - k_O[b] \right). \end{aligned}$$

But since

$$\begin{aligned}\sum_{x\in\Theta}\bar{O}[b](x) &= \sum_{x\in\Theta}\sum_{A\supset x} m_b(A)2^{1-|A|}\\ &= \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\\ &= k_O[b]\end{aligned} \tag{39}$$

i.e., $k_O[b]$ is the normalization factor for $\bar{O}[b]$, the function (30) is a Bayesian b.f., and we can write (as $\overline{\mathcal{P}}(x) = (1/n)$) $\pi[b] = (1 - k_O[b])\overline{\mathcal{P}} + k_O[b]O[b]$.

REFERENCES

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

[2] D. Dubois and H. Prade, "Consonant approximations of belief functions," *Int. J. Approx. Reason.*, vol. 4, no. 5/6, pp. 419–449, Sep./Nov. 1990.

[3] A. B. Yaghlane, T. Denoeux, and K. Mellouli, "Coarsening approximations of belief functions," in *Proc. ECSQARU*, S. Benferhat and P. Besnard, Eds., 2001, pp. 362–373.

[4] T. Denoeux, "Inner and outer approximation of belief structures using a hierarchical clustering approach," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 9, no. 4, pp. 437–460, Aug. 2001.

[5] T. Denoeux and A. B. Yaghlane, "Approximating the combination of belief functions using the fast Moebius transform in a coarsened frame," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 77–101, Oct. 2002.

[6] R. Haenni and N. Lehmann, "Resource bounded and anytime approximation of belief function computations," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 103–154, Oct. 2002.

[7] M. Bauer, "Approximation algorithms and decision making in the Dempster–Shafer theory of evidence—An empirical study," *Int. J. Approx. Reason.*, vol. 17, no. 2/3, pp. 217–237, Aug.–Oct. 1997.

[8] M. Bauer, "Approximations for decision making in the Dempster–Shafer theory of evidence," in *Proc. 12th Conf. Uncertainty Artif. Intell.*, F. Horvitz and E. Jensen, Eds., Portland, OR, Aug. 1–4, 1996, pp. 73–80.

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

[10] J. D. Lowrance, T. D. Garvey, and T. M. Strat, "A framework for evidential-reasoning systems," in *Proc. Nat. Conf. Artif. Intell.*, 1986, pp. 896–903.

[11] P. Smets, "Belief functions versus probability functions," in *Uncertainty and Intelligent Systems*, L. Saitta, B. Bouchon, and R. Yager, Eds. Berlin, Germany: Springer-Verlag, 1988, pp. 17–24.

[12] P. Smets, "Decision making in the TBM: The necessity of the pignistic transformation," *Int. J. Approx. Reason.*, vol. 38, no. 2, pp. 133–147, Feb. 2005.

[13] F. Voorbraak, "A computationally efficient approximation of Dempster–Shafer theory," *Int. J. Man-Mach. Stud.*, vol. 30, no. 5, pp. 525–536, May 1989.

[14] B. Cobb and P. Shenoy, "On transforming belief function models to probability models," Univ. Kansas, Sch. Bus., Lawrence, KS, Feb. 2003. Working Paper 293, Tech. Rep.

[15] B. R. Cobb and P. P. Shenoy, "A comparison of Bayesian and belief function reasoning," *Inf. Syst. Frontiers*, vol. 5, no. 4, pp. 345–358, Dec. 2003.

[16] B. R. Cobb and P. P. Shenoy, "A comparison of methods for transforming belief function models to probability models," in *Proc. ECSQARU*, Aalborg, Denmark, Jul. 2003, pp. 255–266.

[17] B. Cobb and P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models," *Int. J. Approx. Reason.*, vol. 41, no. 3, pp. 314–330, Apr. 2006.

[18] V. Ha and P. Haddawy, "Theoretical foundations for abstraction-based probabilistic planning," in *Proc. 12th Conf. Uncertainty Artif. Intell.*, Aug. 1996, pp. 291–298.

[19] P. Black, "Geometric structure of lower probabilities," in *Random Sets: Theory and Applications*, J. Goutsias, R. P. S. Mahler, and H. T. Nguyen, Eds. New York: Springer-Verlag, 1997, pp. 361–383.

[20] P. Black, "An examination of belief functions and other monotone capacities," Ph.D. dissertation, Dept. Statist., Carnegie Mellon Univ., Pittsburg, PA, 1996.

[21] V. Ha and P. Haddawy, "Geometric foundations for interval-based probabilities," in *KR'98: Principles of Knowledge Representation and Reasoning*, A. G. Cohn, L. Schubert, and S. C. Shapiro, Eds. San Francisco, CA: Morgan Kaufmann, 1998, pp. 582–593. [Online]. Available: citeseer.ist.psu.edu/ha98geometric.html

[22] F. Cuzzolin and R. Frezza, "Geometric analysis of belief space and conditional subspaces," in *Proc. 2nd ISIPTA*, Ithaca, NY, Jun. 26–29, 2001, pp. 122–132.

[23] F. Cuzzolin, "Geometry of upper probabilities," in *Proc. 3rd ISIPTA*, Jul. 2003, pp. 188–203.

[24] F. Cuzzolin, "A geometric approach to the theory of evidence," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, 2007. to be published.

[25] A. P. Dempster, "Upper and lower probability inferences based on a sample from a finite univariate population," *Biometrika*, vol. 54, no. 3/4, pp. 515–528, Dec. 1967.

[26] A. Dempster, "Upper and lower probabilities generated by a random closed interval," *Ann. Math. Stat.*, vol. 39, no. 3, pp. 957–966, Jun. 1968.

[27] A. Dempster, "Upper and lower probabilities inferences for families of hypothesis with monotone density ratios," *Ann. Math. Stat.*, vol. 40, no. 3, pp. 953–969, Jun. 1969.

[28] F. Cuzzolin, "Visions of a generalized probability theory," Ph.D. dissertation, Università di Padova, Dipartimento di Elettronica e Informatica, Padova, Italy, Feb. 2001.

[29] F. Cuzzolin, "Geometry of Dempster's rule of combination," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 2, pp. 961–977, Apr. 2004.

[30] B. Dubrovin, S. Novikov, and A. Fomenko, *Sovremennaja Geometrija. Metody i Prilozenija*. Moscow, Russia: Nauka, 1986.

[31] M. Aigner, *Combinatorial Theory*. New York: Springer-Verlag, 1979.

[32] A. Chateauneuf and J. Y. Jaffray, "Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion," *Math. Soc. Sci.*, vol. 17, no. 3, pp. 263–283, Jun. 1989.

[33] D. Dubois, H. Prade, and P. Smets, "New semantics for quantitative possibility theory," in *Proc. ISIPTA*, 2001, pp. 152–161.

[34] P. Smets, "Constructing the pignistic probability function in a context of uncertainty," in *Proc. Uncertainty Artif. Intell., 5*, M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, Eds., 1990, pp. 29–39.

[35] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, no. 2, pp. 191–234, 1994.

[36] F. Cuzzolin, "The geometry of relative plausibility and belief of singletons," *Ann. Math. Artif. Intell.*, May 2007. submitted for publication.

[37] F. Cuzzolin, "The geometry of relative plausibilities," in *11th Int. Conf. IPMU, Special Session Fuzzy Measures and Integrals, Capacities and Games*, 2006.

[38] P. Smets, "The nature of the unnormalized beliefs encountered in the transferable belief model," in *Proc. 8th Annu. Conf. UAI*, 1992, pp. 292–297.

[39] P. Smets, "The application of the matrix calculus to belief functions," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 1–30, Oct. 2002.

Fabio Cuzzolin was born in Jesolo, Italy, in 1971. He received the Laurea (*magna cum laude*) and Ph.D. degrees from the University of Padova, Padova, Italy, in 1997 and 2001, respectively. His thesis was entitled "Visions of a generalized probability theory."

During his doctoral term, he was with the Autonomous Navigation and Computer Vision Laboratory (NAVLAB), University of Padova. He was also a Postodoctoral Researcher at Politecnico di Milano, Milan, Italy, and was with the UCLA Vision Laboratory, Los Angeles, CA. He is currently with the Perception Project, INRIA Rhône-Alpes, Saint Ismier Cedex, France. His research includes computer vision applications like gesture and action recognition, object pose estimation, and identity recognition from gait. His main field of investigation remains however that of generalized and imprecise probabilities. In particular, he has formulated a geometric approach to the theory of belief functions, focusing mainly on the probabilistic approximation problem, and studied the notion of independence of sources from an algebraic point of view. He collaborates with several journals, among which the *International Journal of Approximate Reasoning*, *Information Fusion*, and the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B.

Dr. Cuzzolin is a member of the Society for Imprecise Probabilities and Their Applications.

Two New Bayesian Approximations of Belief Functions Based on Convex Geometry

Fabio Cuzzolin

***Abstract*—In this paper, we analyze from a geometric perspective the meaningful relations taking place between belief and probability functions in the framework of the geometric approach to the theory of evidence. Starting from the case of binary domains, we identify and study three major geometric entities relating a generic belief function (b.f.) to the set of probabilities $\mathcal{P}$: 1) the dual line connecting belief and plausibility functions; 2) the orthogonal complement of $\mathcal{P}$; and 3) the simplex of consistent probabilities. Each of them is in turn associated with a different probability measure that depends on the original b.f. We focus in particular on the geometry and properties of the orthogonal projection of a b.f. onto $\mathcal{P}$ and its intersection probability, provide their interpretations in terms of degrees of belief, and discuss their behavior with respect to affine combination.**

***Index Terms*—Bayesian belief functions (b.f.), commutativity, geometric approach, intersection probability, orthogonal projection, theory of evidence.**

I. Introduction

UNCERTAINTY measures play a major role in fields like artificial intelligence, where problems involving formalized reasoning are common. The theory of evidence is among the most popular such formalisms, thanks perhaps to its nature of natural extension of the classical Bayesian methodology. Indeed, the notion of *belief function* (b.f.) [1] generalizes that of finite probability, with classical probabilities forming a subclass $\mathcal{P}$ of b.f. called *Bayesian* b.f. B.F.s are defined on the power set $2^\Theta = \{A \subset \Theta\}$ of a finite domain Θ and have the form

$$b(A) = \sum_{B \subset A} m(B)$$

where $m : 2^\Theta \rightarrow [0,1]$ is a second function called *basic probability assignment* (b.p.a.).

The interplay of belief and Bayesian functions is of course of great interest in the theory of evidence. In particular, many people worked on the problem of finding a probabilistic or possibilistic [2] approximation of an arbitrary b.f. A number of papers [3]–[6] have been published on this issue (see [7] and [8] for a review) mainly in order to find efficient implementations of the rule of combination aiming to reduce the number of focal elements. Tessem [9], for instance, incorporated only the highest-valued focal elements in his m_{klx} approximation; a similar approach inspired the *summarization* technique formulated by Lowrance *et al.* [10]. The relation between b.f.s and probabilities is as well the foundation of a popular approach to the theory of evidence, i.e., Smets' "Transferable Belief Model" [11], where beliefs are represented at credal level while decisions are made by resorting to a Bayesian b.f. called *pignistic function* [12]. On his side, Voorbraak [13] proposed to adopt the so-called *relative plausibility function* (pl.f.) $\tilde{pl}_b$, which is the unique probability that assigns to each singleton its normalized plausibility given a b.f. b with plausibility pl_b. He proved that $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities $\tilde{pl}_b \oplus p = b \oplus p \; \forall p \in \mathcal{P}$. Cobb and Shenoy [14]–[16] analyzed the properties of the relative plausibility of singletons [17] and discussed its nature of probability function that is equivalent to the original b.f.

The study of the link between b.f.s and probabilities has also been posed in a geometric setup [18]–[20]. Black in particular dedicated his doctoral thesis to the study of the geometry of b.f.s and other monotone capacities [20]. An abstract of his results can be found in [19], where he uses shapes of geometric loci to give a direct visualization of the distinct classes of monotone capacities. In particular, a number of results about lengths of edges of convex sets representing monotone capacities are given together with their "size" meant as the sum of those lengths. Another close reference is perhaps the work of Ha and Haddawy [18], who proposed an "affine operator" that can be considered a generalization of both b.f.s and interval probabilities and can be used as a tool for constructing convex sets of probability distributions. Uncertainty is modeled as sets of probabilities represented as "affine trees," while actions (modifications of the uncertain state) are defined as tree manipulators. A small number of properties of the affine operator are also presented. In a later work [21], they presented the interval generalization of the probability cross-product operator called convex closure (cc) operator. They analyzed the properties of the cc operator relative to manipulations of sets of probabilities and presented interval versions of Bayesian propagation algorithms based on it. Probability intervals were represented in a computationally efficient fashion by means of a data structure called *pcc-tree*, where branches are annotated with intervals, and nodes are annotated with convex sets of probabilities.

On our side, in a series of recent works [22]–[24], we proposed a geometric interpretation of the theory of evidence in which b.f.s are represented as points of a simplex called *belief space* [22]. As a matter of fact, as a b.f. $b : 2^\Theta \rightarrow [0,1]$ is completely specified by its $2^{|\Theta|} - 1$ belief values $\{b(A), A \subset \Theta,$

Manuscript received August 29, 2006; revised December 14, 2006. This work was supported in part by the VISIONTRAIN project under Contract MRTN-CT-2004-005439. This paper was recommended by Associate Editor E. Santos.

The author is with the Perception Project, INRIA Rhône-Alpes, 38334 Saint Ismier Cedex, France (e-mail: Fabio.Cuzzolin@inrialpes.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2007.895991

$A \neq \emptyset\}$, it can be represented as a point of the Cartesian space $\mathbb{R}^{N-1}$, $N \doteq 2^{|\Theta|}$. In this framework, different uncertainty descriptions like upper and lower probabilities, b.f.s, and probability and possibility measures can be studied in a unified fashion.

