# Two new Bayesian approximations of belief functions based on convex geometry

Fabio Cuzzolin

# Two new Bayesian approximations of belief functions based on convex geometry

Fabio Cuzzolin

**Abstract**

In this paper we analyze from a geometric point of view the meaningful relations which take place between a belief function and the set of 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 the three major geometric entities that relate a generic belief function to the set of probabilities $\mathcal{P}$: the dual line connecting belief and plausibility functions, the orthogonal complement of $\mathcal{P}$, and the simplex of consistent probabilities. These are in turn associated with different probability measures which depend on the original belief function. We describe in particular geometry and properties of the orthogonal projection of a belief function onto $\mathcal{P}$ and the intersection probability, provide their interpretations in terms of degrees of belief, and discuss their behavior with respect to convex closure.

**Index Terms**

Theory of evidence, geometric approach, Bayesian belief functions, intersection probability, orthogonal projection, commutativity.

## I. INTRODUCTION

Uncertainty measures have a major role in fields like artificial intelligence, where problems requiring formalized reasoning are common. The theory of evidence is one of the most popular among those formalisms, thanks perhaps to its nature of quite 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 belief functions*.

F. Cuzzolin is with Perception team, INRIA Rhône-Alpes, Montbonnot, France. Email: Fabio.Cuzzolin@inrialpes.fr

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 belief function. A number of papers [3], [4], [5], [6] have been published on this issue (see [7], [8], [9] for a review), mainly in order to find efficient implementations of the rule of combination aiming to reduce the number of focal elements. Tessem [10], 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.* [11]. The connection between belief functions and probabilities is as well the basement of a popular approach to the theory of evidence, Smets' *pignistic* model [12], in which beliefs are represented at credal level, while decisions are made by resorting to a Bayesian belief function called *pignistic transformation* [13]. On his side, in his 1989 paper [14] F. Voorbraak proposed to adopt the so-called *relative plausibility* function $\tilde{pl}_b$, the unique probability that, given a belief function $b$ with plausibility $pl_b$, assigns to each singleton its normalized plausibility. 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 [15], [16], [17] also analyzed the properties of the relative plausibility of singletons [18] and discussed its nature of probability function that is equivalent to the original belief function.

The study of the interplay between belief functions and probabilities has also been posed in a geometric setup [19], [20], [21]. P. Black, in particular, dedicated his doctoral thesis to the study of the geometry of belief functions and other monotone capacities [21]. An abstract of his results can be found in [20], 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 a work [19] of Ha and Haddawy where they propose an "affine operator" which can be considered a generalization of both belief functions 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 [22] 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*, in which branches are annotated with intervals, and nodes with convex sets of probabilities.

On our side, in a series of recent works [23], [24], [25] we proposed a geometric interpretation of the theory of evidence in which belief functions are represented as points of a simplex called *belief space* [23]. As a matter of fact, as a belief function $b : 2^\Theta \to [0, 1]$ is completely specified by its $N \doteq 2^{|\Theta|} - 1$ belief values $\{b(A), A \subset \Theta, A \neq \emptyset\}$, it can be represented as a point of the Cartesian space $\mathbb{R}^{N-1}$.

In this framework many different uncertainty descriptions like upper and lower probabilities, belief functions, possibility measures can be studied in an 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 belief function, both of them derived from purely geometric considerations. We thoroughly discuss their interpretation and properties, and their relations with the other known Bayesian approximations of belief functions, 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 belief function $b$ is associated with three different geometric entities, namely the line $(b, pl_b)$ joining $b$ with the related plausibility function $pl_b$, the orthogonal complement $\mathcal{P}^\perp$ of the probabilistic subspace $\mathcal{P}$, and the simplex of consistent probabilities $\mathcal{P}[b] = \{p \in \mathcal{P} : p(A) \geq b(A) \ \forall A \subset \Theta\}$. These in turn determine three different probabilities associated with $b$, i.e. the *orthogonal projection* $\pi[b]$ of $b$ onto $\mathcal{P}$, the barycenter of $\mathcal{P}[b]$ or *pignistic function* $BetP[b]$, and the *intersection probability* $p[b]$. In the binary case all those Bayesian functions coincide.

In Section V we prove that, even though 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.), i.e. the natural generalizations of belief functions obtained by relaxing the positivity constraint for b.p.a.s: This intersection yields a Bayesian n.s.f. $\varsigma[b]$.

In Section VI we will see that $\varsigma[b]$ is in turn naturally associated with a Bayesian belief function $p[b]$, which we call intersection probability. We will give two different interpretations of the way in which this probability distributes the masses of the focal elements on $b$ to the elements of $\Theta$, 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 pignistic function and 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 be in turn 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 convex combination operator, and can therefore be expressed as a convex combination of basis pignistic functions, confirming the strict relation between $\pi[b]$ and $BetP[b]$.

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

To improve the readability of the paper the proofs of all major results have been moved to an appendix.

## II. THE THEORY OF EVIDENCE

The *theory of evidence* [1] was introduced in the late Seventies by Glenn Shafer as a way of representing epistemic knowledge, starting from a sequence of seminal works [26], [27], [28], of Arthur Dempster. In this formalism the best representation of chance is a *belief function* (b.f.) rather than a Bayesian mass distribution, assigning probability values to *sets* of possibilities rather than single events.

*Definition 1:* A *basic probability assignment* (b.p.a.) over a finite set (*frame of discernment*

[1]) $\Theta$ is a function $m : 2^\Theta \to [0, 1]$ on its power set $2^\Theta = \{A \subset \Theta\}$ such that

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

Subsets of $\Theta$ associated with non-zero values of $m$ are called *focal elements*.

*Definition 2:* The *belief function* $b : 2^\Theta \to [0, 1]$ associated with a basic probability assignment $m$ on $\Theta$ is defined as:

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

Conversely, the unique basic probability assignment $m_b$ associated with a given belief function $b$ 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 is simply a peculiar belief function assigning non-zero masses to singletons only (*Bayesian* b.f.): $m_b(A) = 0, \ |A| > 1$.

A dual mathematical representation of the evidence encoded by a belief function $b$ is the *plausibility function* (pl.f.)

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

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

$$pl_b(A) \doteq 1 - b(A^c) = 1 - \sum_{B \subset A^c} m_b(B) \tag{2}$$

where $A^c$ denotes the complement of $A$ in $\Theta$.

For each event $A$ $pl_b(A)$ expresses the amount of evidence *not against* $A$. $pl_b$ conveys as much information as $b$, and can be computed from the b.p.a. as $pl_b(A) = \sum_{B \cap A \neq \emptyset} m_b(B) \geq b(A)$.

## III. GEOMETRY OF BELIEF AND PLAUSIBILITY FUNCTIONS

### A. *Belief space*

Motivated by the search for a meaningful probabilistic approximation of belief functions we introduced the notion of *belief space* ([23], [29], [25]), as the space of all the belief functions we can define on a given domain[1]. Consider a frame of discernment $\Theta$ and introduce in the

---

[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.

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 belief function, 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 $\Theta$ is the set of points $\mathcal{B}_\Theta$ of $\mathbb{R}^{N-1}$ corresponding to a belief function.

