On the Properties of the Intersection Probability

Fabio Cuzzolin [a]

[a] *Perception group*
INRIA Rhône-Alpes
655, avenue de l'Europe
38334 SAINT ISMIER CEDEX, France

Abstract

In this paper, drawing inspiration from the commutativity results which hold for a number of Bayesian approximations of belief functions (like pignistic function and relative plausibility of singletons) we study the properties of a new probabilistic approximation of belief functions derived from geometric methods: the intersection probability. The intersection probability inherits its name from the fact that, when combined with a Bayesian function through Dempster's rule, it is equivalent to the intersection of the line joining a pair of belief and plausibility functions with the affine space of Bayesian pseudo belief functions. Its relation with the convex closure operator in the Cartesian space is analyzed, and equivalent conditions under which they commute are given, showing its similarity with orthogonal projection and pignistic transformation. A thorough analysis of the distance between intersection probability and pignistic function in a case study is conducted, and stringent equivalence relations in terms of mass equi-distribution inferred from it.

Key words: *Belief functions, Bayesian approximations, intersection probability, convex closure, commutativity.*

1 Introduction

The *theory of evidence* [23] was introduced in the late Seventies by Glenn Shafer as a way of representing epistemic knowledge, starting from a sequence of seminal works [14–16], of Arthur Dempster. In this formalism the best

Email address: Fabio.Cuzzolin@inrialpes.fr (Fabio Cuzzolin).
URL: http://perception.inrialpes.fr/people/Cuzzolin (Fabio Cuzzolin).

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.

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 [19] approximation of an arbitrary belief function. A number of papers [31,17,18,21,3,2] have been published on this issue, mainly in order to find efficient implementations of the rule of combination aiming to reduce the number of focal elements. Tessem [29], for instance, incorporated only the highest-valued focal elements in his m_{klx} approximation; a similar approach inspired the "summarization" technique formulated in [22]. The connection between belief functions and probabilities is as well the foundation of a popular approach to the theory of evidence, Smets' *Transferable Belief Model* [25], in which beliefs are represented at credal level (as convex sets of probabilities), while decisions are made by resorting to a Bayesian belief function called *pignistic transformation* [24]. On his side, in his 1989 paper [30] 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 through Dempster's [15] rule $\oplus$, $\tilde{pl}_b \oplus p = b \oplus p \ \forall p \in \mathcal{P}$. Cobb and Shenoy [5,6] also analyzed the properties of the relative plausibility of singletons [7] and discussed its nature of probability function that is equivalent to the original belief function.

The approximation problem can be posed in a different setting by investigating the shape of the space of belief functions, describing the location of probability functions in this space, and discussing the correct metric to use to assess the distance between a belief function and a probability. In a series of recent works [13,12,8] we proposed a geometric interpretation of the theory of evidence in which belief functions are represented as points of a simplex called *belief space* [13]. 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 $\mathbb{R}^{N-1}$.
We used the tools provided by the geometric approach to study the interplay of belief and Bayesian functions in the belief space, introducing a new probability $p[b]$ related to a belief function b, which we called *intersection probability*. $p[b]$ is the unique Bayesian belief function determined by the intersection of the line joining a b.f. b and the related pl.f. pl_b with the region of Bayesian (pseudo) b.f. [9].

Here, taking inspiration form similar properties observed in other Bayesian

approximations like pignistic function $BetP[b]$, relative plausibility $\tilde{pl}_b$ and orthogonal projection $\pi[b]$ [9], we analyze the properties of $p[b]$ with respect to two important operators acting on belief functions in the geometric framework: convex combination $Cl(.)$ and Dempster's sum.
As $p[b]$ turns out to be affine to the family of Bayesian approximations commuting with $Cl(.)$ (formed by $\pi[b]$ and $BetP[b]$) we then study the norm of their distance in a significant case study, and formulate more stringent conditions under which they coincide, based on the notion of *equidistribution*, as a step towards a full quantitative assessment of similarities and differences of all Bayesian approximations of belief functions.

1.1 Paper outline

After presenting the basic notions of the theory of evidence (Section 2) we will recall the basics of the geometric approach (Section 3), in particular the definition of belief and plausibility space. As belief functions do not represent every possible vector of the Cartesian space they are immersed in, we introduce the notion of normalized sum function (n.s.f.) (3.3) as pseudo belief function which corresponds to a point outside the belief space.
While examining the simple case of a binary frame (Section 3.4), we will notice that the line $a(b, pl_b)$ joining a pair of related belief and plausibility functions is always orthogonal to the region of Bayesian b.f. $\mathcal{P}$. However, $a(b, pl_b)$ does not intersect $\mathcal{P}$ in general, but it does indeed intersect the region $\mathcal{P}'$ of Bayesian normalized sum functions: This intersection yields a Bayesian n.s.f. $\varsigma[b]$ (Section 3.5), which is in turn naturally associated with a Bayesian belief function $p[b]$, the *intersection probability* (Section 4). $p[b]$ can be given different interpretations in terms of the way it redistributes the mass assignment of b to each element of its domain, according to the difference between plausibility and belief of each singleton.