In this paper, we use tools provided by the geometric approach (Section III) to study the interplay of belief and Bayesian functions in the framework of the belief space. We introduce two new probabilities related to a b.f., which are both derived from purely geometric considerations. We thoroughly discuss their interpretation and properties, and their relations with the other known Bayesian approximations of b.f.s, i.e., pignistic function and relative plausibility of singletons.

A. *Paper Outline*

More precisely, we first look for an insight by considering the simplest case in which the frame of discernment has only two elements (Section IV). It turns out that each b.f. b is associated with three different geometric entities: 1) the simplex of consistent probabilities $\mathcal{P}[b] = \{p \in \mathcal{P} : p(A) \geq b(A)\ \forall A \subset \Theta\}$; 2) the line (b, pl_b) joining b with the related pl.f. pl_b; and 3) the orthogonal complement $\mathcal{P}^{\perp}$ of the probabilistic subspace $\mathcal{P}$. These in turn determine three different probabilities associated with b: 1) the barycenter of $\mathcal{P}[b]$ or *pignistic function* $BetP[b]$; 2) the *intersection probability* $p[b]$; and 3) the *orthogonal projection* $\pi[b]$ of b onto $\mathcal{P}$. In the binary case, all those Bayesian functions coincide.

In Section V, we prove that although the line (b, pl_b) is always orthogonal to $\mathcal{P}$, it does not intersect in general the Bayesian region. However, it does intersect the region of Bayesian *normalized sum functions* (n.s.f.s), i.e., the natural generalizations of b.f.s obtained by relaxing the positivity constraint for b.p.a. This intersection yields a Bayesian n.s.f. $\varsigma[b]$.

In Section VI, we will see that $\varsigma[b]$ is in turn associated with a Bayesian b.f. $p[b]$, which we call intersection probability. We will give two different interpretations of the way this probability distributes the masses of the focal elements of b to the elements of Θ, both depending on the difference between plausibility and belief of singletons. We will also compare the combinatorial and geometric behavior of $p[b]$ with those of the pignistic function and the relative plausibility of singletons.

Section VII will instead be devoted to the study of the orthogonal projection of b onto the probability simplex $\mathcal{P}$. We will show that $\pi[b]$ always exists and is indeed a probability function. After precising the condition under which a b.f. b is orthogonal to $\mathcal{P}$, we will give two equivalent expressions of the orthogonal projection. We will see that $\pi[b]$ can be reduced to another probability signaling the distance of b from orthogonality, and that this "orthogonality flag" can in turn be interpreted as the result of a mass redistribution process analogous to that associated with the pignistic transformation. We will prove that as $BetP[b]$ does, $\pi[b]$ commutes with the affine combination operator and can therefore be expressed as a convex combination of basis pignistic functions, which confirms the strict relation between $\pi[b]$ and $BetP[b]$.

Finally, in Section VIII, we will briefly outline a comparison between the two functions introduced here by comparing their expressions as convex combinations, and formulate some conditions under which they coincide. For the sake of completeness, we will discuss the case of *unnormalized* b.f. (u.b.f.) and argue that, while $p[b]$ is not defined for a generic u.b.f. b, $\pi[b]$ exists and retains its properties.

To improve the readability of this paper, all major proofs have been moved to the Appendix.

II. Theory of Evidence

The *theory of evidence* [1] was introduced in the late 1970s by G. Shafer as a way of representing epistemic knowledge, which was inspired by the sequence of seminal works [25]–[27] of A. Dempster. In this formalism, the best representation of chance is a b.f. rather than a Bayesian mass distribution. A b.f. assigns probability values to *sets* of possibilities rather than single events.

Definition 1: A b.p.a. over a finite set or "frame of discernment" [1] Θ is a function $m : 2^{\Theta} \to [0, 1]$ on its power set $2^{\Theta} = \{A \subset \Theta\}$ such that

$$m(\emptyset) = 0 \quad \sum_{A \subset \Theta} m(A) = 1, \quad m(A) \geq 0 \qquad \forall A \subset \Theta.$$

Subsets of Θ associated with nonzero values of m are called *focal elements*.

Definition 2: The b.f. $b : 2^{\Theta} \to [0, 1]$ associated with a b.p.a. m on Θ is defined as

$$b(A) = \sum_{B \subset A} m(B).$$

Conversely, the unique b.p.a. m_b associated with a given b.f. b can be recovered by means of the *Moebius inversion formula*

$$m_b(A) = \sum_{B \subset A} (-1)^{|A-B|} b(B) \tag{1}$$

so that there is a 1–1 correspondence between the two set functions $m_b \leftrightarrow b$. In the theory of evidence, a probability function or *Bayesian* b.f. is just a special b.f. assigning nonzero masses to singletons only: $m_b(A) = 0$, $|A| > 1$.

A dual mathematical representation of the evidence encoded by a b.f. b is the pl.f.

$$\begin{array}{rcl} pl_b : 2^{\Theta} & \to & [0, 1] \\ A & \mapsto & pl_b(A) \end{array}$$

where the plausibility $pl_b(A)$ of an event A is given by

$$\begin{aligned} pl_b(A) &\doteq 1 - b(A^c) \\ &= 1 - \sum_{B \subset A^c} m_b(B) \\ &= \sum_{B \cap A \neq \emptyset} m_b(B) \geq b(A) \end{aligned} \tag{2}$$

where A^c denotes the complement of A in Θ. For each event A, $pl_b(A)$ expresses the amount of evidence *not against* A.

III. Geometry of Belief and Pl.F.s

A. Belief Space

Motivated by the search for meaningful probabilistic approximations of b.f.s, we introduced the notion of *belief space* [22], [24], [28] as the space of all b.f.s with a given domain.[1] Consider a frame of discernment Θ and introduce in the Cartesian space $\mathbb{R}^{N-1}$, $N = 2^{|\Theta|}$ an orthonormal reference frame $\{X_A : A \subset \Theta, A \neq \emptyset\}$ (note that $\emptyset$ is not included). Each vector $v = \sum_{A\subset\Theta, A\neq\emptyset} v_A X(A)$ in $\mathbb{R}^{N-1}$ is then potentially a b.f., in which each component v_A measures the belief value of $A : v_A = b(A)$. Not every such vector $v \in \mathbb{R}^{N-1}$ however represents a valid b.f.

Definition 3: The *belief space* associated with Θ is the set of points $\mathcal{B}_\Theta$ of $\mathbb{R}^{N-1}$ that correspond to a b.f.

We will assume the domain Θ fixed and denote the belief space with $\mathcal{B}$. To determine which points "are" b.f.s, we can exploit the Moebius inversion lemma (1) by computing the corresponding b.p.a. and checking the axioms m_b must obey. It is not difficult to prove (see [29] for details) that $\mathcal{B}$ is convex. Let us call

$$b_A \doteq b \in \mathcal{B}\; s.t. \quad m_b(A) = 1 \quad m_b(B) = 0, \qquad \forall B \neq A$$

the unique b.f. assigning all the mass to a single subset A of Θ (Ath *basis* b.f.), and $\mathcal{E}_b$ the list of focal elements of b. The following theorem can then be proven [29].

Theorem 1: The set of all b.f.s with focal elements in a given collection L is closed and convex in $\mathcal{B}$, namely

$$\{b : \mathcal{E}_b \subset L\} = Cl(b_A : A \in L)$$

where Cl denotes the cc operator

$$Cl(b_1, \ldots, b_k) = \Bigg\{ b \in \mathcal{B} : b = \alpha_1 b_1 + \cdots + \alpha_k b_k, \qquad \sum_i \alpha_i = 1, \alpha_i \geq 0 \quad \forall i \Bigg\}. \tag{3}$$

The following is then just a consequence of Theorem 1.

Corollary 1: The belief space $\mathcal{B}$ is the cc of all basis b.f.s b_A

$$\mathcal{B} = Cl(b_A,\, A \subset \Theta,\, A \neq \emptyset). \tag{4}$$

The convex space delimited by a collection of (affinely independent [30]) points is called a *simplex*: Fig. 1 illustrates the simplicial form of $\mathcal{B}$. Each b.f. $b \in \mathcal{B}$ can be written as a convex sum as

$$b = \sum_{A\subset\Theta, A\neq\emptyset} m_b(A) b_A. \tag{5}$$

Geometrically, a b.p.a. m_b is nothing but the set of coordinates of b in the simplex $\mathcal{B}$. Clearly, since a probability is a b.f. assigning nonzero masses to singletons only, Theorem 1 implies the following corollary.

[1] Several notations in this paper have been changed with respect to other previous works in order to adopt a more standard symbology for belief and plausibility functions.

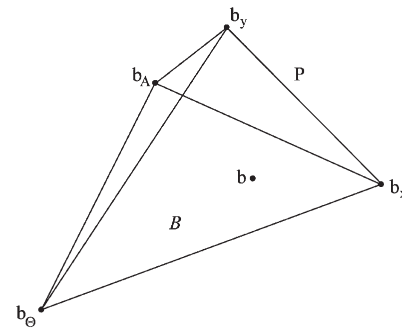

Fig. 1. Simplicial structure of the belief space $\mathcal{B}$. Its vertices are all basis b.f.s b_A represented as vectors of $\mathbb{R}^{N-1}$. The probabilistic subspace is just a subset $Cl(b_x, x \in \Theta)$ of its border.

Corollary 2: The set $\mathcal{P}$ of all Bayesian b.f.s on Θ is the simplex determined by all basis b.f.s associated with singletons[2]

$$\mathcal{P} = Cl(b_x, x \in \Theta).$$

B. Plausibility Space

As pl.f.s are also completely determined by their $N-1$ values $pl_b(A)$, $A \subset \Theta$, $A \neq \emptyset$ on the power set of Θ, they too can be seen as vectors of $\mathbb{R}^{N-1}$. We call *plausibility space* the region $\mathcal{PL}$ of $\mathbb{R}^{N-1}$ whose points correspond to pl.f.s

$$\mathcal{PL} = \Big\{ v \in \mathbb{R}^{N-1} : \exists pl_b : 2^\Theta \to [0,1] \quad s.t.\; v_A = pl_b(A), \qquad \forall A \subset \Theta, A \neq \emptyset \Big\}.$$

In [23], we proved the following proposition.

Proposition 1: $\mathcal{PL}$ is a simplex $\mathcal{PL} = Cl(pl_A, A \subset \Theta, A \neq \emptyset)$ whose vertices are

$$pl_A = -\sum_{B\subset A} (-1)^{|B|} b_B. \tag{6}$$

The vertex pl_A of the plausibility space turns out to be the plausibility vector associated with the basis b.f. b_A, $pl_A = pl_{b_A}$. Again, every plausibility vector pl_b can be uniquely expressed as a combination of the basis b.f.s b_A. We have that[3]

$$\begin{aligned} pl_b &= \sum_{B\subset\Theta} pl_b(B) X_B \\ &= \sum_{B\subset\Theta} pl_b(B) \cdot \sum_{A\supset B} b_A (-1)^{|A\setminus B|} \\ &= \sum_{A\subset\Theta} b_A \left(\sum_{B\subset A} (-1)^{|A\setminus B|} pl_b(B) \right) \end{aligned}$$

[2] With a harmless abuse of notation, we will denote the basis belief function associated with a singleton x by b_x instead of $b_{\{x\}}$. Accordingly, we will write $m_b(x), pl_b(x)$ instead of $m_b(\{x\}), pl_b(\{x\})$.

[3] Note that $pl_b(\emptyset) = 0$, so that the expression is well defined although $X_\emptyset$ does not exist.

(since by Moebius transform $X_B = \sum_{A \supset B} b_A (-1)^{|A \setminus B|}$) which yields

$$pl_b = \sum_{A \subset \Theta} \mu_b(A) b_A \tag{7}$$

where (see [23])

$$\begin{aligned} \mu_b(A) &\doteq \sum_{B \subset A} (-1)^{|A \setminus B|} pl_b(B) \\ &= (-1)^{|A|+1} \sum_{B \supset A} m_b(B), \quad A \neq \emptyset \end{aligned} \tag{8}$$

($\mu_b(\emptyset) = 0$) is the Moebius inverse of the pl.f. called *basic plausibility assignment* (b.pl.a.). The Bayesian region $\mathcal{P} = Cl(b_x, x \in \Theta)$ is part of the border of both belief and plausibility spaces.

C. N.S.F.s

It may be confusing to think of belief and pl.f.s as points of the same Cartesian space. However, this is a simple consequence of the fact that both are defined on the same domain, i.e., the power set of Θ. As Θ is finite, they can both be seen as real-valued vectors with the same number $N - 1 = 2^{|\Theta|} - 1$ of components.

Furthermore, as belief and plausibility spaces do not exhaust the whole $\mathbb{R}^{N-1}$, it is natural to wonder whether points "outside" them have any meaningful interpretation in this framework [29]. In fact, following the same principle, each vector $v = [v_1, \ldots, v_A, \ldots, v_\Theta]' \in \mathbb{R}^{N-1}$ can be thought of as a function $\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ s.t. $\varsigma(A) = v_A$. Each of these functions ς has a Moebius inverse $m_\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ such that

$$\varsigma(A) = \sum_{B \subset A} m_\varsigma(B)$$

i.e., *each* vector ς of $\mathbb{R}^{N-1}$ can be thought of as a *sum function* (see [31] for a brief introduction). However, m_ς does not in general meet the positivity constraint: $m_\varsigma(A) \not\geq 0 \, \forall A \subset \Theta$.