We will assume the domain $\Theta$ fixed, and denote the belief space with $\mathcal{B}$. To determine which points of "are" belief functions 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 [30] for the details) that $\mathcal{B}$ is convex. Let us call

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

the unique belief function assigning all the mass to a single subset $A$ of $\Theta$ ($A$-th *basis belief function*). It can be proved that [30], denoting with $\mathcal{E}_b$ the list of focal elements of $b$,

*Theorem 1:* The set of all the belief functions with focal elements in a given collection $L$ is closed and convex in $\mathcal{B}$:

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

where $Cl$ denotes the convex closure operator:

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

The following is then just a consequence of Theorem 1.

*Corollary 1:* The belief space $\mathcal{B}$ coincides with the convex closure of all the basis belief functions $b_A$,

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

The convex space delimited by a collection of points is called *simplex*: Figure 1 illustrates the simplicial form of $\mathcal{B}$. Moreover, each belief function $b \in \mathcal{B}$ can be written as a convex sum as

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

Geometrically, the 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 belief function assigning non zero masses to singletons only, Theorem 1 implies that
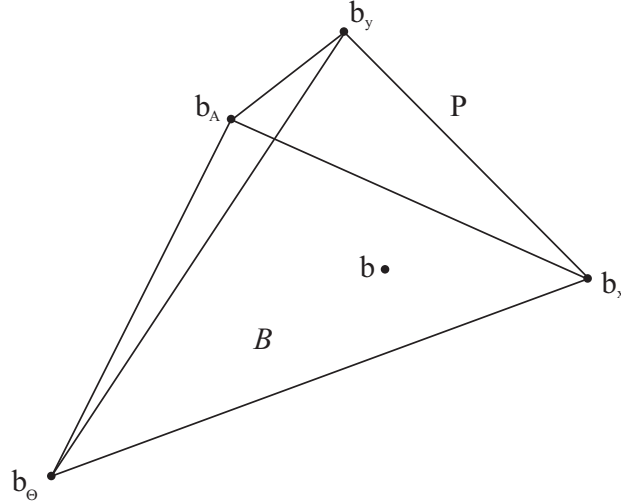
Fig. 1.    Simplicial structure of the belief space $\mathcal{B}$: its vertices are all the basis belief functions $b_A$ represented as vectors of $\mathbb{R}^{N-1}$. The probabilistic subspace is just a subset $Cl(b_x, x \in \Theta)$ of its border.

*Corollary 2:* The set $\mathcal{P}$ of all the Bayesian belief functions on $\Theta$ is a subset of the border of $\mathcal{B}$, precisely the simplex determined by all the basis functions associated with singletons [2]:

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

The interpretation of belief functions as convex sets of probabilities also fits in this framework, as the credal sets [31] $\{p \in \mathcal{P} : p(A) \geq b(A) \ \forall A \subset \Theta\}$ are nothing but simplices in $\mathcal{P}$ [25].

### B. Plausibility space

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

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

In [24] we proved that

[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 $m_b(\{x\}), pl_b(\{x\})$.

*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 belief function $b_A$, $pl_A = pl_{b_A}$. Again, every plausibility vector $pl_b$ can be uniquely expressed as a combination of the basis belief functions $b_A$. We have that[3]

$$pl_b = \sum_{B \subset \Theta} pl_b(B) X_B = \sum_{B \subset \Theta} pl_b(B) \cdot \sum_{A \supset B} b_A (-1)^{|A \setminus B|} = \sum_{A \subset \Theta} b_A \Big( \sum_{B \subset A} (-1)^{|A \setminus B|} pl_b(B) \Big)$$

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

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

where (see [24])

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

($\mu_b(\emptyset) = 0$) is the Moebius inverse of the plausibility function, 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. Normalized sum functions

It may be confusing to think of belief and plausibility functions as points of the same Cartesian space. However, this is a simple consequence of the fact that both are defined on the same domain, the power set of $\Theta$. As $\Theta$ 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 [30]. In fact, following the same principle, each vector $v = [v_1, ..., v_A, ..., v_\Theta]' \in \mathbb{R}^{N-1}$ can be thought of as a function $\varsigma : 2^\Theta \setminus \emptyset \rightarrow \mathbb{R}$ s.t. $\varsigma(A) = v_A$. As the Moebius transformation is invertible, for each of these functions $\varsigma$ there always exists another function $m_\varsigma : 2^\Theta \setminus \emptyset \rightarrow \mathbb{R}$ such that

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

---

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

i.e. *each* vector $\varsigma$ of $\mathbb{R}^{N-1}$ can be thought of as a *sum function* (see [32] for a brief introduction). However, $m_\varsigma$ does not in general meet the positivity constraint $m_\varsigma(A) \geq 0 \; \forall \, A \subset \Theta$.

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 the points of this section are sum functions meeting the normalization axiom,

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

or *normalized sum functions* (n.s.f.). n.s.f are the natural extensions of belief functions in this geometric framework. Analogously to the case of belief functions, we can call *Bayesian normalized sum functions* the n.s.f. $\varsigma$ 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. Figure 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

$$b = [b(x) = m_b(x), b(y) = m_b(y)]'$$
$$pl_b = [pl_b(x) = 1 - m_b(y), pl_b(y) = 1 - m_b(x)]'$$

respectively. 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 the two simplices are symmetric with respect to the Bayesian region $\mathcal{P}$. Furthermore, each pair of functions $(b, pl_b)$ determines a line which is *orthogonal* to $\mathcal{P}$, where $b$ and $pl_b$ lie on symmetric positions on the two sides of the Bayesian region.

Let us denote with $a(v_1, .., v_k)$ the affine subspace of some Cartesian space $\mathbb{R}^m$ generated by the points $v_1, ..., v_k \in \mathbb{R}^m$, 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 the normalized sum functions on $\Theta_2$. The region $\mathcal{P}'$ of all the Bayesian n.s.f. is obviously (recalling Equation (9)) the line

$$\mathcal{P}' = \{\varsigma \in \mathbb{R}^2 : m_\varsigma(x) + m_\varsigma(y) = 1\} = a(\mathcal{P})$$
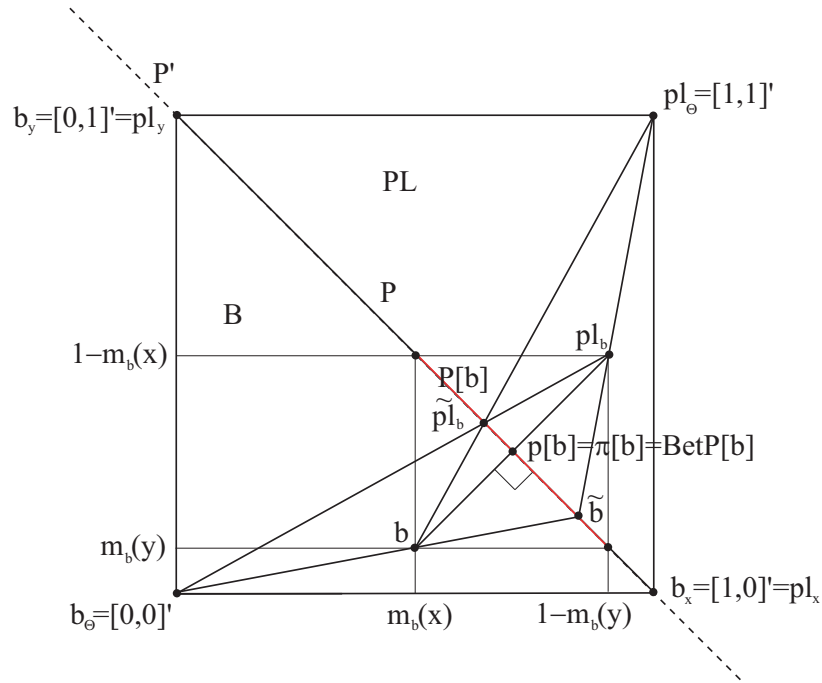
Fig. 2. In a binary frame $\Theta_2 = \{x, y\}$ both belief $\mathcal{B}$ and plausibility space $\mathcal{PL}$ are simplices with vertices $\{b_\Theta = [0, 0]', b_x = [1, 0]', b_y = [0, 1]'\}$ and $\{pl_\Theta = [1, 1]', pl_x = b_x, pl_y = b_y\}$ respectively. A belief function $b$ and the corresponding plausibility function $pl_b$ are always located in symmetric positions with respect to the set $\mathcal{P}$ of probabilities on $\Theta$. 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 $b_\Theta$ respectively. The other Bayesian functions related to $b$ all coincide with the center of the segment of consistent probabilities $\mathcal{P}[b]$.