Sections 5 and 6 form the core of the paper. After recalling important commutativity results which hold for other Bayesian approximations (5.1) (which can divided in two groups according to the operator they relate to), we analyze the properties of the intersection probability with respect to Dempster' rule (5.2) and convex combination (5.3).
As $p[b]$ perfectly represents the intersection n.s.f. $\varsigma[b]$ when combined with any Bayesian b.f., it deserves the name of "intersection probability". On the other side, it shows an interesting affinity with the pignistic transformation, as it commutes with the convex closure operator in $\mathbb{R}^{N-1}$, under certain conditions (Section 5.3.2).
In Section 6 we go further in our analysis by comparing the intersection probability with the other two Bayesian approximations which commute with $Cl(.)$, orthogonal projection and pignistic function. After studying the norm of their

difference in the case study of a ternary frame (6.1), we prove that mass equidistribution among events of the same size is a sufficient condition for their equivalence (6.2,6.3).
More stringent conditions are to be looked for in the near future.

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

2 The theory of evidence

Definition 1 *A* basic probability assignment *(b.p.a.) over a finite set (*frame of discernment *[23]) Θ is a function $m : 2^\Theta \rightarrow [0,1]$ on its power set $2^\Theta = \{A \subseteq \Theta\}$ such that*

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

Subsets of Θ associated with non-zero values of m are called *focal elements*.

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

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

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*

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

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 special 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.) $pl_b : 2^\Theta \rightarrow [0,1]$, $A \mapsto pl_b(A)$, where the plausibility $pl_b(A)$ of an event A is given by

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

and A^c denotes the complement of A in Θ.
Belief functions can be combined by means of Dempster's rule, yielding their *orthogonal sum*.

Definition 3 *The* orthogonal sum *or* Dempster's sum *of two belief functions* b_1, b_2 *is a new belief function* $b_1 \oplus b_2$ *with b.p.a.*

$$m_{b_1 \oplus b_2}(A) = \frac{\sum_{B \cap C = A} m_{b_1}(B) m_{b_2}(C)}{\sum_{B \cap C \neq \emptyset} m_{b_1}(B) m_{b_2}(C)}. \tag{3}$$

We denote by $k(b_1, b_2)$ the denominator of (3).

3 Geometry of belief and plausibility functions

In the theory of evidence the question of how to approximate a belief function with a probability or Bayesian b.f. naturally arises, specially in contexts in which point-wise estimates of a quantity of interest are needed. The approximation problem can be posed in a geometric setting by investigating the shape of the space in which belief and probability functions live. We then introduced the notion of *belief space* [13,8] as the space of all the belief functions we can define on a given domain [1].

3.1 Belief space

Given a frame of discernment Θ, a belief function $b : 2^\Theta \to [0, 1]$ is completely specified by its $N - 1$ belief values

$$\{b(A), A \subseteq \Theta, A \neq \emptyset\},$$

$N \doteq 2^{|\Theta|}$, and can then be represented as a point of $\mathbb{R}^{N-1}$. We can introduce an orthonormal reference frame $\{X_A : A \subseteq \Theta, A \neq \emptyset\}$ so that each vector $v = \sum_{A \subseteq \Theta, A \neq \emptyset} v_A X_A$ in $\mathbb{R}^{N-1}$ is potentially a belief function, in which each component v_A measures the belief value of A: $v_A = b(A)$. We call *belief space* $\mathcal{B}$ the set of points of $\mathbb{R}^{N-1}$ corresponding to a belief function. It can be proven that [13] the belief space $\mathcal{B}$ coincides with the convex closure of all the basis belief functions b_A,

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

where Cl denotes the convex closure operator:

$$Cl(b_1, ..., b_k) = \left\{ b \in \mathcal{B} : b = \alpha_1 b_1 + \cdots + \alpha_k b_k, \sum_i \alpha_i = 1,\ \alpha_i \geq 0\ \forall i \right\} \tag{5}$$

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

and

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

is the unique belief function assigning all the mass to a single subset A of Θ (A-th *basis belief function*).
More precisely, as the basis b.f. $\{b_A, \emptyset \subsetneq A \subseteq \Theta\}$ are affinely independent [2] $\mathcal{B}$ is a *simplex*, i.e. the convex closure $Cl(x_1, ..., x_{k+1})$ of $k+1$ affinely independent points $x_1, ..., x_{k+1}$ of the Cartesian space $\mathbb{R}^k$.
Each belief function $b \in \mathcal{B}$ can be written as a convex sum as

$$b = \sum_{\emptyset \subsetneq A \subseteq \Theta} m_b(A) b_A. \tag{6}$$

Geometrically, then, the b.p.a. m_b is nothing but the set of coordinates of b in the simplex $\mathcal{B}$.
The set $\mathcal{P}$ of all the Bayesian belief functions on Θ is a subset of the border of $\mathcal{B}$, precisely the simplex determined by all the basis functions associated with singletons [3]: $\mathcal{P} = Cl(b_x, x \in \Theta)$.