The section $\{v \in \mathbb{R}^{N-1} : v_\Theta = 1\}$ of $\mathbb{R}^{N-1}$ corresponds to the constraint $\varsigma(\Theta) = 1$, so that all points of this section are sum functions meeting the normalization axiom

$$\sum_{A \subset \Theta} m_\varsigma(A) = 1.$$

Normalized sum functions (N.S.F.s) are natural extensions of b.f.s in this geometric framework. Analogous to the case of b.f.s, we call *Bayesian n.s.f.* any n.s.f. ς such that

$$\sum_{x \in \Theta} m_\varsigma(x) = 1. \tag{9}$$

IV. Belief and Probability in the Binary Case

It may be helpful to visually render these concepts in a simple example. Fig. 2 shows the geometry of belief and plausibility spaces for a binary frame $\Theta_2 = \{x, y\}$. As $|\Theta| = 2$, b.f. and pl.f. "are" vectors $[v_x, v_y, v_\Theta]'$ of a space with $N - 1 = 2^2 - 1 = 3$ dimensions. However, since $b(\Theta) = pl_b(\Theta) = 1$ for all b, we can neglect the component $v_\Theta \equiv 1$ and represent belief and plausibility vectors as points of a plane with coordinates

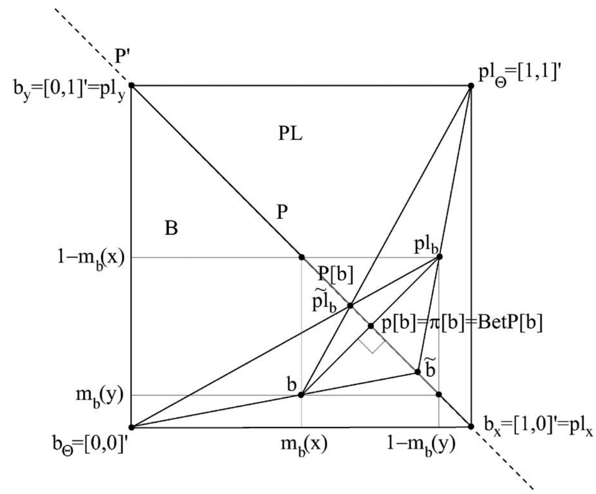

Fig. 2. In a binary frame $\Theta_2 = \{x, y\}$, both belief $\mathcal{B}$ and plausibility $\mathcal{PL}$ space are simplices with vertices $\{b_\Theta = [0,0]', b_x = [1,0]', b_y = [0,1]'\}$ and $\{pl_\Theta = [1,1]', pl_x = b_x, pl_y = b_y\}$, respectively. A b.f. b and the corresponding pl.f. pl_b are always located in symmetric positions with respect to the segment $\mathcal{P}$ of probabilities on Θ_2. The associated relative plausibility $\tilde{pl}_b$ and belief $\tilde{b}$ of singletons are shown as the intersections of the probabilistic subspace with the line joining pl_b and $b_\Theta = [0,0]'$ and the line passing through b and b_Θ, respectively. The other Bayesian functions related to b all coincide with the center of the segment of consistent probabilities $\mathcal{P}[b]$.

$$\begin{aligned} b &= [b(x) = m_b(x), b(y) = m_b(y)]' \\ pl_b &= [pl_b(x) = 1 - m_b(y), pl_b(y) = 1 - m_b(x)]'. \end{aligned}$$

In this case, the b.pl.a. of b is $\mu_b(x) = (-1)^2 \sum_{B \supset x} m_b(B) = m_b(x) + m_b(\Theta) = pl_b(x)$, $\quad \mu_b(y) = (-1)^2 \sum_{B \supset y} m_b(B) = m_b(y) + m_b(\Theta) = pl_b(y)$, and $pl_b = pl_b(x) b_x + pl_b(y) b_y$. We can notice that both $\mathcal{B}$ and $\mathcal{PL}$ are symmetric with respect to the Bayesian region $\mathcal{P}$. Furthermore, each pair of functions (b, pl_b) determines a line *orthogonal* to $\mathcal{P}$, where b and pl_b lie on symmetric positions on the two sides of the Bayesian segment $\mathcal{P} = Cl(b_x, b_y)$.

Let us denote with $a(v_1, \ldots, v_k)$ the affine subspace of a Cartesian space $\mathbb{R}^m$ generated by some points $v_1, \ldots, v_k \in \mathbb{R}^m$, i.e., the set $\{v \in \mathbb{R}^m : v = \alpha_1 v_1 + \cdots + \alpha_k v_k, \sum_i \alpha_i = 1\}$.

In the binary case, the plane $\mathbb{R}^2$ in which $\mathcal{B}, \mathcal{PL}$ lie is the affine space of n.s.f.s on Θ_2. The region $\mathcal{P}'$ of all Bayesian n.s.f. is obviously (9) the line

$$\mathcal{P}' = \{\varsigma \in \mathbb{R}^2 : m_\varsigma(x) + m_\varsigma(y) = 1\} = a(\mathcal{P})$$

and coincides with the affine space $a(\mathcal{P}) = a(b_x, x \in \Theta)$ generated by $\mathcal{P}$.

Consider now the set of probabilities $\mathcal{P}[b]$ dominating b (*consistent* probabilities), i.e., the Bayesian b.f. such that $p(A) \geq b(A) \; \forall A \subset \Theta$. In the simple binary case, the probabilities consistent with b form a segment (1-D simplex) in $\mathcal{P}$ (see Fig. 2) whose center of mass $\overline{\mathcal{P}}$ is known [23], [32], [33] to be Smets' *pignistic function* [34], [35]

$$\begin{aligned} BetP[b] &= \sum_{x\in\Theta} b_x \sum_{A\supset x} \frac{m_b(A)}{|A|} \\ &= b_x \left(m_b(x) + \frac{m_b(\Theta)}{2} \right) + b_y \left(m_b(y) + \frac{m_b(\Theta)}{2} \right). \end{aligned} \tag{10}$$

We can notice however that it also coincides with the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$, and the intersection $p[b]$ of the line $a(b, pl_b)$ with the Bayesian simplex $\mathcal{P}$

$$p[b] = \pi[b] = BetP[b] = \overline{\mathcal{P}}[b].$$

Epistemic notions like consistency and pignistic transformation seem then to be related to geometric properties such as orthogonality. It is natural to wonder whether this is true in general or is just an artifact of the binary frame.

It is worth to notice incidentally that the *relative plausibility of singletons* $\tilde{pl}_b$ [13]

$$\tilde{pl}_b(x) \doteq \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)} \tag{11}$$

although consistent with b does *not* follow the same scheme. The same can be said of the *relative belief* of singletons, i.e., the Bayesian function

$$\tilde{b}(x) \doteq \frac{m_b(x)}{\sum_{y\in\Theta} m_b(y)}$$

assigning to each singleton x its normalized mass (see Fig. 2). We will consider their behavior separately in the near future [36].

In the following, we will instead study two other geometric loci related to b, in particular the line $a(b, pl_b)$ and the orthogonal complement $\mathcal{P}^\perp$ of $\mathcal{P}$, and introduce the two Bayesian b.f.s associated with them, i.e., orthogonal projection $\pi[b]$ and intersection probability $p[b]$. We will compare them with both pignistic function and relative plausibility of singletons, and with each other. We will provide interpretations of $\pi[b]$, $p[b]$ in terms of degrees of belief and discuss their behavior with respect to affine combination.

V. Geometry of the Dual Line

Let us then first consider the "dual line" connecting a pair of belief and plausibility measures supporting the same evidence. As a matter of fact, orthogonality turns out to be a general feature of $a(b, pl_b)$. As we just saw in the binary case, $b(\Theta) = pl_b(\Theta) = 1\ \forall b$, so that we can consider b, pl_b as points of $\mathbb{R}^{N-2}$.

A. Orthogonality

Let us consider the affine subspace $a(\mathcal{P}) = a(b_x, x \in \Theta)$ generated by the simplex of Bayesian b.f.s. This can be written as the translated version of a vector space

$$a(\mathcal{P}) = b_x + span(b_y - b_x\ \forall y \in \Theta, y \neq x)$$

where $span(b_y - b_x)$ denotes the vector space generated by the $n-1$ vectors $b_y - b_x$ ($n = |\Theta|$). After recalling that, by definition

$$b_B(A) = \begin{cases} 1, & A \supset B \\ 0, & \text{else} \end{cases} \tag{12}$$

we can point out that these vectors show a rather peculiar symmetry

$$b_y - b_x(A) = \begin{cases} 1, & A \supset \{y\}, A \not\supset \{x\} \\ 0, & A \supset \{x\},\{y\} \text{ or } A \not\supset \{x\},\{y\} \\ -1, & A \not\supset \{y\}, A \supset \{x\} \end{cases} \tag{13}$$

that can be usefully exploited.

Lemma 1: $[b_y - b_x](A^c) = -[b_y - b_x](A)\ \forall A \subset \Theta$.

Proof: By (12) $[b_y - b_x](A) = 1 \Rightarrow A \supset \{y\}, A \not\supset \{x\} \Rightarrow A^c \supset \{x\}, A^c \not\supset \{y\} \Rightarrow [b_y - b_x](A^c) = -1$ and vice-versa. On the other side, $[b_y - b_x](A) = 0 \Rightarrow A \supset \{y\}, A \supset \{x\}$ or $A \not\supset \{y\}$, $A \not\supset \{x\}$. In the first case, $A^c \not\supset \{x\},\{y\}$, and in the second one, $A^c \supset \{x\},\{y\}$. In both cases, $[b_y - b_x](A^c) = 0$. ■

Theorem 2: The line connecting pl_b and b in $\mathbb{R}^{N-2}$ is orthogonal to the affine space generated by the probabilistic simplex, i.e., $b - pl_b \perp a(\mathcal{P})$.

Proof [4]*:* Having denoted with X_A the Ath axis of the orthonormal reference frame $\{X_A : A \neq \Theta, \emptyset\}$ in $\mathbb{R}^{N-2}$ (see Section III), we can write their difference as

$$pl_b - b = \sum_{\emptyset \subsetneq A \subsetneq \Theta} [pl_b(A) - b(A)]\, X_A$$

where

$$\begin{aligned} [pl_b - b](A^c) &= pl_b(A^c) - b(A^c) \\ &= 1 - b(A) - b(A^c) \\ &= 1 - b(A^c) - b(A) \\ &= pl_b(A) - b(A) \\ &= [pl_b - b](A). \end{aligned} \tag{14}$$

The scalar product $\langle \cdot, \cdot \rangle$ between the vector $pl_b - b$ and the basis vectors of $a(\mathcal{P})$ is then

$$\langle pl_b - b, b_y - b_x \rangle = \sum_{\emptyset \subsetneq A \subsetneq \Theta} [pl_b - b](A) \cdot [b_y - b_x](A)$$

which by (14) becomes

$$\sum_{|A| \leq \lfloor |\Theta|/2 \rfloor, A \neq \emptyset} [pl_b - b](A) \Big\{ [b_y - b_x](A) + [b_y - b_x](A^c) \Big\}$$

whose addenda are all nil by Lemma 1. ■

[4] In fact, the proof is valid for $A = \Theta, \emptyset$ too.

B. Intersection With the Region of Bayesian N.S.F.s

One might be tempted to conclude that since $a(b, pl_b)$ and $\mathcal{P}$ are always orthogonal, their intersection *is* the orthogonal projection of b onto $\mathcal{P}$ as in the binary case. Unfortunately, this is not the case for in general they *do not intersect* each other.

As a matter of fact, b and pl_b belong to a (2^{n-2})-dimensional Euclidean space, while the dimension of $\mathcal{P}$ is only $n-1$. If $n = 2$, $n - 1 = 1$ and $2^n - 2 = 2$ so that $a(\mathcal{P})$ divides the plane into two half-planes with b on one side and pl_b on the other side (see Fig. 2).

Formally, for a point on the line $a(b, pl_b)$ to be a probability, we need to find a value of α such that $b + \alpha(pl_b - b) \in \mathcal{P}$. Its components obviously are $b(A) + \alpha[pl_b(A) - b(A)]$ for any subset $A \subset \Theta, A \neq \Theta, \emptyset$ and in particular when $A = \{x\}$ is a singleton

$$b(x)+\alpha\left[pl_b(x)-b(x)\right] = b(x)+\alpha\left[1-b(x^c)-b(x)\right]. \quad (15)$$

A necessary condition for this point to belong to $\mathcal{P}$ is the normalization constraint for singletons

$$\sum_{x\in\Theta} b(x) + \alpha \sum_{x\in\Theta} (1 - b(x^c) - b(x)) = 1$$

$$\Rightarrow \alpha = \frac{1 - \sum_{x\in\Theta} b(x)}{\sum_{x\in\Theta}(1 - b(x^c) - b(x))} \doteq \beta[b] \quad (16)$$

which yields a single candidate value $\beta[b]$ for the line coordinate of the intersection.