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-dimensional simplex) in $\mathcal{P}$ (see Figure 2 again), whose center of mass is well known [33], [34], [24] to be Smets' *pignistic function* ([35], [36], [15])

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

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

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

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

even though it is 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

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

assigning to each singleton $x$ its normalized mass (see Figure 2 again). We will consider their behavior separately in the near future [37].

In the following we will study each of these geometric entities related to $b$, in particular the line $a(b, pl_b)$, and the orthogonal complement of $\mathcal{P}$. We will provide a unified picture of the geometric interplay of belief and probability by studying the properties of two Bayesian functions derived from geometrical considerations, the orthogonal projection $\pi[b]$ and the intersection probability $p[b]$. We will compare them with both the pignistic function and the 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 (in the case of the orthogonal projection) their behavior with respect to convex combination.

## V. Geometry of the dual line

Let us then first study 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 belief functions. 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 remembering that, by definition,

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

we can see that these vectors show a rather peculiar symmetry

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

that can be usefully exploited.

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

*Proof:* Following (12) we can appreciate that

$$[b_y - b_x](A) = 1 \Rightarrow A \supset \{y\}, A \not\supset \{x\} \Rightarrow A^c \supset \{x\}, A^c \not\supset \{y\} \Rightarrow [b_y - b_x](A^c) = -1$$

and vice-versa, while $[b_y - b_x](A) = 0 \Rightarrow A \supset \{y\}, A \supset \{x\}$ or $A \not\supset \{y\}, A \not\supset \{x\}$. In the first case $A^c \not\supset \{x\}, \{y\}$, in the second one $A^c \supset \{x\}, \{y\}$; therefore 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 $A$-th 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

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

where

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

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

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

---

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

which by Equation (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. ∎

## B. Intersection with the region of Bayesian normalized sum functions

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 $N - 2 = (2^n - 2)$-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 Figure 2 again). Formally, for a point on the line $a(b, pl_b)$ to be a probability we need to find a value of $\alpha$ 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,

$$b(x) + \alpha[pl_b(x) - b(x)] = b(x) + \alpha[1 - b(x^c) - b(x)]. \tag{15}$$

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

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

which yields a single candidate value $\beta[b]$ for the line coordinate of the intersection. Using the terminology of Section III-C, the candidate projection

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

(having called $\mathcal{P}'$ the set of all the Bayesian n.s.f. in $\mathbb{R}^{N-2}$) is *a Bayesian n.s.f.*, but is not guaranteed to be a Bayesian *belief* function. For normalized sum functions, 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

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

*Theorem 3:* The coordinates of $\varsigma[b]$ with respect to the basis Bayesian belief functions $\{b_x, x \in \Theta\}$ can be expressed in terms of the basic probability assignment $m_b$ of $b$ as follows:

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

where

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

*Proof:* The numerator of Equation (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

$$\sum_{y \in \Theta} [pl_b(y) - b(y)] = \sum_{y \in \Theta} (1 - b(y^c) - b(y)) = \sum_{y \in \Theta} \left[ \sum_{B \subset \Theta} m_b(B) - \sum_{B \subset y^c} m_b(B) - \sum_{B \subset y} m_b(B) \right] =$$

$$= \sum_{y \in \Theta} \sum_{B \supset y, B \neq y} m_b(B) = \sum_{|B|>1} m_b(B)|B|.$$

$\blacksquare$

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

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

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

*Theorem 4:* $\varsigma[b] \in \mathcal{P}$ 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 $\mathcal{P}$.

In other words, for 2-additive belief functions $\varsigma[b]$ is nothing but the *mean probability* function $\frac{b+pl_b}{2}$. In the general case, instead, the reflection of $b$ through $\mathcal{P}$ not only does not coincide with $pl_b$, but it is not even a plausibility function [38].

## VI. INTERSECTION PROBABILITY

We have seen that the line $a(b, pl_b)$ is always orthogonal to $\mathcal{P}$, even in the general case. However, $(b, pl_b)$ does not intersect the probabilistic subspace, in general, but it does intersect the region of Bayesian normalized sum functions in $\varsigma[b]$ (17). But of course (since $\sum_x m_{\varsigma[b]}(x) = 1$)
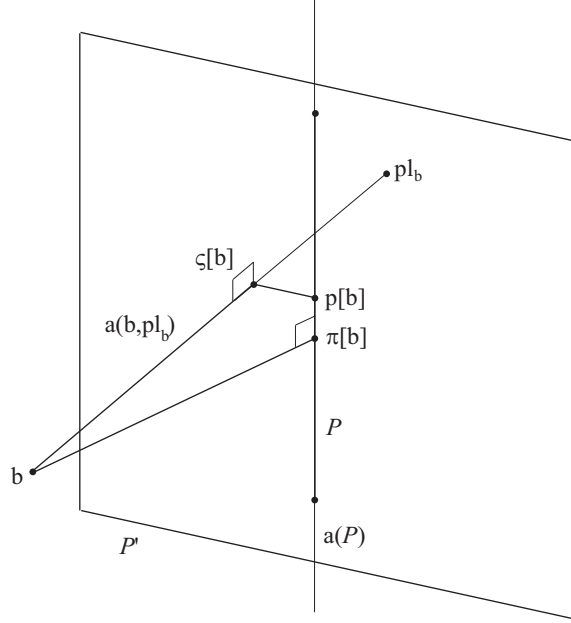
Fig. 3. The geometry of the line $a(b, pl_b)$ and the relative locations of $p[b]$, $\varsigma[b]$ and $\pi[b]$. Each b.f. $b$ and the related pl.f. $pl_b$ lie on opposite sides of the hyperplane $\mathcal{P}'$ of the Bayesian n.s.f. which divides $\mathbb{R}^{N-2}$ into two parts. The line $a(b, pl_b)$ 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.