3.2 Plausibility space

As plausibility functions are also completely determined by their $N-1$ values $pl_b(A)$, $\emptyset \subsetneq A \subseteq \Theta$ on the power set of Θ, 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.
In [12] we proved that $\mathcal{PL}$ is again a simplex

$$\mathcal{PL} = Cl(pl_A, \emptyset \subsetneq A \subseteq \Theta), \ \ pl_A = -\sum_{B \subseteq A} (-1)^{|B|} b_B$$

whose vertices are the plausibility vectors 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

$$pl_b = \sum_{A \subseteq \Theta} \mu_b(A) b_A$$

[2] An *affine combination* of k points $v_1, ..., v_k \in \mathbb{R}^N$ is a sum $\alpha_1 v_1 + \cdots + \alpha_k v_k$ whose coefficients sum to one: $\sum_i \alpha_i = 1$. The affine subspace generated by the points $v_1, ..., v_k \in \mathbb{R}^N$ is the set $a(v_1, ..., v_k) \doteq \{v \in \mathbb{R}^N : v = \alpha_1 v_1 + \cdots + \alpha_k v_k, \sum_i \alpha_i = 1\}$. If $v_1, ..., v_k$ generate an affine space of dimension k they are said to be *affinely independent*.

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

where [12]

$$\mu_b(A) \doteq \sum_{B\subseteq A} (-1)^{|A\setminus B|} pl_b(B) = (-1)^{|A|+1} \sum_{B\supseteq A} m_b(B), \;\; A \neq \emptyset \qquad (7)$$

($\mu_b(\emptyset) = 0$) is the Moebius inverse of the plausibility function, which we call *basic plausibility assignment* (b.pl.a.).

3.3 Normalized sum functions or pseudo belief 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 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 [10]. 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 $\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ s.t. $\varsigma(A) = v_A$. As the Moebius transformation (1) is invertible, for each of these functions ς there always exists another function $m_\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ such that

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

i.e. *each* vector ς of $\mathbb{R}^{N-1}$ can be thought of as a *sum function* [1]. 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 the points of this section are sum functions meeting the normalization axiom,

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

or *normalized sum functions* (n.s.f.). n.s.f. (also called *pseudo belief functions* in the literature [27]) are natural extensions of b.f. in this geometric framework. Analogously to the case of belief functions, we can call *Bayesian normalized sum function* a n.s.f. ς such that

$$\sum_{x\in\Theta} m_\varsigma(x) = 1. \qquad (8)$$

3.4 Belief and probability in the binary case

It may be helpful to visually render these concepts in a simple example. Figure 1 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 (recalling Equation 7) is

$$\mu_b(x) = (-1)^2 \sum_{B \supseteq x} m_b(B) = m_b(x) + m_b(\Theta) = pl_b(x),$$
$$\mu_b(y) = (-1)^2 \sum_{B \supseteq y} m_b(B) = m_b(y) + m_b(\Theta) = pl_b(y)$$

and $pl_b = pl_b(x)b_x + pl_b(y)b_y$. We can notice that the two simplices are symmet-