Using the terminology in Section III-C, the candidate projection

$$\varsigma[b] \doteq b + \beta[b](pl_b - b) = a(b, pl_b) \cap \mathcal{P}' \quad (17)$$

(having called $\mathcal{P}'$ the set of all Bayesian n.s.f.s in $\mathbb{R}^{N-2}$) is *a Bayesian* n.s.f. but is not guaranteed to be a Bayesian b.f. For n.s.f.s, the condition $\sum_{x\in\Theta} m_\varsigma(x) = 1$ implies $\sum_{|A|>1} m_\varsigma(A) = 0$, so that $\mathcal{P}'$ can be written as

$$\mathcal{P}' = \left\{ \varsigma = \sum_{A\subset\Theta} m_\varsigma(A) b_A \in \mathbb{R}^{N-2} : \sum_{|A|=1} m_\varsigma(A) = 1, \right.$$
$$\left. \sum_{|A|>1} m_\varsigma(A) = 0 \right\}. \quad (18)$$

Theorem 3: The coordinates of $\varsigma[b]$ with respect to the basis Bayesian b.f.s $\{b_x, x \in \Theta\}$ can be expressed in terms of the b.p.a. m_b of b as

$$m_{\varsigma[b]}(x) = m_b(x) + \beta[b] \sum_{A\supset x, A\neq x} m_b(A) \quad (19)$$

where

$$\beta[b] = \frac{1 - \sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta}(pl_b(x) - m_b(x))} = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}. \quad (20)$$

Proof: The numerator of (16) is trivially $\sum_{|B|>1} m(B)$. On the other side

$$1 - b(x^c) - b(x) = \sum_{B\subset\Theta} m_b(B) - \sum_{B\subset x^c} m_b(B) - m_b(x)$$
$$= \sum_{B\supset x, B\neq x} m_b(B)$$

so that the denominator of $\beta[b]$ becomes

$$\sum_{y\in\Theta} [pl_b(y) - b(y)] = \sum_{y\in\Theta} (1 - b(y^c) - b(y))$$
$$= \sum_{y\in\Theta} \sum_{B\supset y, B\neq y} m_b(B)$$
$$= \sum_{|B|>1} m_b(B)|B|$$

yielding (20). Equation (19) comes directly from (15) when we recall that $b(x) = m_b(x), \varsigma(x) = m_\varsigma(x)\ \forall x \in \Theta$. ■

Equation (19) ensures that $m_{\varsigma[b]}(x)$ is positive for each $x \in \Theta$. A symmetric version can be obtained after realizing that $(\sum_{|B|=1} m_b(B) / \sum_{|B|=1} m_b(B)|B|) = 1$, so that we can write

$$m_{\varsigma[b]}(x) = b(x) \frac{\sum_{|B|=1} m_b(B)}{\sum_{|B|=1} m_b(B)|B|}$$
$$+ [pl_b - b](x) \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}. \quad (21)$$

It is easy to prove that the line $a(b, pl_b)$ intersects the probabilistic subspace *only for 2-additive* b.f.s (the proof can be found in the Appendix).

Theorem 4: $\varsigma[b] \in \mathcal{P}$ if and only if (iff) b is 2-additive, i.e., $m_b(A) = 0 |A| > 2$, and in this case, pl_b is the reflection of b through $\mathcal{P}$.

For 2-additive b.f.s, $\varsigma[b]$ is nothing but the *mean probability* function $(b + pl_b)/2$. In the general case however, the reflection of b through $\mathcal{P}$ not only does not coincide with pl_b but is also not even a pl.f. [37].

VI. Intersection Probability

We have seen that although the line $a(b, pl_b)$ is always orthogonal to $\mathcal{P}$, it does not intersect the probabilistic subspace in general, but it does intersect the region of Bayesian n.s.f.s in $\varsigma[b]$ (17). But of course (since $\sum_x m_{\varsigma[b]}(x) = 1$) $\varsigma[b]$ is naturally associated with a Bayesian b.f., assigning an equal amount of mass to each singleton and 0 to each $A : |A| > 1$, namely

$$p[b] \doteq \sum_{x\in\Theta} m_{\varsigma[b]}(x) b_x \quad (22)$$

where $m_{\varsigma[b]}(x)$ is given by (19). It is easy to see that $p[b]$ is a probability, since by definition $m_{p[b]}(A) = 0$ for $|A| > 1$, $m_{p[b]}(x) = m_{\varsigma[b]}(x) \geq 0 \quad \forall x \in \Theta$, and $\sum_{x\in\Theta} m_{p[b]}(x) = \sum_{x\in\Theta} m_{\varsigma[b]}(x) = 1$ by construction. We call $p[b]$ the

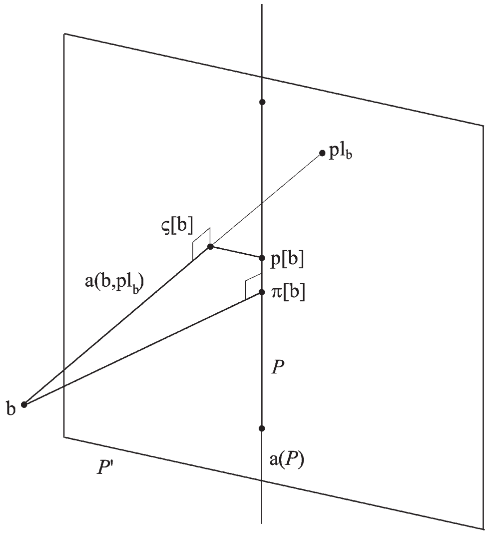

Fig. 3. Geometry of the line $a(b, pl_b)$ and relative locations of $p[b]$, $\varsigma[b]$, and $\pi[b]$. Each b.f. b and the related pl.f. pl_b lie on opposite sides of the hyperplane $\mathcal{P}'$ of the Bayesian n.s.f. that divides $\mathbb{R}^{N-2}$ into two parts. The line $a(b, pl_b)$ connecting them always intersects $\mathcal{P}'$ but not necessarily $a(\mathcal{P})$ (vertical line). This intersection $\varsigma[b]$ is naturally associated with a probability $p[b]$ (in general distinct from the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$) having the same components in the base $\{b_x, x \in \Theta\}$ of $a(\mathcal{P})$. $\mathcal{P}$ is a simplex (a segment in the figure) in $a(\mathcal{P}) : \pi[b]$ and $p[b]$ are both "true" probabilities.

intersection probability. The geometry of $\varsigma[b]$ and $p[b]$ with respect to the regions of Bayesian b.f. and n.s.f. is sketched in Fig. 3.

A. Interpretations

1) Non-Bayesianity Flag and Relative Plausibility: A first interpretation of this new probability is immediate after noticing that

$$\beta[b] = \frac{1 - \sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta} pl_b(x) - \sum_{x\in\Theta} m_b(x)} = \frac{1 - k_{\tilde{b}}}{k_{\tilde{pl}_b} - k_{\tilde{b}}}$$

where

$$k_{\tilde{b}} = \sum_{x\in\Theta} m_b(x)$$
$$k_{\tilde{pl}_b} = \sum_{x\in\Theta} pl_b(x) = \sum_{A\subset\Theta} m_b(A)|A|$$

are the normalization factors for $\tilde{b}$ and $\tilde{pl}_b$, respectively, so that $p[b]$ can be rewritten as

$$p[b](x) = m_b(x) + (1 - k_{\tilde{b}})\frac{pl_b(x) - m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}}. \tag{23}$$

When b is Bayesian, $pl_b(x) - m_b(x) = 0\ \forall x \in \Theta$. If b is not Bayesian, there exists at least a singleton x such that $pl_b(x) - m_b(x) > 0$. The Bayesian b.f.

$$R[b](x) \doteq \frac{\sum_{A\supset x, A\neq x} m_b(A)}{\sum_{|A|>1} m_b(A)|A|} = \frac{pl_b(x) - m_b(x)}{\sum_{y\in\Theta}(pl_b(y) - m_b(y))}$$

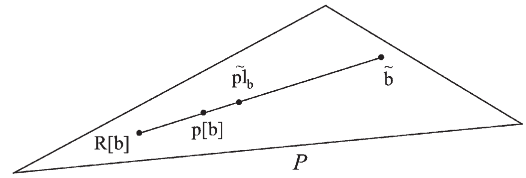

Fig. 4. Location of intersection probability $p[b]$ and relative plausibility of singletons $\tilde{pl}_b$ with respect to the non-Bayesianity flag $R[b]$. They both lie on the segment joining $R[b]$ and the relative belief of singletons $\tilde{b}$, but $\tilde{pl}_b$ is closer to $\tilde{b}$ than $p[b]$.

then measures the relative contribution of each singleton x to the non-Bayesianity of b. Equation (23) shows in fact that the non-Bayesian mass $1 - k_{\tilde{b}}$ is assigned by $p[b]$ to each singleton according to its relative contribution $R[b](x)$ to the non-Bayesianity of b.

The flag probability $R[b]$ also relates the intersection probability $p[b]$ to other two classical Bayesian approximations, i.e., the relative plausibility $\tilde{pl}_b$ and belief $\tilde{b}$ of singletons, as (23) reads as

$$p[b] = k_{\tilde{b}}\tilde{b} + (1 - k_{\tilde{b}})R[b]. \tag{24}$$

Geometrically, since $k_{\tilde{b}} = \sum_{x\in\Theta} m_b(x) \leq 1$, $p[b]$ belongs to the segment linking $R[b]$ with the relative belief of singletons $\tilde{b}$ with convex coordinate the total mass of singletons $k_{\tilde{b}}$. But now, the relative pl.f. can also be written in terms of $\tilde{b}$ and $R[b]$ as by definition

$$\begin{aligned} R[b](x) &= \frac{pl_b(x) - m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \\ &= \frac{pl_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} - \frac{m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \\ &= \tilde{pl}_b(x)\frac{k_{\tilde{pl}_b}}{k_{\tilde{pl}_b} - k_{\tilde{b}}} - \tilde{b}(x)\frac{k_{\tilde{b}}}{k_{\tilde{pl}_b} - k_{\tilde{b}}} \end{aligned}$$

since $\tilde{pl}_b(x) = pl_b(x)/k_{\tilde{pl}_b}$, and $\tilde{b}(x) = m_b(x)/k_{\tilde{b}}$, so that

$$\tilde{pl}_b = \left(\frac{k_{\tilde{b}}}{k_{\tilde{pl}_b}}\right)\tilde{b} + \left(1 - \frac{k_{\tilde{b}}}{k_{\tilde{pl}_b}}\right)R[b].$$

Both $\tilde{pl}_b$ and $p[b]$ belong to $Cl(R[b], \tilde{b})$. However, as $k_{\tilde{pl}_b} = \sum_{A\subset\Theta} m_b(A)|A| \geq 1$, $k_{\tilde{b}}/k_{\tilde{pl}_b} \leq k_{\tilde{b}}$, which in turn implies that $p[b]$ is closer to $R[b]$ than the relative pl.f. $\tilde{pl}_b$ (see Fig. 4). The convex coordinate of $\tilde{pl}_b$ in $Cl(R[b], \tilde{b})$ measures the ratio between total mass and plausibility of singletons. Obviously, when $k_{\tilde{b}} = 0$ ($\tilde{b}$ does not exists), $p[b] = \tilde{pl}_b = R[b]$ by (23).

2) Meaning of the Ratio $\beta[b]$ and Pignistic Function: To shed more light on $p[b]$ and get an alternative interpretation of the intersection probability, it is useful to compare $p[b]$ as expressed in (23) with another classical Bayesian approximation of b, i.e., the pignistic function

$$BetP[b](x) \doteq \sum_{A\supset x} \frac{m_b(A)}{|A|} = m_b(x) + \sum_{A\supset x, A\neq x} \frac{m_b(A)}{|A|}.$$

We can notice that in $BetP[b]$, the mass of each event A, $|A| > 1$ is considered *separately*, and its mass $m_b(A)$ is *equally* shared among the elements of A. In $p[b]$, instead, it is the *total* mass $\sum_{|A|>1} m_b(A) = 1 - k_{\tilde{b}}$ of nonsingletons that is considered, and this total mass is distributed *proportionally* to their non-Bayesian contribution to each element of Θ.

How should $\beta[b]$ be interpreted then? If we write $p[b](x)$ as

$$p[b](x) = m_b(x) + \beta[b]\left(pl_b(x) - m_b(x)\right) \tag{25}$$

we can observe that a fraction measured by $\beta[b]$ of its non-Bayesian contribution $pl_b(x) - m_b(x)$ is *uniformly* assigned to each singleton. This leads to another parallelism between $p[b]$ and $BetP[b]$. It suffices to note that if $|A| > 1$

$$\beta[b_A] = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} = \frac{1}{|A|}$$

so that both $p[b](x)$ and $BetP[b](x)$ assume the form

$$m_b(x) + \sum_{A \supset x, A \neq x} m_b(A)\beta_A$$

where $\beta_A = const = \beta[b]$ for $p[b]$, while $\beta_A = \beta[b_A]$ in case of the pignistic function.

Under which condition $p[b]$ and pignistic function coincide? A sufficient condition can be achieved by decomposing $\beta[b]$ as

$$\begin{aligned}\beta[b] &= \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} \\ &= \frac{\sum_{k=2}^{n}\sum_{|B|=k} m_b(B)}{\sum_{k=2}^{n}\left(k\sum_{|B|=k} m_b(B)\right)} \\ &= \frac{\sigma^2 + \cdots + \sigma^n}{2\sigma^2 + \cdots + n\sigma^n}\end{aligned} \tag{26}$$

after defining $\sigma^k \doteq \sum_{|B|=k} m_b(B)$.

Theorem 5: Intersection probability and pignistic function coincide if $\exists k \in [2, \ldots, n]$ such that $\sigma^i = 0\ \forall i \neq k$, i.e., the focal elements of b have size 1 or k only.

Proof: $p[b] = BetP[b]$ is equivalent to

$$\begin{aligned}m_b(x) + \sum_{A \supset x, A \neq x} m_b(A)\beta[b] &= m_b(x) + \sum_{A \supset x, A \neq x} \frac{m_b(A)}{|A|} \\ &\equiv \sum_{A \supset x, A \neq x} m_b(A)\beta[b] \\ &= \sum_{A \supset x, A \neq x} \frac{m_b(A)}{|A|}.\end{aligned}$$

If $\exists k : m_b(A) = 0$ for $|A| \neq k$, then $\beta[b] = 1/k$, and the equality is met. ■

In particular, this is true when $\Sigma^i = 0$, $i > 2$, i.e., when b is 2-additive. The condition of Theorem 5 is in fact a rather straightforward generalization of the concept of 2-additivity.