$\varsigma[b]$ is naturally associated with a Bayesian *belief function*, assigning an equal amount of mass to each singleton and 0 to each $A : |A| > 1$, namely

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

where $m_{\varsigma[b]}(x)$ is given by Equation (19). It is easy to see that $p[b]$ is a probability, since by definition $m_{p[b]}(A) = 0$ for $|A| > 1$, $m_{p[b]}(x) = m_{\varsigma[b]}(x) \geq 0 \ \forall x \in \Theta$, and $\sum_{x \in \Theta} m_{p[b]}(x) = \sum_{x \in \Theta} m_{\varsigma[b]}(x) = 1$ by construction. The geometry of $\varsigma[b]$ and $p[b]$ with respect to the regions of Bayesian b.f and n.s.f. is sketched in Figure 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

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

are the normalization factors for $\tilde{b}$, $\tilde{pl}_b$ 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 belief function

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

measures then 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, 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 $k_{\tilde{b}}$. On the other side, the relative plausibility function can also be written in terms of $\tilde{b}$ and $R[b]$ as, by definition,

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

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

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

This means that both $\tilde{pl}_b$ also belongs $Cl(R[b], \tilde{b})$. However, as $k_{\tilde{pl}_b} \doteq \sum_{A \subset \Theta} m_b(A)|A| \geq 1$, $k_{\tilde{b}}/k_{\tilde{pl}_b} \leq k_{\tilde{b}}$ which in turn implies that $p[b]$ is closer to $R[b]$ than the relative plausibility function $\tilde{pl}_b$ (see Figure 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 Equation (23).
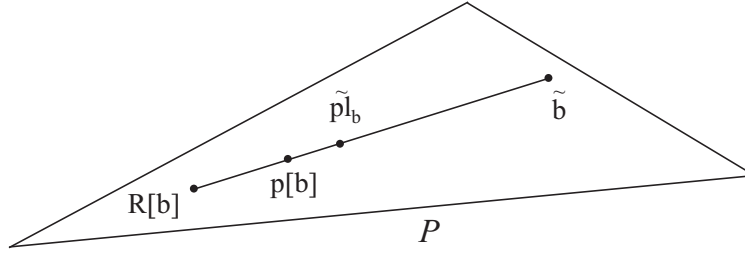
Fig. 4.   Geometric location of $p[b]$ and relative plausibility of singletons with respect to the non-Bayesianity flag. They both lie on the segment joining $R[b]$ with the relative belief of singletons $\tilde{b}$, but $\tilde{pl}_b$ is closer to $\tilde{b}$ than $p[b]$.

*2) Meaning of the ratio $\beta[b]$ and pignistic function:* To shed more light on $p[b]$ and get an alternative interpretation of this probability it may be useful to compare $p[b]$ as expressed in Equation (23) with another classical Bayesian "relative" of $b$, the pignistic function

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

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 non-singletons which is considered, and this total mass is distributed *proportionally* to their non-Bayesian contribution to each element of $\Theta$.

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

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

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 necessary and sufficient condition can be achieved by decomposing $\beta[b]$ as

$$\beta[b] = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} = \frac{\sum_{k=2}^{n} \sum_{|B|=k} m_b(B)}{\sum_{k=2}^{n} k \cdot \sum_{|B|=k} m_b(B)} = \frac{\Sigma_2 + \cdots + \Sigma_n}{2\Sigma_2 + \cdots + n\Sigma_n} \qquad (26)$$

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

*Theorem 5:* $p[b]$ and pignistic function coincide iff $\exists\ k \in [2, ..., n]$ 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

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

but that means $\beta[b] = 1/|A|$ for all $A \supset x, A \neq x$. Since $\beta[b]$ does not depend on $|A|$, this can be true only if $\exists\ k : m_b(A) = 0$ for $|A| \neq k$, and $\beta[b] = 1/k$. But this is equivalent to $\Sigma_i = 0$ for $i \neq k$, in which case $\beta[b] = \frac{\Sigma_k}{k\Sigma_k} = 1/k$. ∎

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 then a ternary frame $\Theta = \{x, y, z\}$, and a belief function $b$ with b.p.a.

$$m_b(x) = 0.1, \quad m_b(y) = 0, \quad m_b(z) = 0.2,$$

$$m_b(\{x, y\}) = 0.3, \quad m_b(\{x, z\}) = 0.1, \quad m_b(\{y, z\}) = 0, \quad m_b(\Theta) = 0.3.$$

Recalling Equation (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

$$pl_b(x) - m_b(x) = m_b(\Theta) + m_b(\{x, y\}) + m_b(\{x, z\}) = 0.7,$$

$$pl_b(y) - m_b(y) = m_b(\{x, y\}) + m_b(\Theta) = 0.6,$$

$$pl_b(z) - m_b(z) = m_b(\{x, z\}) + m_b(\Theta) = 0.4$$

so that the non-Bayesian flag is $R(x) = 0.7/1.7$, $R(y) = 0.6/1.7$, $R(z) = 0.4/1.7$. For each singleton then the original b.p.a. $m_b(x)$ is increased by a share of the mass of non singletons $1 - k_{\tilde{b}} = 0.7$ proportional to the value of $R(x)$,

$$p[b](x) = m_b(x) + (1 - k_{\tilde{b}})R(x) = 0.1 + 0.7 * 0.7/1.7 = 0.388,$$

$$p[b](y) = m_b(y) + (1 - k_{\tilde{b}})R(y) = 0 + 0.7 * 0.6/1.7 = 0.247,$$

$$p[b](z) = m_b(z) + (1 - k_{\tilde{b}})R(z) = 0.2 + 0.7 * 0.4/1.7 = 0.365.$$

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

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

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

$$p[b](x) = m_b(x) + \beta[b](pl_b(x) - m_b(x)) = 0.1 + 0.7/1.7 * 0.7,$$
$$p[b](y) = m_b(y) + \beta[b](pl_b(y) - m_b(y)) = 0 + 0.7/1.7 * 0.6,$$
$$p[b](z) = m_b(z) + \beta[b](pl_b(z) - m_b(z)) = 0.2 + 0.7/1.7 * 0.4.$$

## VII. ORTHOGONAL PROJECTION

We have seen that even though the line $a(b, pl_b)$ is always orthogonal to the probabilistic sub-space, its intersection with the region $\mathcal{P}'$ of the Bayesian n.s.f. is not always in $\mathcal{P}$. Nevertheless, an orthogonal projection $\pi[b]$ of $b$ onto $a(\mathcal{P})$ is obviously guaranteed to exist no matter the behavior of $a(b, pl_b)$, as $a(\mathcal{P})$ is nothing but a linear subspace in the space of all the normalized sum functions (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 $\Theta$.

### A. Orthogonality condition

We need to find a necessary and sufficient condition for an arbitrary vector $v = \sum_{A \subset \Theta} v_A X_A$ of 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

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

that remembering Equation (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 $a(\mathcal{P})$ will then be expressed as

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

[5]In fact the proof is again valid for $A = \Theta, \emptyset$ too, see Section VIII-B.

If the vector $v$, in particular, is a belief function ($v_A = b(A)$)

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

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 $|B \cup \{x\}| = |B \cup \{y\}| = |B| + 1$) and the equation reduces to, after erasing the common factor $2^{n-2}$,

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

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

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

or

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

From (29) we can see that $\pi[b]$ *is indeed a probability*, since both $1 + |A^c| 2^{1-|A|} \geq 0$ and $1 - |A| 2^{1-|A|} \geq 0 \ \forall |A| = 1, ..., n$. 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 probabilistic approximation of the 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. Nevertheless, if we process Equation (29) we can reduce $\pi$ to a new Bayesian function strictly related to the pignistic function.

*Theorem 7:* $\pi[b] = \bar{\mathcal{P}}(1 - k_O[b]) + k_O[b] O[b]$, where $\bar{\mathcal{P}}$ 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 belief function.

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 joining the uniform, non-informative probability and the Bayesian function $O[b]$.