3) Example: Let us see a simple example to briefly discuss the two interpretations of $p[b]$ introduced above. Consider a ternary frame $\Theta = \{x, y, z\}$, and a b.f. b with b.p.a. given by

$$\begin{aligned} m_b(x) &= 0.1 & m_b(y) &= 0 \\ m_b(z) &= 0.2 & m_b(\{x,y\}) &= 0.3 \\ m_b(\{x,z\}) &= 0.1 & m_b(\{y,z\}) &= 0 \\ m_b(\Theta) &= 0.3. \end{aligned}$$

Recalling (23), the total mass of singletons is $k_{\tilde{b}} = 0.1 + 0 + 0.2 = 0.3$, while the non-Bayesian contributions of x, y, z are respectively

$$\begin{aligned} pl_b(x) - m_b(x) &= m_b(\Theta) + m_b(\{x,y\}) + m_b(\{x,z\}) = 0.7 \\ pl_b(y) - m_b(y) &= m_b(\{x,y\}) + m_b(\Theta) = 0.6 \\ pl_b(z) - m_b(z) &= m_b(\{x,z\}) + m_b(\Theta) = 0.4 \end{aligned}$$

so that the non-Bayesian flag has values $R(x) = 0.7/1.7$, $R(y) = 0.6/1.7$, $R(z) = 0.4/1.7$.

For each singleton then, the original b.p.a. $m_b(x)$ is increased by a share of the mass of nonsingletons $1 - k_{\tilde{b}} = 0.7$ proportional to the value of $R(x)$, i.e.,

$$\begin{aligned} p[b](x) &= m_b(x) + (1 - k_{\tilde{b}})R(x) \\ &= 0.1 + 0.7 * 0.7/1.7 \\ &= 0.388 \\ p[b](y) &= m_b(y) + (1 - k_{\tilde{b}})R(y) \\ &= 0 + 0.7 * 0.6/1.7 \\ &= 0.247 \\ p[b](z) &= m_b(z) + (1 - k_{\tilde{b}})R(z) \\ &= 0.2 + 0.7 * 0.4/1.7 \\ &= 0.365. \end{aligned}$$

Equivalently, the line coordinate $\beta[b]$ of $p[b]$ is equal to

$$\begin{aligned} &\frac{1 - k_{\tilde{b}}}{m_b(\{x,y\})\,|\{x,y\}| + m_b(\{x,z\})\,|\{x,z\}| + m_b(\Theta)|\Theta|} \\ &= \frac{0.7}{0.3 * 2 + 0.1 * 2 + 0.3 * 3} = \frac{0.7}{1.7} \end{aligned}$$

and measures the share of $pl_b(x) - m_b(x)$ assigned to each singleton

$$\begin{aligned} p[b](x) &= m_b(x) + \beta[b]\left(pl_b(x) - m_b(x)\right) \\ &= 0.1 + 0.7/1.7 * 0.7 \\ p[b](y) &= m_b(y) + \beta[b]\left(pl_b(y) - m_b(y)\right) \\ &= 0 + 0.7/1.7 * 0.6 \\ p[b](z) &= m_b(z) + \beta[b]\left(pl_b(z) - m_b(z)\right) \\ &= 0.2 + 0.7/1.7 * 0.4. \end{aligned}$$

VII. Orthogonal Projection

Although the intersection of the line $a(b, pl_b)$ with the region $\mathcal{P}'$ of the Bayesian n.s.f. is not always in $\mathcal{P}$, an orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ is obviously guaranteed to exist as $a(\mathcal{P})$ is nothing but a linear subspace in the space of n.s.f.s (such as b). An explicit calculation of $\pi[b]$, however, requires a description of the orthogonal complement of $a(\mathcal{P})$ in $\mathbb{R}^{N-2}$. Let us denote with $n = |\Theta|$ the cardinality of Θ.

A. Orthogonality Condition

We need to find a necessary and sufficient condition for an arbitrary vector $v = \sum_{A\subset\Theta} v_A X_A$ to be orthogonal[5] to the probabilistic subspace $a(\mathcal{P})$. If we compute the scalar product $\langle v, b_y - b_x\rangle$ between v and the generators $b_y - b_x$ of $a(P)$, we get

$$\left\langle \sum_{A\subset\Theta} v_A X_A, b_y - b_x \right\rangle = \sum_{A\subset\Theta} v_A [b_y - b_x](A)$$

which remembering (13) becomes

$$\langle v, b_y - b_x\rangle = \sum_{A\supset y, A\not\supset x} v_A - \sum_{A\supset x, A\not\supset y} v_A.$$

The orthogonal complement $a(\mathcal{P})^\perp$ of $a(\mathcal{P})$ can then be expressed as

$$v(\mathcal{P})^\perp = \left\{ v : \sum_{A\supset y, A\not\supset x} v_A = \sum_{A\supset x, A\not\supset y} v_A \ \forall y \neq x \right\}.$$

If the vector v in particular is a b.f. ($v_A = b(A)$)

$$\sum_{A\supset y, A\not\supset x} b(A) = \sum_{A\supset y, A\not\supset x} \sum_{B\subset A} m_b(B) = \sum_{B\subset\{x\}^c} m_b(B) 2^{n-1-|B\cup\{y\}|}$$

since $2^{n-1-|B\cup\{y\}|}$ is the number of subsets A of $\{x\}^c$ containing both B and y, and the orthogonality condition becomes

$$\sum_{B\subset\{x\}^c} m_b(B) 2^{n-1-|B\cup\{y\}|} = \sum_{B\subset\{y\}^c} m_b(B) 2^{n-1-|B\cup\{x\}|}, \quad \forall y \neq x.$$

Now, sets $B \subset \{x, y\}^c$ appear in both summations with the same coefficient (since in that case $|B \cup \{x\}| = |B \cup \{y\}| = |B| + 1$), and the equation, after erasing the common factor 2^{n-2}, reduces to

$$\sum_{B\supset y, B\not\supset x} m_b(B) 2^{1-|B|} = \sum_{B\supset x, B\not\supset y} m_b(B) 2^{1-|B|}, \qquad \forall y \neq x \tag{27}$$

which expresses the desired orthogonality condition.

[5] The proof is again valid for $A = \Theta, \emptyset$ too. See Section VIII-A.

Theorem 6: The orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ can be expressed in terms of the b.p.a. m_b of b as

$$\pi[b](x) = \sum_{A\supset x} m_b(A) 2^{1-|A|} + \sum_{A\subset\Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right) \tag{28}$$

$$\pi[b](x) = \sum_{A\supset x} m_b(A) \left(\frac{1 + |A^c| 2^{1-|A|}}{n} \right) + \sum_{A\not\supset x} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right). \tag{29}$$

Equation (29) shows that $\pi[b]$ *is indeed a probability*, since both $1 + |A^c| 2^{1-|A|} \geq 0$ and $1 - |A| 2^{1-|A|} \geq 0 \ \ \forall |A| = 1, \ldots, n$. This is not at all trivial, as $\pi[b]$ is the projection of b onto the *affine* space $a(\mathcal{P})$ and could have in principle assigned negative masses to one or more singletons. $\pi[b]$ is hence another valid candidate to the role of the probabilistic approximation of b.f. b.

B. Orthogonality Flag

Theorem 6 does not apparently provide any intuition about the meaning of $\pi[b]$ in terms of degrees of belief. In fact, if we process (29), we can reduce π to a new Bayesian function strictly related to the pignistic function.

Theorem 7: $\pi[b] = \overline{\mathcal{P}}(1 - k_O[b]) + k_O[b] O[b]$, where $\overline{\mathcal{P}}$ is the uniform probability, and

$$O[b](x) = \frac{\bar{O}[b](x)}{k_O[b]} = \frac{\sum_{A\supset x} m_b(A) 2^{1-|A|}}{\sum_{A\subset\Theta} m_b(A) |A| 2^{1-|A|}} = \frac{\sum_{A\supset x} \frac{m_b(A)}{2^{|A|}}}{\sum_{A\subset\Theta} \frac{m_b(A)|A|}{2^{|A|}}} \tag{30}$$

is a Bayesian b.f.

As $0 \leq |A| 2^{1-|A|} \leq 1$ for all $A \subset \Theta$, $k_O[b]$ assumes values in the interval [0, 1]. Theorem 7 then implies that the orthogonal projection is always located on the line segment $Cl(\overline{\mathcal{P}}, O[b])$ joining the uniform, noninformative probability, and the Bayesian function $O[b]$.

By (30), it turns out that $\pi[b] = \overline{\mathcal{P}}$ iff $O[b] = \overline{\mathcal{P}}$ (since $k_O[b] > 0$). The meaning of $O[b]$ becomes clear when noticing that condition (27) (under which a b.f. b is orthogonal to $a(\mathcal{P})$) can be rewritten as

$$\begin{aligned} \sum_{B\supset y, B\not\supset x} m_b(B) 2^{1-|B|} + \sum_{B\supset y, x} m_b(B) 2^{1-|B|} \\ = \sum_{B\supset x, B\not\supset y} m_b(B) 2^{1-|B|} + \sum_{B\supset y, x} m_b(B) 2^{1-|B|} \\ \equiv \sum_{B\supset y} m_b(B) 2^{1-|B|} = \sum_{B\supset x} m_b(B) 2^{1-|B|} \\ \equiv \bar{O}[b](x) = const \\ \equiv O[b](x) = const = \overline{\mathcal{P}} \qquad \forall x \in \Theta. \end{aligned}$$

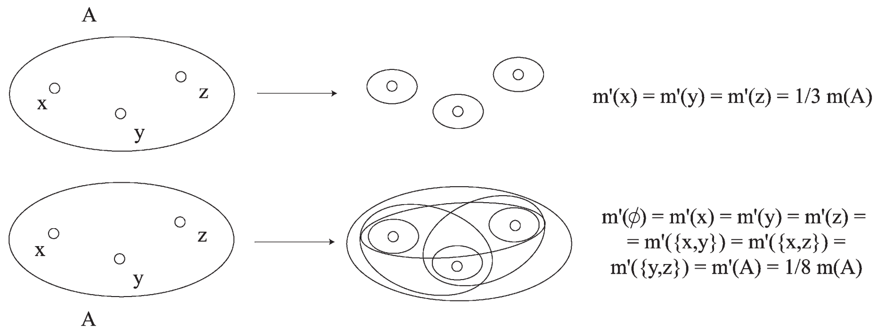

Fig. 5. Redistribution processes associated with pignistic transformation and orthogonal projection. (Top) In the pignistic transformation, the mass of each focal element A is distributed among its elements. (Bottom) In the orthogonal projection instead (through the orthogonality flag), the mass of each f.e. A is divided among all its subsets $B \subset A$. In both cases, the related relative plausibility of singletons yields a Bayesian b.f.

Therefore, $\pi[b] = \overline{\mathcal{P}}$ iff $b \perp a(\mathcal{P})$, and $O - \overline{\mathcal{P}}$ measures the nonorthogonality of b with respect to $\mathcal{P}$. $O[b]$ then deserves the name of *orthogonality flag*.

C. Interpretation in Terms of Plausibilities and Redistribution Processes

A compelling link can be drawn between orthogonal projection and pignistic function by means of the orthogonality flag $O[b]$. Let us define the two b.f.s

$$b_{\|} \doteq \frac{1}{k_{\|}} \sum_{A \subset \Theta} \frac{m_b(A)}{|A|} b_A$$

$$b_{2\|} \doteq \frac{1}{k_{2\|}} \sum_{A \subset \Theta} \frac{m_b(A)}{2^{|A|}} b_A$$

where $k_{\|}$ and $k_{2\|}$ are the normalization factors needed to make them two admissible b.f.

Theorem 8: $O[b]$ is the relative plausibility of singletons of $b_{2\|}$, and $BetP[b]$ is the relative plausibility of singletons of $b_{\|}$.

Proof: By definition of pl.f.

$$pl_{b_{2\|}}(x) = \sum_{A \supset x} m_{b_{2\|}}(A) = \frac{1}{k_{2\|}} \sum_{A \supset x} \frac{m_b(A)}{2^{|A|}} = \frac{\bar{O}[b]}{2k_{2\|}}$$

$$\sum_{x \in \Theta} pl_{b_{2\|}}(x) = \frac{1}{k_{2\|}} \sum_{x \in \Theta} \sum_{A \supset x} \frac{m_b(A)}{2^{|A|}} = \frac{k_O[b]}{2k_{2\|}}$$

by (39). Hence, $\tilde{pl}_{b_{2\|}}(x) = \bar{O}[b]/k_O[b] = O[b]$. Equivalently

$$pl_{b_{\|}}(x) = \sum_{A \supset x} m_{b_{\|}}(A) = \frac{1}{k_{\|}} \sum_{A \supset x} \frac{m_b(A)}{|A|} = \frac{1}{k_{\|}} BetP[b](x)$$

and since $\sum_x BetP[b](x) = 1$, $\tilde{pl}_{b_{\|}}(x) = BetP[b](x)$. ■

The two functions $b_{\|}$ and $b_{2\|}$ represent two different processes acting on b (see Fig. 5). The first one redistributes the mass of each focal element among its *singletons* (yielding directly a Bayesian b.f. $BetP[b]$). The second one distributes the b.p.a. of each event A among its *subsets* $B \subset A$ ($\emptyset, A$ included). In this second case, we get a u.b.f. [38] b^U

$$m_{b^U}(A) = \sum_{B \supset A} \frac{m_b(B)}{2^{|B|}}$$

whose relative belief of singletons $\tilde{b}^U$ is in fact the orthogonality flag $O[b]$.

1) Example: Let us consider again as an example the b.f. b on the ternary frame seen in Section VI-A3. To get the orthogonality flag $O[b]$, we need to apply the redistribution process of Fig. 5 (bottom) to each focal element of b. In this case, their masses are divided among their subsets as

$$m(x) = 0.1 \mapsto m'(x) = m'(\emptyset) = 0.1/2 = 0.05$$
$$m(z) = 0.2 \mapsto m'(z) = m'(\emptyset) = 0.2/2 = 0.1$$
$$m(\{x,y\}) = 0.3 \mapsto m'(\{x,y\}) = m'(x) = m'(y) = m'(\emptyset) = 0.3/4 = 0.075$$
$$m(\{x,z\}) = 0.1 \mapsto m'(\{x,z\}) = m'(x) = m'(z) = m'(\emptyset) = 0.1/4 = 0.025$$
$$m(\Theta) = 0.3 \mapsto m'(\Theta) = m'(\{x,y\}) = m'(\{x,z\}) = m'(\{y,z\}) = m'(x) = m'(y) = m'(z) = m'(\emptyset) = 0.3/8 = 0.0375.$$

By summing all contributions related to singletons on the right-hand side, we get

$$m_{b^U}(x) = 0.05 + 0.075 + 0.025 + 0.0375 = 0.1875$$
$$m_{b^U}(y) = 0.075 + 0.0375 = 0.1125$$
$$m_{b^U}(z) = 0.1 + 0.025 + 0.0375 = 0.1625$$

whose sum is the normalization factor

$$k_O[b] = m_{b^U}(x) + m_{b^U}(y) + m_{b^U}(z) = 0.4625$$

so that by normalizing, we get $O[b] = [0.405, 0.243, 0.351]'$. The orthogonal projection $\pi[b]$ is finally the convex combination of $O[b]$ and $\overline{\mathcal{P}} = [1/3, 1/3, 1/3]'$ with coordinate $k_O[b]$

$$\begin{aligned}\pi[b] &= \overline{\mathcal{P}}(1 - k_O[b]) + k_O[b]O[b] \\ &= [1/3,1/3,1/3]'(1-0.4625)+0.4625[0.405, 0.243, 0.351]' \\ &= [0.366, 0.291, 0.342]'.\end{aligned}$$

D. Orthogonal Projection and Affine Combination

As a confirmation of this relationship, orthogonal projection and pignistic function both commute with affine combination.

Theorem 9: Orthogonal projection and affine combination commute, i.e., if $\alpha_1 + \alpha_2 = 1$

$$\pi[\alpha_1 b_1 + \alpha_2 b_2] = \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2].$$

Proof: By Theorem 7, $\pi[b] = (1 - k_O[b])\overline{\mathcal{P}} + \bar{O}[b]$, where $k_O[b] = \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}$, and $\bar{O}[b](x) = \sum_{A\supset x} m_b(A)2^{1-|A|}$. Hence

$$\begin{aligned}k_O[\alpha_1 b_1+\alpha_2 b_2] &= \sum_{A\subset\Theta} (\alpha_1 m_{b_1}(A)+\alpha_2 m_{b_2}(A))\,|A|2^{1-|A|} \\ &= \alpha_1 k_O[b_1]+\alpha_2 k_O[b_2], \\ \bar{O}[\alpha_1 b_1+\alpha_2 b_2](x) &= \sum_{A\supset x} (\alpha_1 m_{b_1}(A)+\alpha_2 m_{b_2}(A))\,2^{1-|A|} \\ &= \alpha_1 \bar{O}[b_1]+\alpha_2 \bar{O}[b_2]\end{aligned}$$

which in turn implies (since $\alpha_1 + \alpha_2 = 1$)

$$\begin{aligned}\pi[\alpha_1 b_1 + \alpha_2 b_2] = &\ (1 - \alpha_1 k_O[b_1] - \alpha_2 k_O[b_2])\,\overline{\mathcal{P}} \\ &+ \alpha_1 \bar{O}[b_1] + \alpha_2 \bar{O}[b_2] \\ &= \alpha_1 \left[(1 - k_O[b_1])\,\overline{\mathcal{P}} + \bar{O}[b_1]\right] \\ &\quad + \alpha_2 \left[(1 - k_O[b_2])\,\overline{\mathcal{P}} + \bar{O}[b_2]\right] \\ &= \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2].\end{aligned}$$

■

This property can be used to find an alternative expression of the orthogonal projection as the *convex combination of the pignistic functions associated with all basis b.f.s.*

Lemma 2: The orthogonal projection of a basis b.f. b_A is given by $\pi[b_A] = (1 - |A|2^{1-|A|})\overline{\mathcal{P}} + |A|2^{1-|A|}\overline{\mathcal{P}}_A$, where $\overline{\mathcal{P}}_A = (1/|A|)\sum_{x\in A} b_x$ is the center of mass of all the probabilities with support in A.

Proof: By (30), $k_O[b_A] = |A|2^{1-|A|}$, so that

$$\bar{O}[b_A](x) = \begin{cases} 2^{1-|A|}, & x \in A \\ 0, & x \notin A \end{cases} \Rightarrow O[b_A](x) = \begin{cases} \frac{1}{|A|}, & x \in A \\ 0, & x \notin A \end{cases}$$

i.e., $O[b_A] = (1/|A|)\sum_{x\in A} b_x = \overline{\mathcal{P}}_A$. ■

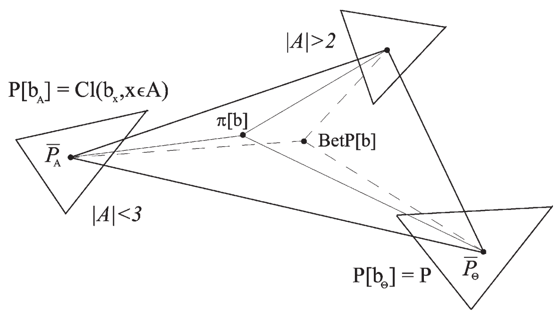

Fig. 6. Orthogonal projection $\pi[b]$ and pignistic function $BetP[b]$ are both located on the simplex whose vertices are all the basis pignistic functions, i.e., the uniform probabilities associated with each single event A. However, the convex coordinates of $\pi[b]$ are weighted by a factor $k_O[b_A] = |A|2^{1-|A|}$, yielding a point that is closer to vertices related to lower size events.

Theorem 10: The orthogonal projection can be expressed as a convex combination of all noninformative probabilities with support on a single event A as

$$\begin{aligned}\pi[b] &= \overline{\mathcal{P}}\left(1 - \sum_{A\neq\Theta}\alpha_A\right) + \sum_{A\neq\Theta}\alpha_A\overline{\mathcal{P}}_A \\ \alpha_A &\doteq m_b(A)|A|2^{1-|A|}.\end{aligned} \quad (31)$$

Proof:

$$\pi[b] = \pi\left[\sum_{A\subset\Theta} m_b(A) b_A\right] = \sum_{A\subset\Theta} m_b(A)\pi[b_A]$$

by Theorem 9, which by Lemma 2 becomes

$$\begin{aligned}&\sum_{A\subset\Theta} m_b(A)\left[(1 - |A|2^{1-|A|})\overline{\mathcal{P}} + |A|2^{1-|A|}\overline{\mathcal{P}}_A\right] \\ &= \left(1 - \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\right)\overline{\mathcal{P}} + \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\overline{\mathcal{P}}_A \\ &= \left(1 - \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\right)\overline{\mathcal{P}} + \sum_{A\neq\Theta} m_b(A)|A|2^{1-|A|}\overline{\mathcal{P}}_A \\ &\quad + m_b(\Theta)|\Theta|2^{1-|\Theta|}\overline{\mathcal{P}}\end{aligned}$$

i.e., (31). ■

As $\overline{\mathcal{P}}_A = BetP[b_A]$, we recognize that

$$\begin{aligned}BetP[b] &= \sum_{A\subset\Theta} m_b(A) BetP[b_A] \\ \pi[b] &= \sum_{A\neq\Theta}\alpha_A BetP[b_A] + \left(1 - \sum_{A\neq\Theta}\alpha_A\right) BetP[b_\Theta]\end{aligned} \quad (32)$$

with $\alpha_A = m_b(A)k_O[b_A]$. Both orthogonal projection and pignistic function are convex combinations of all basis pignistic functions. However, as $k_O[b_A] = |A|2^{1-|A|} < 1$ for $|A| > 2$, the orthogonal projection turns out to be closer to the vertices associated with events of lower cardinality (see Fig. 6).

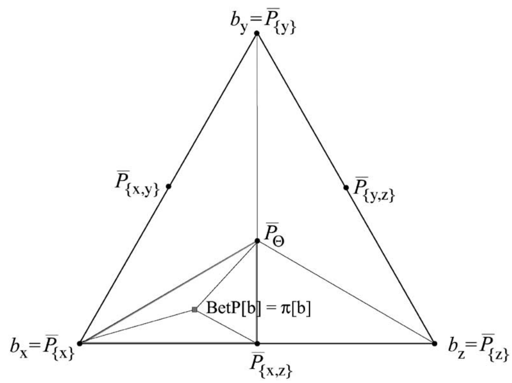

Fig. 7. Orthogonal projection and pignistic function for the b.f. (33) on the ternary frame $\Theta_3 = \{x, y, z\}$.

1) Example—Ternary Case: Let us consider as an example a ternary frame $\Theta_3 = \{x, y, x\}$ and a b.f. on Θ_3 with b.p.a.

$$\begin{aligned} m_b(x) &= 1/3 \\ m_b(\{x, z\}) &= 1/3 \\ m_b(\Theta_3) &= 1/3 \\ m_b(A) &= 0, \quad A \neq \{x\}, \{x, z\}, \Theta_3. \end{aligned} \tag{33}$$

According to (31)

$$\begin{aligned} \pi[b] &= 1/3\overline{\mathcal{P}}_{\{x\}} + 1/3\overline{\mathcal{P}}_{\{x,z\}} + (1 - 1/3 - 1/3)\overline{\mathcal{P}} \\ &= \frac{1}{3} b_x + \frac{1}{3}\frac{b_x + b_z}{2} + \frac{1}{3}\frac{b_x + b_y + b_z}{3} \\ &= b_x\left(\frac{1}{3} + \frac{1}{6} + \frac{1}{9}\right) + b_z\left(\frac{1}{6} + \frac{1}{9}\right) + b_y\frac{1}{9} \\ &= \frac{11}{18} b_x + \frac{1}{9} b_y + \frac{5}{18} b_z \end{aligned}$$

and the orthogonal projection is the barycenter of the simplex $Cl(\overline{\mathcal{P}}_{\{x\}}, \overline{\mathcal{P}}_{\{x,z\}}, \overline{\mathcal{P}})$ (see Fig. 7). On the other side

$$\begin{aligned} BetP[b](x) &= \frac{m_b(x)}{1} + \frac{m_b(x,z)}{2} + \frac{m_b(\Theta_3)}{3} = \frac{11}{18} \\ BetP[b](y) &= \frac{1}{9} \\ BetP[b](z) &= \frac{1}{6} + \frac{1}{9} = \frac{5}{18} \end{aligned}$$

i.e., $BetP[b] = \pi[b]$. This is true for each b.f. $b \in \mathcal{B}_3$, since for (32) if $|\Theta| = 3$ then $\alpha_A = m_b(A)$ for $|A| \leq 2$, and $1 - \sum_A \alpha_A = 1 - \sum_{A \neq \Theta} m_b(A) = m_b(\Theta)$.

2) Distance Between $BetP$ and π in the Quaternary Case: To get a hint of the relationship between orthogonal projection and pignistic function in the general case, let us compare their expressions in the simplest case in which they are distinct: a frame $\Theta = \{x, y, z, w\}$ of size 4. Their analytic expressions for the generic element $x \in \Theta$ are

$$\begin{aligned} BetP[b](x) &= m_b(x) + \frac{1}{2}\left(m_b(\{x, y\}) + m_b(\{x, z\}) \right. \\ &\qquad \left. + \; m_b(\{x, w\})\right) \\ &\quad + \frac{1}{3}\left(m_b(\{x, y, z\}) + m_b(\{x, y, w\}) \right. \\ &\qquad \left. + \; m_b(\{x, z, w\})\right) \\ &\quad + \frac{1}{4} m_b(\Theta) \\ \pi[b](x) &= m_b(x) + \frac{1}{2}\left(m_b(\{x, y\}) + m_b(\{x, z\}) \right. \\ &\qquad \left. + \; m_b(\{x, w\})\right) \\ &\quad + \frac{5}{16}\left(m_b(\{x, y, z\}) + m_b(\{x, y, w\}) \right. \\ &\qquad \left. + \; m_b(\{x, z, w\})\right) \\ &\quad + \frac{1}{16} m_b(\{y, z, w\}) + \frac{1}{4} m_b(\Theta). \end{aligned} \tag{34}$$

They are very similar to each other. Basically, the difference is that $\pi[b]$ also counts the masses of focal elements in $\{x\}^c$ (with a small contribution), while $BetP[b]$ by definition does not.