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

$$\sum_{B \supset y, B \not\supset x} m_b(B)2^{1-|B|} + \sum_{B \supset y, x} m_b(B)2^{1-|B|} = \sum_{B \supset x, B \not\supset y} m_b(B)2^{1-|B|} + \sum_{B \supset y, x} m_b(B)2^{1-|B|} \equiv$$

$$\sum_{B \supset y} m_b(B)2^{1-|B|} = \sum_{B \supset x} m_b(B)2^{1-|B|} \equiv \bar{O}[b](x) = const \equiv O[b](x) = const = \bar{\mathcal{P}} \quad \forall x \in \Theta.$$

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

### C. Interpretation in terms of plausibilities, 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 belief functions

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

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 $b_{2||}$; $BetP[b]$ is the relative plausibility of singletons of $b_{||}$.

*Proof:* By definition of plausibility function

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

by Equation (38). 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 Figure 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 an unnormalized [39] belief function $b^U$

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

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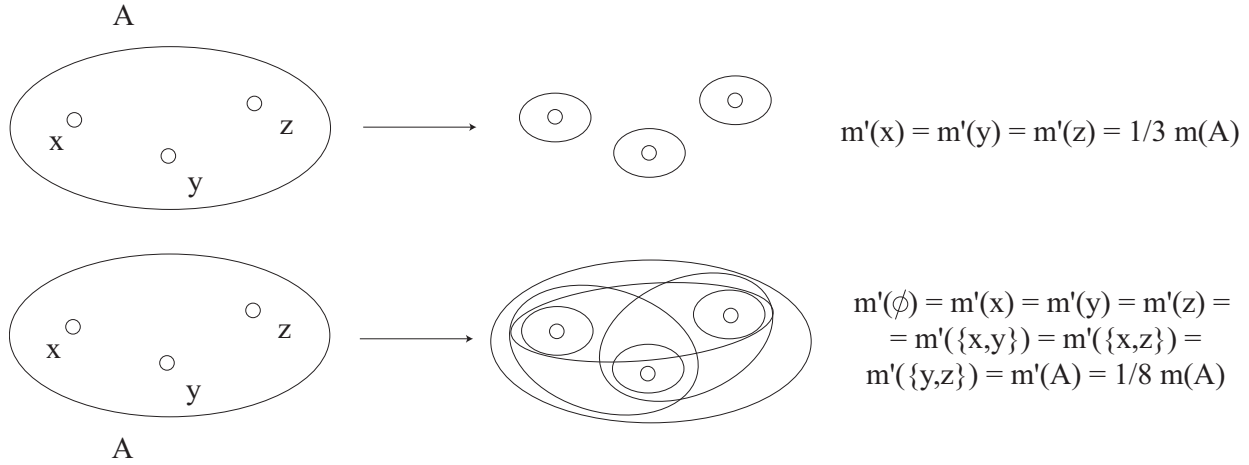


Fig. 5. Redistribution processes associated with pignistic transformation and orthogonal projection. In the pignistic transformation (top) the mass of each focal element is distributed among its elements. In the orthogonal projection (bottom), instead (through the orthogonality flag), the mass of each f.e. is divided among its subsets. In both cases, the related relative belief of singletons yields a Bayesian belief function.

*1) Example:* Let us consider again as an example the belief function $b$ on the ternary frame

$$m_b(x) = 0.1, \quad m_b(y) = 0, \quad m_b(z) = 0.2,$$
$$m_b(\{x, y\}) = 0.3, \quad m_b(\{x, z\}) = 0.1, \quad m_b(\{y, z\}) = 0, \quad m_b(\Theta) = 0.3$$

seen in Section VI-A.3. To get the orthogonality flag $O[b]$ we need to apply the redistribution process of Figure 5 to each focal element of $b$. In this case their masses are divided among their

subsets as follows:

$$m(x) = 0.1 \qquad \mapsto \quad m'(x) = m'(\emptyset) = 0.1/2 = 0.05$$

$$m(z) = 0.2 \qquad \mapsto \quad m'(z) = m'(\emptyset) = 0.2/2 = 0.1$$

$$m(\{x,y\}) = 0.3 \quad \mapsto \quad m'(\{x,y\}) = m'(x) = m'(y) = m'(\emptyset) = 0.3/4 = 0.075$$

$$m(\{x,z\}) = 0.1 \quad \mapsto \quad m'(\{x,z\}) = m'(x) = m'(z) = m'(\emptyset) = 0.1/4 = 0.025$$

$$m(\Theta) = 0.3 \qquad \mapsto \quad m'(\Theta) = m'(\{x,y\}) = m'(\{x,z\}) = m'(\{y,z\}) =$$

$$= m'(x) = m'(y) = m'(z) = m'(\emptyset) = 0.3/8 = 0.0375.$$

By summing the 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

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

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

$$\pi[b] = \bar{\mathcal{P}}(1 - k_O[b]) + k_O[b]O[b] = [1/3\ 1/3\ 1/3]' \cdot (1 - 0.4625) + 0.4625 \cdot [0.405\ 0.243\ 0.351]'$$

$$= [0.366\ 0.291\ 0.342]'.$$

### D. Orthogonal projection and convex combination

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

*Theorem 9:* Orthogonal projection and convex combination commute, i.e.

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

*Proof:* By Theorem 7 $\pi[b] = (1 - k_O[b])\bar{\mathcal{P}} + \bar{O}[b]$ 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

$$k_O[\alpha_1 b_1 + \alpha_2 b_2] = \sum_{A \subset \Theta} \big(\alpha_1 m_{b_1}(A) + \alpha_2 m_{b_2}(A)\big)|A|2^{1-|A|} = \alpha_1 k_O[b_1] + \alpha_2 k_O[b_2],$$

$$\bar{O}[\alpha_1 b_1 + \alpha_2 b_2](x) = \sum_{A \supset x} \big(\alpha_1 m_{b_1}(A) + \alpha_2 m_{b_2}(A)\big)2^{1-|A|} = \alpha_1 \bar{O}[b_1] + \alpha_2 \bar{O}[b_2]$$

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

$$\pi[\alpha_1 b_1 + \alpha_2 b_2] = (1 - \alpha_1 k_O[b_1] - \alpha_2 k_O[b_2])\bar{\mathcal{P}} + \alpha_1 \bar{O}[b_1] + \alpha_2 \bar{O}[b_2] =$$

$$= \alpha_1 \big[(1 - k_O[b_1])\bar{\mathcal{P}} + \bar{O}[b_1]\big] + \alpha_2 \big[(1 - k_O[b_2])\bar{\mathcal{P}} + \bar{O}[b_2]\big] = \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2].$$

$\blacksquare$

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

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

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

$$\bar{O}[b_A](x) = \begin{cases} 2^{1-|A|} & x \in A \\ 0 & x \notin A \end{cases} \quad \Rightarrow \quad O[b_A](x) = \begin{cases} \frac{1}{|A|} & x \in A \\ 0 & x \notin A \end{cases} = \frac{1}{|A|}\sum_{x \in A} b_x = \bar{\mathcal{P}}_A.$$

$\blacksquare$

*Theorem 10:* The orthogonal projection can be expressed as a convex combination of all the non-informative probabilities with support on a single event $A$ as

$$\pi[b] = \bar{\mathcal{P}}\Big(1 - \sum_{A \neq \Theta} \alpha_A\Big) + \sum_{A \neq \Theta} \alpha_A \bar{\mathcal{P}}_A, \quad \alpha_A \doteq m_b(A)|A|2^{1-|A|}. \tag{31}$$