After computing their difference

$$\begin{aligned} BetP[b](x) - \pi[b](x) = \frac{1}{48}\,&[m_b(\{x, y, z\}) + m_b(\{x, y, w\}) \\ &+ m_b(\{x, z, w\}) - 3m_b(\{y, z, w\})] \end{aligned}$$

we can study their L_2 distance as b varies. After introducing the notation

$$\begin{aligned} y_1 &\doteq m_b(\{x, y, z\}) \quad & y_2 &\doteq m_b(\{x, y, w\}) \\ y_3 &\doteq m_b(\{x, z, w\}) \quad & y_4 &\doteq m_b(\{y, z, w\}) \end{aligned}$$

we can maximize (minimize) the norm

$$\begin{aligned} \|BetP[b] - \pi[b]\|^2 &\doteq \sum_x |BetP[b](x) - \pi[b](x)|^2 \\ &= (y_1 + y_2 + y_3 - 3y_4)^2 \\ &\quad + (y_1 + y_2 + y_4 - 3y_3)^2 \\ &\quad + (y_1 + y_3 + y_4 - 3y_2)^2 \\ &\quad + (y_2 + y_3 + y_4 - 3y_1)^2 \end{aligned}$$

by imposing $(\partial/\partial y_i)\|BetP[b](\mathbf{y}) - \pi[b](\mathbf{y})\|^2 = 0$ subject to $y_1 + y_2 + y_3 + y_4 = 1$. The unique solution turns out to be

$$\mathbf{y} = [1/4, 1/4, 1/4, 1/4]'$$

which corresponds to [after replacing this solution into (34)] $BetP[b] = \pi[b] = \overline{\mathcal{P}}$, where $\overline{\mathcal{P}} = [1/4, 1/4, 1/4, 1/4]'$ is the uniform probability on Θ. In other words, the distance between pignistic function and orthogonal projection is minimal (zero) when all size 3 subsets have the same mass.

It is then natural to suppose that their difference must be maximal when all the mass is concentrated on a single size-3 event. This is in fact correct: $\|BetP[b] - \pi[b]\|^2$ is maximal and equal to $1^2 + 1^2 + 1^2 + (-3)^2 = 12$ when $y_i = 1,\ y_j = 0\ \forall j \neq i$, i.e., the mass of one among $\{x, y, z\}, \{x, y, w\}, \{x, z, w\}$, $\{y, z, w\}$ is one.

VIII. Brief Discussion

The intuition for both the novel probabilistic approximations of a b.f. we introduced in this paper is provided by the analysis of the interplay between belief and probability spaces in the context of the geometric approach to the theory of evidence. Both intersection probability and orthogonal projection are related to the notion of orthogonality: the orthogonality of the dual line and that of $\pi[b] - b$ with respect to $\mathcal{P}$. Nevertheless, they possess different interpretations in terms of mass assignment, and relate in significant but distinct ways with the pignistic transformation.

An interesting parallel between $p[b]$ and $\pi[b]$ comes from their geometric description as points of a segment. Theorem 7 and (24)

$$\pi[b] = k_O[b]O[b] + \overline{\mathcal{P}}(1 - k_O[b])$$
$$p[b] = k_{\tilde{b}}\tilde{b} + (1 - k_{\tilde{b}})R[b]$$

state that they can both be written as convex combinations that depend on some flag probabilities associated with them, namely orthogonality and non-Bayesianity flag, respectively

$$\pi[b] \leftrightarrow O[b]$$
$$p[b] \leftrightarrow R[b].$$

It is then worth to study the condition under which $p[b]$ and orthogonal projection $\pi[b]$ are the same probability.

A trivial consequence of Theorem 4 is that when b is 2-additive, $\pi[b] = p[b] = \varsigma[b]$. This though gives us just "pointwise" information on the relationship between intersection probability and orthogonal projection. It would definitively be worth conducting a study of the distance between all Bayesian approximations of b.f.s, $BetP$, π, p, $\tilde{pl}_b$, $\tilde{b}$ as b varies in $\mathcal{B}$, in order to understand how they depend on the b.p.a. of b. We started doing this for the pair $BetP[b]$, $\pi[b]$ in the case of quaternary frames (Section VII-D2), getting some interesting results. We reserve to explore this direction thoroughly in the near future.

A. U.B.F.s

We also wish to add a remark on the validity of the results presented in this paper. They have been in fact obtained for "classical" b.f.s for which the mass assigned to the empty set is $0 : b(\emptyset) = m_b(\emptyset) = 0$. However, it makes sense in certain situations to work with u.b.f.s [38], i.e., b.f.s admitting nonzero support $m_b(\emptyset) \neq 0$ for the empty set [39]. $m_b(\emptyset)$ is an indicator of the amount of conflict in the evidence carried by a b.f. b but can also be interpreted as the possibility that the existing frame of discernment does not exhaust all the possible outcomes of the problem. U.B.F.s are naturally associated with vectors with $N = 2^{|\Theta|}$ coordinates. A new set of basis u.b.f. can then be defined

$$\{b_A \in \mathbb{R}^N, \emptyset \subseteq A \subseteq \Theta\}$$

this time including a vector $b_\emptyset \doteq [1\ 0\ \cdots 0]'$. Note also that in this case $b_\Theta = [0\ \cdots 0\ 1]'$.

It is natural to wonder whether the above discussion, and in particular definition and properties of $p[b]$ and $\pi[b]$, retains its validity. Let us consider again the binary case. We now have to use four coordinates associated with all events in Θ: $\emptyset$, $\{x\}$, $\{y\}$, and Θ. Remember that in the case of u.b.f.

$$b(A) = \sum_{\emptyset \subsetneq B \subseteq A} m_b(B), \qquad A \neq \emptyset$$

i.e., the contribution of the empty set is not considered when computing the belief value of an event $A \neq \emptyset$.[6] The corresponding basis belief and pl.f.s are then

$$\begin{aligned}
b_\emptyset &= [1, 0, 0, 0]' & pl_\emptyset &= [0, 0, 0, 0]' \\
b_x &= [0, 1, 0, 1]' & pl_x &= [0, 1, 0, 1]' = b_x \\
b_y &= [0, 0, 1, 1]' & pl_y &= [0, 0, 1, 1]' = b_y \\
b_\Theta &= [0, 0, 0, 1]' & pl_\Theta &= [0, 1, 1, 1]'.
\end{aligned}$$

A striking difference with the "classical" case is that $b(\Theta) = 1 - m_b(\emptyset) = pl_b(\Theta)$, which implies that both belief and plausibility spaces are *not* in general subsets of the section $v_\Theta = 1$ of $\mathbb{R}^N$. In other words, u.b.f. and u.pl.f. are not n.s.f.s (Section III-C).

More precisely, b, pl_b are n.s.f. iff $b(\emptyset) \neq 0$. As a consequence, *the line* $a(b, pl_b)$ *is not guaranteed to intersect the affine space* $\mathcal{P}'$ *of the Bayesian* n.s.f.

Consider for instance the line connecting $b_\emptyset$ and $pl_\emptyset$ in the binary case

$$\alpha\, b_\emptyset + (1 - \alpha)\, pl_\emptyset = \alpha\, [1, 0, 0, 0]', \qquad \alpha \in \mathbb{R}.$$

As $\mathcal{P}' = \{[a, b, (1 - b), -a]', a, b \in \mathbb{R}\}$, there clearly is no value $\alpha \in \mathbb{R}$ s.t. $\alpha \cdot [1, 0, 0, 0]' \in \mathcal{P}'$.

Simple calculations show that in fact $a(b, pl_b) \cap \mathcal{P}' \neq \emptyset$ iff $b(\emptyset) = 0$ (i.e., b is "classical") or (trivially) $b \in \mathcal{P}$. This is true in the general case.

Proposition 2: $p[b]$ and $\beta[b]$ are well defined for classical b.f.s only.

It is interesting to note that however the orthogonality results of Section V-A *are still valid* since Lemma 1 does not involve the empty set, while the proof of Theorem 2 is valid for the components $A = \emptyset, \Theta$ too (as $b_y - b_x(A) = 0$ for $A = \emptyset, \Theta$).

Proposition 3: $a(b, pl_b)$ is orthogonal to $\mathcal{P}$ for each u.b.f. b, although $\varsigma[b] = a(b, pl_b) \cap \mathcal{P}' \neq \emptyset$ iff b is a b.f.

Analogously, the orthogonality condition (27) does not concern the mass of the empty set. The orthogonal projection $\pi[b]$ of a u.b.f. b is then well defined (check Theorem 6's proof), and

[6] In the unnormalized case, the notation b is usually reserved for *implicability* functions, while belief functions are denoted by Bel [12].

it is still given by (28) and (29), where this time the summations on the right-hand side include the empty set too

$$\pi[b](x) = \sum_{A \supset x} m_b(A) 2^{1-|A|} + \sum_{\emptyset \subseteq A \subset \Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right)$$

$$\pi[b](x) = \sum_{A \supset x} m_b(A) \left(\frac{1 + |A^c| 2^{1-|A|}}{n} \right) + \sum_{\emptyset \subseteq A \not\supset x} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right).$$

IX. Conclusion

In this paper, we introduced two new probabilistic approximations of b.f.s, which are both derived from purely geometric considerations. They are indeed associated with two different geometric loci: the dual line passing through b and pl_b in the belief space; and the orthogonal complement of the probability subspace.

After proving that the line $a(b, pl_b)$ is always orthogonal to $\mathcal{P}$ and intersects the region of the Bayesian n.s.f. $\mathcal{P}'$, we introduced the probability $p[b]$ associated with this intersection and discussed two interpretations of $p[b]$ in terms of non-Bayesian contributions of singletons.

On the other side, after precising the condition under which a b.f. b is orthogonal to $\mathcal{P}$, we gave two equivalent expressions of the orthogonal projection of b onto $\mathcal{P}$. We saw that $\pi[b]$ can be reduced to another probability signaling the distance of b from orthogonality, and that this "orthogonality flag" can in turn be interpreted as the result of a mass redistribution process analogous to that associated with the pignistic transformation. We proved that $\pi[b]$ commutes with the affine combination operator and can therefore be expressed as a convex combination of basis pignistic functions, which confirms the strict relation between $\pi[b]$ and $BetP[b]$.

We finally studied the difference between intersection probability and orthogonal projection, and discussed which results retain their validity in the case of u.b.f.s.

We have seen when discussing the binary case that, while $BetP[b]$, $p[b]$, and $\pi[b]$ belong to the same "family" of Bayesian approximations of b (as they coincide under 2-additivity), the relative plausibility $\tilde{p}[b]$ and belief $\tilde{b}$ of singletons [13] do not fit in the same scheme. In the near future, we will show that $\tilde{p}[b]$ turns out to be the best Bayesian approximation of a b.f. in the framework of Dempster's combination rule, and investigate the dual geometry of relative plausibility and belief of singletons [36]. Naturally enough, the geometric approach can also be exploited to study the problem of approximating a b.f. with a possibility measure or "consistent" b.f. [2]. Last but not least, it will be definitively worth to seek for a complete picture of the conditions under which all different Bayesian approximations of b coincide as a crucial contribution to a full understanding their semantics.

Appendix
Proofs

Proof of Theorem 4

By definition (17), $\varsigma[b]$ can be written in terms of the reference frame $\{b_A, A \subset \Theta\}$ as

$$\sum_{A \subset \Theta} m_b(A) b_A + \beta[b] \left(\sum_{A \subset \Theta} \mu_b(A) b_A - \sum_{A \subset \Theta} m_b(A) b_A \right) = \sum_{A \subset \Theta} b_A \left[m_b(A) + \beta[b] \left(\mu_b(A) - m_b(A) \right) \right]$$

since $\mu_b(.)$ is the Moebius inverse of $pl_b(.)$. For $\varsigma[b]$ to be a Bayesian b.f., accordingly, all the components related to nonsingleton subsets need to be zero

$$m_b(A) + \beta[b] \left(\mu_b(A) - m_b(A) \right) = 0, \qquad \forall A : |A| > 1.$$

This condition in turn reduces to (recalling expression (20) of $\beta[b]$)

$$\mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \left[\sum_{|B|>1} m_b(B)|B| - \sum_{|B|>1} m_b(B) \right] = 0$$
$$\equiv \mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>1} m_b(B)(|B| - 1) = 0 \qquad (35)$$

$\forall A : |A| > 1$. But now, $\sum_{|B|>1} m_b(B)(|B|-1) = \sum_{|B|>1} m_b(B) + \sum_{|B|>2} m_b(B)(|B| - 2)$, so that (35) reads as

$$[\mu_b(A) + m_b(A)] \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>2} m_b(B)(|B| - 2) = 0$$
$$\equiv [m_b(A) + \mu_b(A)]\, M_1[b] + m_b(A) M_2[b] = 0 \qquad (36)$$

$\forall A : |A| > 1$, after defining $M_1[b] \doteq \sum_{|B|>1} m_b(B)$, and $M_2[b] \doteq \sum_{|B|>2} m_b(B)(|B| - 2)$, respectively.

Now, it is easy to note that

$$M_1[b] = 0 \Leftrightarrow m_b(B) = 0 \qquad \forall B : |B| > 1 \Leftrightarrow b \in \mathcal{P}$$
$$M_2[b] = 0 \Leftrightarrow m_b(B) = 0 \qquad \forall B : |B| > 2$$

as all the terms inside the summations are nonnegative by definition of b.p.a.. We can distinguish three cases: 1) $M_1 = 0 = M_2$ ($b \in \mathcal{P}$); 2) $M_1 \neq 0$ but $M_2 = 0$, and finally 3) $M_1 \neq 0 \neq M_2$. If $M_1 = M_2 = 0$, then b is a probability (trivially), while if $M_1 \neq 0 \neq M_2$, then (36) implies $m_b(A) = \mu_b(A) = 0$, $|A| > 1$ i.e., $b \in \mathcal{P}$, which is a contradiction.