*Proof:*

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

by Theorem 9, which by Lemma 2 becomes

$$\sum_{A \subset \Theta} m_b(A)\big[(1 - |A|2^{1-|A|})\bar{\mathcal{P}} + |A|2^{1-|A|}\bar{\mathcal{P}}_A\big] = \Big(1 - \sum_{A \subset \Theta} m_b(A)|A|2^{1-|A|}\Big)\bar{\mathcal{P}} +$$

$$+ \sum_{A \subset \Theta} m_b(A)|A|2^{1-|A|}\bar{\mathcal{P}}_A = \Big(1 - \sum_{A \subset \Theta} m_b(A)|A|2^{1-|A|}\Big)\bar{\mathcal{P}} + \sum_{A \neq \Theta} m_b(A)|A|2^{1-|A|}\bar{\mathcal{P}}_A +$$

$$+ m_b(\Theta)|\Theta|2^{1-|\Theta|}\bar{\mathcal{P}}$$

i.e. Equation (31). $\blacksquare$

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

$$BetP[b] = \sum_{A \subset \Theta} m_b(A) BetP[b_A], \quad \pi[b] = \sum_{A \neq \Theta} \alpha_A BetP[b_A] + \Big(1 - \sum_{A \neq \Theta} \alpha_A\Big) BetP[b_\Theta] \tag{32}$$

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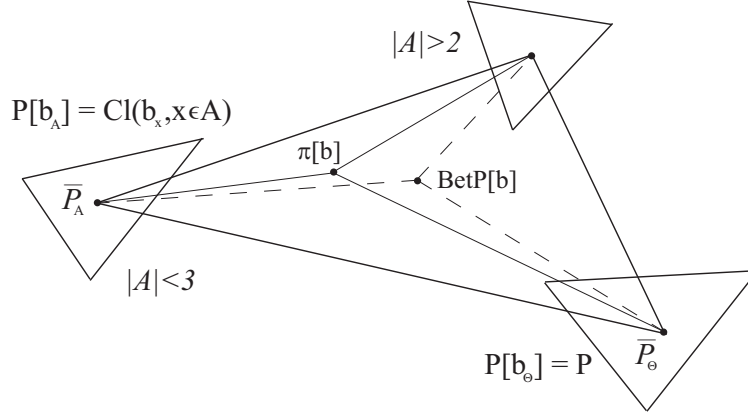
Fig. 6.   Orthogonal projection $\pi[b]$ and pignistic function $BetP[b]$ are both located on the simplex whose vertices are all the basis pignistic functions, i.e. the uniform probabilities associated with each single event $A$. However, the convex coordinates of $\pi[b]$ are weighted by a factor $k_O[b_A]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

*1) Example: ternary case:* Let us consider as an example a ternary frame $\Theta_3 = \{x, y, x\}$, and a belief function on $\Theta_3$ with b.p.a.

$$m_b(x) = 1/3, \quad m_b(\{x,z\}) = 1/3, \quad m_b(\Theta_3) = 1/3, \quad m_b(A) = 0 \; A \neq \{x\}, \{x,z\}, \Theta_3.$$

According to Equation (31)

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

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

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

i.e. $BetP[b] = \pi[b]$. This is true for each belief function $b \in \mathcal{B}_3$, since for Equation (32) when $|\Theta| = 3$ $\alpha_A = m_b(A)$ 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 $\pi$ 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.
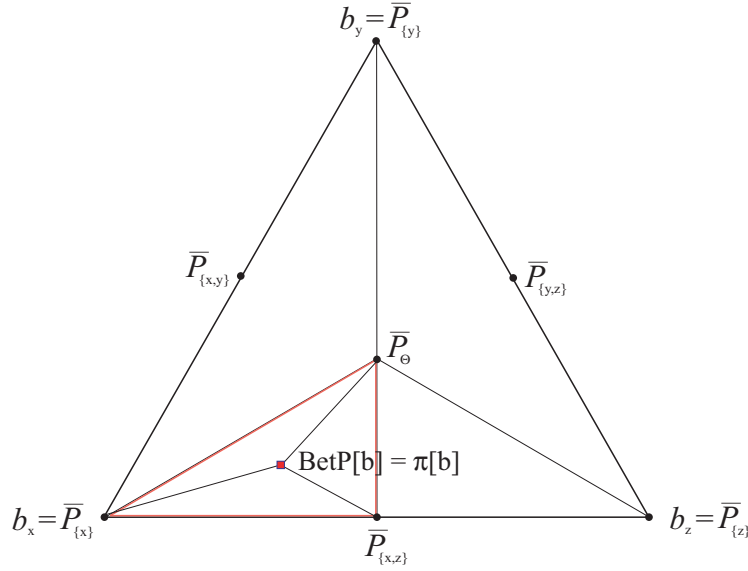
Fig. 7. Orthogonal projection and pignistic function in the ternary case $\Theta_3 = \{x, y, z\}$.

Their analytic expressions for the element $x \in \Theta$ are

$$BetP[b](x) = m_b(x) + \tfrac{1}{2}(m_b(\{x,y\}) + m_b(\{x,z\}) + m_b(\{x,w\})) +$$
$$+\tfrac{1}{3}(m_b(\{x,y,z\}) + m_b(\{x,y,w\}) + m_b(\{x,z,w\})) + \tfrac{1}{4}m_b(\Theta);$$
$$\pi[b](x) = m_b(x) + \tfrac{1}{2}(m_b(\{x,y\}) + m_b(\{x,z\}) + m_b(\{x,w\})) +$$
$$+\tfrac{5}{16}(m_b(\{x,y,z\}) + m_b(\{x,y,w\}) + m_b(\{x,z,w\})) + \tfrac{1}{16}m_b(\{y,z,w\}) + \tfrac{1}{4}m_b(\Theta).$$

(33)

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

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

we can analyze the behavior of their $L_2$ distance as $b$ varies. After introducing the simpler notation

$$y_1 = m_b(\{x,y,z\}), \quad y_2 = m_b(\{x,y,w\}), \quad y_3 = m_b(\{x,z,w\}), \quad y_4 = m_b(\{y,z,w\}),$$

we can maximize (minimize) the norm

$$\|BetP[b] - \pi[b]\|^2 = (y_1 + y_2 + y_3 - 3y_4)^2 + (y_1 + y_2 + y_4 - 3y_3)^2 + (y_1 + y_3 + y_4 - 3y_2)^2 + (y_2 + y_3 + y_4 - 3y_1)^2$$

by imposing $\frac{\partial}{\partial y_i}\|BetP[b](\mathbf{y}) - \pi[b](\mathbf{y})\|^2 = 0$ 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 (33)) $BetP[b] = \pi[b] = \bar{\mathcal{P}}$ where $\bar{\mathcal{P}} = [1/4\ 1/4\ 1/4\ 1/4]'$ is the uniform probability on $\Theta$. 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. A BRIEF DISCUSSION

In this paper we introduced two novel probabilistic approximations of a belief function $b$, the intuition for both of them provided by the geometric 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. Recalling Theorem 7 and Equation (24)

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

we can appreciate that they can both be written as convex combinations, which depend on the flag probabilities associated with them, namely the orthogonality and non-Bayesianity flag respectively:

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

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

## A. Analytic comparison

A trivial consequence of Theorem 4 is that when $b$ is 2-additive, $\pi[b] = p[b] = \varsigma[b]$. The inverse implication is also true.