The only nontrivial case is then $M_2 = 0$, where condition (36) becomes

$$M_1[b]\, [m_b(A) + \mu_b(A)] = 0, \qquad \forall A : |A| > 1.$$

For all $|A| > 2$, we have that $m_b(A) = \mu_b(A) = 0$ (since $M_2 = 0$), and the constraint is met. If $|A| = 2$, instead $\mu_b(A) = (-1)^{|A|+1} \sum_{B \supset A} m_b(B) = (-1)^{2+1} m_b(A) = -m_b(A)$ (since $m_b(B) = 0 \, \forall B \supset A, |B| > 2$) so that $\mu_b(A) + m_b(A) = 0$, and the constraint is again met. Finally, as the coordinate $\beta[b]$ of $\varsigma[b]$ on the line $a(b, pl_b)$ can then be re-written as

$$\beta[b] = \frac{M_1[b]}{M_2[b] + 2M_1[b]} \tag{37}$$

if $M_2 = 0$, then $\beta[b] = 1/2$, and $\varsigma[b] = (b + pl_b)/2$.

Proof of Theorem 6

Finding the orthogonal projection $\pi[b]$ of b onto $a(\mathcal{P})$ is equivalent to imposing the condition $\langle \pi[b] - b, b_y - b_x \rangle = 0 \; \forall \, y \neq x$. Replacing the masses of $\pi - b$

$$\begin{cases} \pi(x) - m_b(x), & x \in \Theta \\ -m_b(A), & |A| > 1 \end{cases}$$

into (27) yields, after extracting the singletons x from the summation, the system

$$\begin{cases} \pi(y) = \pi(x) + \sum\limits_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} + m_b(y) \\ \qquad - m_b(x) - \sum\limits_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \qquad \forall y \neq x \\ \sum\limits_{y \in \Theta} \pi(y) = 1. \end{cases} \tag{38}$$

After replacing the first $n-1$ equations of (38) into the normalization constraint, we get

$$\pi(x) + \sum_{y \neq x} \Bigg[\pi(x) + m_b(y) - m_b(x) + \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} - \sum_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \Bigg] = 1$$

which is equivalent to

$$\begin{aligned} n\pi(x) = & \, 1 + (n-1) m_b(x) - \sum_{y \neq x} m_b(y) \\ & + \sum_{y \neq x} \sum_{A \supset x, A \not\supset y, |A|>1} m_b(A) 2^{1-|A|} \\ & - \sum_{y \neq x} \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|}. \end{aligned}$$

But now

$$\sum_{y \neq x} \sum_{A \supset y, A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} = \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A|$$

as all events A not containing x do contain some $y \neq x$, and they are counted $|A|$ times (i.e., once for each element they contain). Instead

$$\begin{aligned} & \sum_{y \neq x} \sum_{A \supsetneq x, A \not\supset y} m_b(A) 2^{1-|A|} \\ & \quad = \sum_{A \supset x, 1 < |A| < n} m_b(A) 2^{1-|A|} (n - |A|) \\ & \quad = \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} (n - |A|) \end{aligned}$$

for $n - |A| = 0$ when $A = \Theta$. Hence, $\pi(x)$ is equal to

$$\begin{aligned} & \frac{1}{n} \Bigg[1 + (n-1) m_b(x) - \sum_{y \neq x} m_b(y) - \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A| \\ & \qquad + \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} (n - |A|) \Bigg] \\ & = \frac{1}{n} \Bigg[n \, m_b(x) + 1 - \sum_{y \in \Theta} m_b(y) + n \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad - \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} |A| - \sum_{A \not\supset x, |A|>1} m_b(A) 2^{1-|A|} |A| \Bigg]. \end{aligned}$$

We then just need to note that $-\sum_{y \in \Theta} m_b(y) = -\sum_{|A|=1} m_b(A) |A| 2^{1-|A|}$, so that the orthogonal projection can be finally expressed as

$$\begin{aligned} \pi(x) & = \frac{1}{n} \Bigg[n \, m_b(x) + n \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad + 1 - \sum_{A \subset \Theta} m_b(A) |A| 2^{1-|A|} \Bigg] \\ & = m_b(x) + \sum_{A \supsetneq x} m_b(A) 2^{1-|A|} \\ & \qquad + \sum_{A \subset \Theta} m_b(A) \left(\frac{1 - |A| 2^{1-|A|}}{n} \right) \end{aligned}$$

i.e., (28), and since

$$\begin{aligned} 2^{1-|A|} + \frac{1}{n} - \frac{|A|}{n} 2^{1-|A|} & = \frac{1 + 2^{1-|A|} (n - |A|)}{n} \\ & = \frac{1 + 2^{1-|A|} |A^c|}{n} \end{aligned}$$

we get the second form (29).

Proof of Theorem 7

By (28), we can write

$$\begin{aligned} \pi[b](x) & = \bar{O}[b](x) + \frac{1}{n} \Bigg(\sum_{A \subset \Theta} m_b(A) \\ & \qquad - \sum_{A \subset \Theta} m_b(A) |A| 2^{1-|A|} \Bigg) \\ & = \bar{O}[b](x) + \frac{1}{n} \left(1 - k_O[b] \right). \end{aligned}$$

But since

$$\begin{aligned}\sum_{x\in\Theta}\bar{O}[b](x) &= \sum_{x\in\Theta}\sum_{A\supset x} m_b(A)2^{1-|A|}\\ &= \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\\ &= k_O[b]\end{aligned} \tag{39}$$

i.e., $k_O[b]$ is the normalization factor for $\bar{O}[b]$, the function (30) is a Bayesian b.f., and we can write (as $\overline{\mathcal{P}}(x) = (1/n)$) $\pi[b] = (1 - k_O[b])\overline{\mathcal{P}} + k_O[b]O[b]$.

References

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

[2] D. Dubois and H. Prade, "Consonant approximations of belief functions," *Int. J. Approx. Reason.*, vol. 4, no. 5/6, pp. 419–449, Sep./Nov. 1990.

[3] A. B. Yaghlane, T. Denoeux, and K. Mellouli, "Coarsening approximations of belief functions," in *Proc. ECSQARU*, S. Benferhat and P. Besnard, Eds., 2001, pp. 362–373.

[4] T. Denoeux, "Inner and outer approximation of belief structures using a hierarchical clustering approach," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 9, no. 4, pp. 437–460, Aug. 2001.

[5] T. Denoeux and A. B. Yaghlane, "Approximating the combination of belief functions using the fast Moebius transform in a coarsened frame," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 77–101, Oct. 2002.

[6] R. Haenni and N. Lehmann, "Resource bounded and anytime approximation of belief function computations," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 103–154, Oct. 2002.

[7] M. Bauer, "Approximation algorithms and decision making in the Dempster–Shafer theory of evidence—An empirical study," *Int. J. Approx. Reason.*, vol. 17, no. 2/3, pp. 217–237, Aug.–Oct. 1997.

[8] M. Bauer, "Approximations for decision making in the Dempster–Shafer theory of evidence," in *Proc. 12th Conf. Uncertainty Artif. Intell.*, F. Horvitz and E. Jensen, Eds., Portland, OR, Aug. 1–4, 1996, pp. 73–80.

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

[10] J. D. Lowrance, T. D. Garvey, and T. M. Strat, "A framework for evidential-reasoning systems," in *Proc. Nat. Conf. Artif. Intell.*, 1986, pp. 896–903.

[11] P. Smets, "Belief functions versus probability functions," in *Uncertainty and Intelligent Systems*, L. Saitta, B. Bouchon, and R. Yager, Eds. Berlin, Germany: Springer-Verlag, 1988, pp. 17–24.

[12] P. Smets, "Decision making in the TBM: The necessity of the pignistic transformation," *Int. J. Approx. Reason.*, vol. 38, no. 2, pp. 133–147, Feb. 2005.

[13] F. Voorbraak, "A computationally efficient approximation of Dempster–Shafer theory," *Int. J. Man-Mach. Stud.*, vol. 30, no. 5, pp. 525–536, May 1989.

[14] B. Cobb and P. Shenoy, "On transforming belief function models to probability models," Univ. Kansas, Sch. Bus., Lawrence, KS, Feb. 2003. Working Paper 293, Tech. Rep.

[15] B. R. Cobb and P. P. Shenoy, "A comparison of Bayesian and belief function reasoning," *Inf. Syst. Frontiers*, vol. 5, no. 4, pp. 345–358, Dec. 2003.

[16] B. R. Cobb and P. P. Shenoy, "A comparison of methods for transforming belief function models to probability models," in *Proc. ECSQARU*, Aalborg, Denmark, Jul. 2003, pp. 255–266.

[17] B. Cobb and P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models," *Int. J. Approx. Reason.*, vol. 41, no. 3, pp. 314–330, Apr. 2006.

[18] V. Ha and P. Haddawy, "Theoretical foundations for abstraction-based probabilistic planning," in *Proc. 12th Conf. Uncertainty Artif. Intell.*, Aug. 1996, pp. 291–298.

[19] P. Black, "Geometric structure of lower probabilities," in *Random Sets: Theory and Applications*, J. Goutsias, R. P. S. Mahler, and H. T. Nguyen, Eds. New York: Springer-Verlag, 1997, pp. 361–383.

[20] P. Black, "An examination of belief functions and other monotone capacities," Ph.D. dissertation, Dept. Statist., Carnegie Mellon Univ., Pittsburg, PA, 1996.

[21] V. Ha and P. Haddawy, "Geometric foundations for interval-based probabilities," in *KR'98: Principles of Knowledge Representation and Reasoning*, A. G. Cohn, L. Schubert, and S. C. Shapiro, Eds. San Francisco, CA: Morgan Kaufmann, 1998, pp. 582–593. [Online]. Available: citeseer.ist.psu.edu/ha98geometric.html

[22] F. Cuzzolin and R. Frezza, "Geometric analysis of belief space and conditional subspaces," in *Proc. 2nd ISIPTA*, Ithaca, NY, Jun. 26–29, 2001, pp. 122–132.

[23] F. Cuzzolin, "Geometry of upper probabilities," in *Proc. 3rd ISIPTA*, Jul. 2003, pp. 188–203.

[24] F. Cuzzolin, "A geometric approach to the theory of evidence," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, 2007. to be published.

[25] A. P. Dempster, "Upper and lower probability inferences based on a sample from a finite univariate population," *Biometrika*, vol. 54, no. 3/4, pp. 515–528, Dec. 1967.

[26] A. Dempster, "Upper and lower probabilities generated by a random closed interval," *Ann. Math. Stat.*, vol. 39, no. 3, pp. 957–966, Jun. 1968.

[27] A. Dempster, "Upper and lower probabilities inferences for families of hypothesis with monotone density ratios," *Ann. Math. Stat.*, vol. 40, no. 3, pp. 953–969, Jun. 1969.

[28] F. Cuzzolin, "Visions of a generalized probability theory," Ph.D. dissertation, Università di Padova, Dipartimento di Elettronica e Informatica, Padova, Italy, Feb. 2001.

[29] F. Cuzzolin, "Geometry of Dempster's rule of combination," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 2, pp. 961–977, Apr. 2004.

[30] B. Dubrovin, S. Novikov, and A. Fomenko, *Sovremennaja Geometrija. Metody i Prilozenija*. Moscow, Russia: Nauka, 1986.

[31] M. Aigner, *Combinatorial Theory*. New York: Springer-Verlag, 1979.

[32] A. Chateauneuf and J. Y. Jaffray, "Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion," *Math. Soc. Sci.*, vol. 17, no. 3, pp. 263–283, Jun. 1989.

[33] D. Dubois, H. Prade, and P. Smets, "New semantics for quantitative possibility theory," in *Proc. ISIPTA*, 2001, pp. 152–161.

[34] P. Smets, "Constructing the pignistic probability function in a context of uncertainty," in *Proc. Uncertainty Artif. Intell., 5*, M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, Eds., 1990, pp. 29–39.

[35] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, no. 2, pp. 191–234, 1994.

[36] F. Cuzzolin, "The geometry of relative plausibility and belief of singletons," *Ann. Math. Artif. Intell.*, May 2007. submitted for publication.

[37] F. Cuzzolin, "The geometry of relative plausibilities," in *11th Int. Conf. IPMU, Special Session Fuzzy Measures and Integrals, Capacities and Games*, 2006.

[38] P. Smets, "The nature of the unnormalized beliefs encountered in the transferable belief model," in *Proc. 8th Annu. Conf. UAI*, 1992, pp. 292–297.

[39] P. Smets, "The application of the matrix calculus to belief functions," *Int. J. Approx. Reason.*, vol. 31, no. 1/2, pp. 1–30, Oct. 2002.

Fabio Cuzzolin was born in Jesolo, Italy, in 1971. He received the Laurea (*magna cum laude*) and Ph.D. degrees from the University of Padova, Padova, Italy, in 1997 and 2001, respectively. His thesis was entitled "Visions of a generalized probability theory."

During his doctoral term, he was with the Autonomous Navigation and Computer Vision Laboratory (NAVLAB), University of Padova. He was also a Postodoctoral Researcher at Politecnico di Milano, Milan, Italy, and was with the UCLA Vision Laboratory, Los Angeles, CA. He is currently with the Perception Project, INRIA Rhône-Alpes, Saint Ismier Cedex, France. His research includes computer vision applications like gesture and action recognition, object pose estimation, and identity recognition from gait. His main field of investigation remains however that of generalized and imprecise probabilities. In particular, he has formulated a geometric approach to the theory of belief functions, focusing mainly on the probabilistic approximation problem, and studied the notion of independence of sources from an algebraic point of view. He collaborates with several journals, among which the *International Journal of Approximate Reasoning*, *Information Fusion*, and the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B.

Dr. Cuzzolin is a member of the Society for Imprecise Probabilities and Their Applications.