*Theorem 11:* The orthogonal projection $\pi[b]$ and $p[b]$ coincide iff $b$ is 2-additive, i.e.

$$m_b(A) = 0 \quad \forall A : |A| > 2.$$

*Proof:* We just need to compare expressions (19) and (28) to see that the two quantities are the same iff

$$\begin{cases} \frac{1+|A^c|2^{1-|A|}}{n} = \beta[b] & A \supset x, A \neq x \\ m_b(A) = 0 & A \not\supset x, |A| > 2. \end{cases}$$

The second condition is true as $\frac{1-|A|2^{1-|A|}}{n} \neq 0 \; \forall A : |A| \neq 1, 2$. But now, as $\beta[b]$ does not depend on the size $|A|$ of $A$, the first condition requires $m_b(A) = 0$ for $|A| \neq k$ for some $k \in [2, ..., n]$, with

$$\frac{1 + |A^c|2^{1-|A|}}{n} = \frac{1 + (n-k)2^{1-k}}{n} = \beta[b] = \frac{1}{k}$$

(using the form (26) of $\beta[b]$). As the resulting equation $n = k(1 + (n-k)2^{1-k})$ is met by $k = 2$ only we have as desired.                                                                    ∎

Theorem 11 gives 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 the Bayesian functions we now know, $BetP, \pi, p, \tilde{pl}_b, \tilde{b}$ as $b$ varies in $\mathcal{B}$, in order to understand how they are influenced by the basic probability assignment. We started to do this for the pair $BetP[b], \pi[b]$ in the case of binary frames (Section VII-D.2), getting some interesting results. We reserve to explore this direction thoroughly in the near future.

## B. Unnormalized belief functions

We may also want to add a remark on the validity of the results presented in this paper. They have been in fact obtained for "classical" belief functions, 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 *unnormalized* belief functions (u.b.f.) [39], i.e. belief functions admitting non-zero support $m_b(\emptyset) \neq 0$ for the empty set [40]. $m_b(\emptyset)$ is an indicator of the amount of conflict in the evidence carried by a belief function $b$, but can also be interpreted as the possibility that the existing frame of discernment does not exhaust all the possible outcomes of the problem. Unnormalized b.f.

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 their validity. Let us consider again the binary case. We now have to use four coordinates, associated with all the events in $\Theta$: $\emptyset$, $\{x\}$, $\{y\}$, and $\Theta$. 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 plausibility functions 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 normalized sum functions (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:* $p[b]$, $\beta[b]$ are well defined for classical belief functions only.

---

[6]Notice that in the unnormalized case the notation $b$ is usually used for *implicability* functions, while belief functions are denoted by $Bel$ [13].

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 component $A = \emptyset, \Theta$ 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. $b$, even though $\varsigma[b] = a(b, pl_b) \cap \mathcal{P}' \neq \emptyset$ 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 well defined (check Theorem 6's proof), and it is still given by Equations (28),(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. CONCLUSIONS

In this paper we used the geometric approach to the theory of evidence to introduce two new probabilities related to a belief function, both of them 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 be in turn 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 convex combination operator, and can therefore be expressed as a convex combination of basis pignistic functions, confirming the strict relation between $\pi[b]$ and $BetP[b]$.

We finally studied the difference between $p[b]$ and $\pi[b]$, and discussed which results retain their validity for unnormalized belief functions.

We have seen when discussing the binary case that, while $BetP[b], p[b]$ and $\pi[b]$ belong to the same "family" of Bayesian relatives of $b$ (as they coincide under 2-additivity), relative plausibility $\tilde{p}[b]$ and belief $\tilde{b}$ of singletons [14] 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 belief function in the framework of Dempster's combination rule, and investigate the dual geometry of relative plausibility and belief of singletons [37]. Naturally enough the geometric approach can be adopted to study the problem of approximating a belief function with a possibility measure (consonant b.f.).

To complete the study we started in this paper we would need to depict a complete picture of the conditions under which all the different Bayesian relatives of $b$ coincide. Furthermore, a contribution to understanding their semantics could come from the study of the convex geometric loci $\{b \in \mathcal{B} : Bayes[b] = const\}$ of all the belief functions for which a certain Bayesian approximation (say $BetP[b], p[b]$ or $\pi[b]$) is constant.

## 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

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

since $\mu_b(.)$ is the Moebius inverse of $pl_b(.)$. For $\varsigma[b]$ to be a Bayesian *belief* function, accordingly, all the components related to non-singleton subsets need to be zero,

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

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

$$\mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \Big[ \sum_{|B|>1} m_b(B)|B| - \sum_{|B|>1} m_b(B) \Big] = 0 \ \forall A : |A| > 1 \ \equiv$$

$$\mu_b(A) \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>1} m_b(B)(|B| - 1) = 0. \tag{34}$$

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 expression (34) becomes

$$[\mu_b(A) + m_b(A)] \sum_{|B|>1} m_b(B) + m_b(A) \sum_{|B|>2} m_b(B)(|B| - 2) = 0 \ \forall A : |A| > 1 \ \equiv \tag{35}$$

$$[m_b(A) + \mu_b(A)] \cdot M_1[b] + m_b(A) \cdot M_2[b] = 0$$

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

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

$$M_2[b] = 0 \Leftrightarrow m_b(B) = 0 \quad \forall B : |B| > 2$$

as all the terms inside the summations are non-negative by definition of basic probability
assignment. We can distinguish three cases: $M_1 = 0 = M_2$ ($b \in \mathcal{P}$), $M_1 \neq 0$ but $M_2 = 0$, and
finally $M_1 \neq 0 \neq M_2$. If $M_1 = M_2 = 0$ then $b$ is a probability (trivially), while if $M_1 \neq 0 \neq M_2$
then Equation (35) implies $m_b(A) = \mu_b(A) = 0$, $|A| > 1$ i.e. $b \in \mathcal{P}$, which is a contradiction.

The only non-trivial case is then $M_2 = 0$, in which condition (35) becomes

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

But in this case if $|A| > 2$ then $m_b(A) = \mu_b(A) = 0$ (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

$$\beta[b] = \frac{M_1[b]}{M_2[b] + 2M_1[b]}, \tag{36}$$

if $M_2 = 0$ then $\beta[b] = 1/2$ and $\varsigma[b] = \frac{b+pl_b}{2}$.

*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 Equation (27) yields, after extracting the singletons $x$ from the summation, the system

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

$$\tag{37}$$

After replacing the first $n-1$ equations into the normalization constraint we get

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

which is equivalent to

$$n\pi(x) = 1 + (n-1)m_b(x) - \sum_{y \neq x} m_b(y) +$$

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

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 the $A$'s not containing $x$ do contain some $y \neq x$, and they are counted $|A|$ times (i.e. one time for each element they contain). Instead

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

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

$$\frac{1}{n} \left[ 1 + (n-1)m_b(x) - \sum_{y \neq x} m_b(y) - \sum_{A \not\supset x, |A| > 1} m_b(A) 2^{1-|A|} |A| + \sum_{A \supset x, |A| > 1} m_b(A) 2^{1-|A|} (n - |A|) \right]$$

$$= \frac{1}{n} \left[ n \cdot m_b(x) + 1 - \sum_{y \in \Theta} m_b(y) + n \sum_{A \supset x, |A| > 1} m_b(A) 2^{1-|A|} + \right.$$

$$\left. - \sum_{A \supset x, |A| > 1} m_b(A) 2^{1-|A|} |A| - \sum_{A \not\supset x, |A| > 1} m_b(A) 2^{1-|A|} |A| \right].$$

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

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

$$= m_b(x) + \sum_{A \supset x, |A| > 1} m_b(A) 2^{1-|A|} + \sum_{A \subset \Theta} m_b(A) \left( \frac{1 - |A| 2^{1-|A|}}{n} \right)$$

i.e. Equation (28), and since $2^{1-|A|} + \frac{1}{n} - \frac{|A|}{n} 2^{1-|A|} = \frac{1 + 2^{1-|A|}(n - |A|)}{n} = \frac{1 + 2^{1-|A|}|A^c|}{n}$ we get the second form (29).

*Proof of Theorem 7*

By Equation (29) we can write

$$\pi[b](x) = \bar{O}[b](x) + \frac{1}{n}\Big(\sum_{A\subset\Theta} m_b(A) - \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|}\Big) = \bar{O}[b](x) + \frac{1}{n}\Big(1 - k_O[b]\Big).$$

But since

$$\sum_{x\in\Theta}\bar{O}[b](x) = \sum_{x\in\Theta}\sum_{A\supset x} m_b(A)2^{1-|A|} = \sum_{A\subset\Theta} m_b(A)|A|2^{1-|A|} = k_O[b], \tag{38}$$

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

## REFERENCES

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

[2] D. Dubois and H. Prade, "Consonant approximations of belief functions," *International Journal of Approximate Reasoning*, vol. 4, pp. 419–449, 1990.

[3] A. B. Yaghlane, T. Denoeux, and K. Mellouli, "Coarsening approximations of belief functions," in *Proceedings of ECSQARU'2001*, 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. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9(4), pp. 437–460, 2001.

[5] T. Denoeux and A. B. Yaghlane, "Approximating the combination of belief functions using the fast moebius transform in a coarsened frame," *International Journal of Approximate Reasoning*, vol. 31(1-2), pp. 77–101, October 2002.

[6] R. Haenni and N. Lehmann, "Resource bounded and anytime approximation of belief function computations," *International Journal of Approximate Reasoning*, vol. 31(1-2), pp. 103–154, October 2002.

[7] T. Denouex, "Inner and outer approximation of belief structures using a hierarchical clustering approach," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9(4), pp. 437–460, 2001.

[8] M. Bauer, "Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study," *International Journal of Approximate Reasoning*, vol. 17, pp. 217–237, 1997.

[9] ——, "Approximations for decision making in the Dempster-Shafer theory of evidence," in *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, F. Horvitz, E.; Jensen, Ed., Portland, 1-4 August 1996, pp. 73–80.

[10] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61:2, pp. 315–329, 1993.

[11] J. D. Lowrance, T. D. Garvey, and T. M. Strat, "A framework for evidential-reasoning systems," in *Proceedings of the National Conference on Artificial Intelligence*, A. A. for Artificial Intelligence, Ed., 1986, pp. 896–903.

[12] P. Smets, "Belief functions versus probability functions," in *Uncertainty and Intelligent Systems*, S. L. Bouchon B. and Y. R., Eds.   Springer Verlag, Berlin, 1988, pp. 17–24.

[13] P. smets, "Decision making in the tbm: the necessity of the pignistic transformation," *International Journal of Approximate Reasoning*, vol. 38(2), pp. 133–147, February 2005.

[14] F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory," *International Journal on Man-Machine Studies*, vol. 30, pp. 525–536, 1989.

[15] B. Cobb and P. Shenoy, "On transforming belief function models to probability models," University of Kansas, School of Business, Working Paper No. 293, Tech. Rep., February 2003.

[16] B. R. Cobb and P. P. Shenoy, "A comparison of bayesian and belief function reasoning," *Information Systems Frontiers*, vol. 5(4), pp. 345–358, 2003.

[17] ——, "A comparison of methods for transforming belief function models to probability models," in *Proceedings of ECSQARU'2003, Aalborg, Denmark*, July 2003, pp. 255–266.

[18] B. Cobb and P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models," *International Journal of Approximate Reasoning*, vol. 41(3), pp. 314–330, April 2006.

[19] V. Ha and P. Haddawy, "Theoretical foundations for abstraction-based probabilistic planning," in *Proc. of the $12^{th}$ Conference on Uncertainty in Artificial Intelligence*, August 1996, pp. 291–298.

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

[21] ——, "An examination of belief functions and other monotone capacities," PhD Dissertation, Department of Statistics, Carnegie Mellon University, 1996, pgh. PA 15213.

[22] 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, California: Morgan Kaufmann, 1998, pp. 582–593. [Online]. Available: citeseer.ist.psu.edu/ha98geometric.html

[23] F. Cuzzolin and R. Frezza, "Geometric analysis of belief space and conditional subspaces," in *Proceedings of the $2^{nd}$ International Symposium on Imprecise Probabilities and their Applications (ISIPTA2001), Cornell University, Ithaca, NY*, 26-29 June 2001.

[24] F. Cuzzolin, "Geometry of upper probabilities," in *Proceedings of the $3^{rd}$ Internation Symposium on Imprecise Probabilities and Their Applications (ISIPTA'03)*, July 2003.

[25] ——, "Geometrical structure of belief space and conditional subspaces," *submitted to the IEEE Transactions on Systems, Man and Cybernetics part C*, January 2005.

[26] A. P. Dempster, "Upper and lower probability inferences based on a sample from a finite univariate population," *Biometrika*, vol. 54, pp. 515–528, 1967.

[27] A. Dempster, "Upper and lower probabilities generated by a random closed interval," *Annals of Mathematical Statistics*, vol. 39, pp. 957–966, 1968.

[28] ——, "Upper and lower probabilities inferences for families of hypothesis with monotone density ratios," *Annals of Mathematical Statistics*, vol. 40, pp. 953–969, 1969.

[29] F. Cuzzolin, "Visions of a generalized probability theory," PhD Dissertation, Università di Padova, Dipartimento di Elettronica e Informatica, 19 February 2001.

[30] ——, "Geometry of Dempster's rule of combination," *IEEE Transactions on Systems, Man and Cybernetics part B*, vol. 34:2, pp. 961–977, April 2004.

[31] M. Zaffalon and E. Fagiuoli, "Tree-based credal networks for classification." [Online]. Available: citeseer.ist.psu.edu/579709.html

[32] M. Aigner, *Combinatorial Theory*.  New York: Classics in Mathematics, Springer, 1979.

[33] A. Chateauneuf and J. Y. Jaffray, "Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion," *Mathematical Social Sciences*, vol. 17, pp. 263–283, 1989.

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

[35] P. Smets, "Constructing the pignistic probability function in a context of uncertainty," in *Uncertainty in Artificial Intelligence, 5*, M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, Eds.　Elsevier Science, 1990, pp. 29–39.

[36] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[37] F. Cuzzolin, "The geometry of relative plausibility and belief of singletons," *submitted to the International Journal of Approximate Reasoning*, 2006.

[38] ——, "The geometry of relative plausibilities," in *Proceedings of the $11^{th}$ International Conference on Information Processing and Management of Uncertainty IPMU'06, special session on "Fuzzy measures and integrals, capacities and games*.

[39] P. Smets, "The nature of the unnormalized beliefs encountered in the transferable belief model," in *Proceedings of the 8th Annual Conference on Uncertainty in Artificial Intelligence (UAI-92)*.　San Mateo, CA: Morgan Kaufmann, 1992, pp. 292–29.

[40] P. smets, "The application of the matrix calculus to belief functions," *International Journal of Approximate Reasoning*, vol. 31(1-2), pp. 1–30, October 2002.

1997