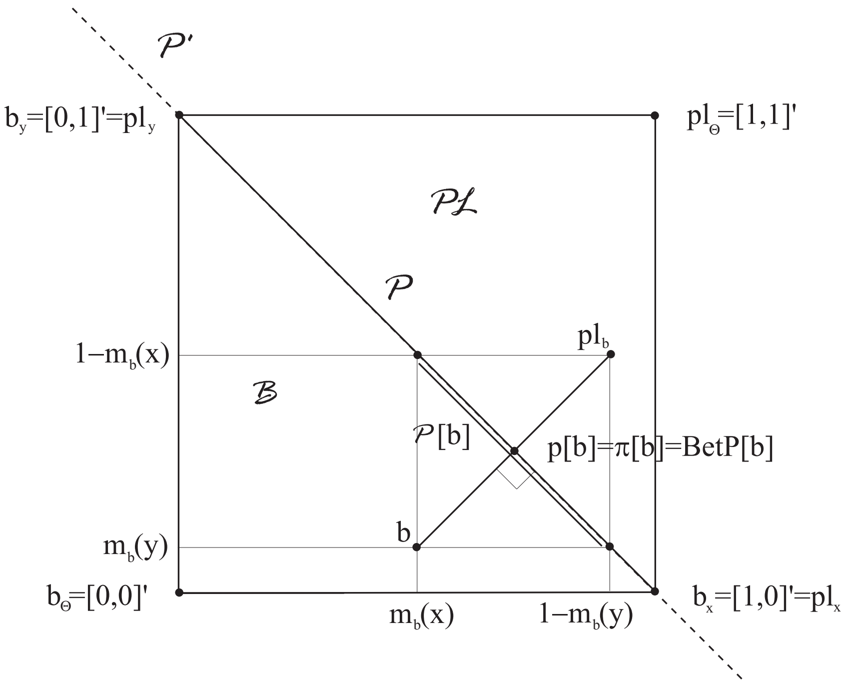

Fig. 1. In a binary frame $\Theta_2 = \{x, y\}$ both belief $\mathcal{B}$ and plausibility $\mathcal{PL}$ space are simplices with vertices $\{b_\Theta = [0, 0]', b_x = [1, 0]', b_y = [0, 1]'\}$ and $\{pl_\Theta = [1, 1]', pl_x = b_x, pl_y = b_y\}$ respectively. A b.f. b and the corresponding pl.f. pl_b are always located in symmetric positions with respect to the set $\mathcal{P}$ of probabilities on Θ. The pignistic function $BetP[b]$ and the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$ coincide with the center of the segment of consistent probabilities $\mathcal{P}[b]$, and with the intersection $p[b]$ of the line $a(b, pl_b)$ with $\mathcal{P}$.

ric 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.

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 Θ_2. The region $\mathcal{P}'$ of all the Bayesian n.s.f. is obviously (recalling Equation (8)) 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 (see Footnote 2) $a(\mathcal{P}) = a(b_x, x \in \Theta)$ generated by $\mathcal{P}$.
In the simple binary case the set of probabilities compatible with b ($\mathcal{P}[b] = \{p \in \mathcal{P} : p(A) \geq b(A)$ for all A) form a segment (1-dimensional simplex) in $\mathcal{P}$ (see Figure 1 again), whose center of mass $\bar{\mathcal{P}}$ is well known [4,20,12] to be Smets' *pignistic function* [26,28]

$$BetP[b] = \sum_{x \in \Theta} b_x \sum_{A \supseteq x} \frac{m_b(A)}{|A|} = b_x\Big(m_b(x) + \frac{m_b(\Theta)}{2}\Big) + b_y\Big(m_b(y) + \frac{m_b(\Theta)}{2}\Big). \quad (9)$$

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 $\mathcal{P}$:

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

3.5 Geometry of the dual line

When we study the *dual line* connecting a pair of belief and plausibility measures supporting the same evidence in the general case of an arbitrary frame, we realize that orthogonality still holds [9].

Proposition 1 *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. $b - pl_b \perp a(\mathcal{P})$.*

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. In fact, 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 ($n \doteq |\Theta|$), 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 1 again).

However, the dual line $a(b, pl_b)$ *does* intersect the region $\mathcal{P}'$ of Bayesian normalized sum functions (see Section 3.3 and Figure 2) in the point

$$\varsigma[b] \doteq a(b, pl_b) \cap \mathcal{P}' = b + \beta[b](pl_b - b) \quad (10)$$

with

$$\beta[b] = \frac{1 - \sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta}\Big(pl_b(x) - m_b(x)\Big)} = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|}. \tag{11}$$

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 [9]:

$$m_{\varsigma[b]}(x) = m_b(x) + \beta[b] \sum_{A \supsetneq x} m_b(A). \tag{12}$$

Equation (12) ensures that $m_{\varsigma[b]}(x)$ is positive for each $x \in \Theta$.

4 Intersection probability

Even though the line $a(b, pl_b)$ does not intersect the probabilistic subspace in the general case, since $\sum_x m_{\varsigma[b]}(x) = 1$ and $m_{\varsigma[b]}(x)$ is positive for each $x \in \Theta$, the quantity $\varsigma[b]$ is naturally associated with a Bayesian *belief* function, assigning an equal amount of mass to each singleton $m_{p[b]}(x) = m_{\varsigma[b]}(x)$ and $m_{p[b]}(A) = 0$ to each A s.t. $|A| > 1$

$$p[b] \doteq \sum_{x\in\Theta} m_{\varsigma[b]}(x) b_x \tag{13}$$

where $m_{\varsigma[b]}(x)$ is given by Equation (12). We call (13) “intersection probability” [9]. 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 2.

4.1 Interpretation

A first interpretation of this new probability, and an interesting parallelism with other two Bayesian approximations of b, are 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

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

are the normalization factors for the *relative plausibility of singletons* [30] $\tilde{pl}_b$ and the *relative belief of singletons* $\tilde{b}$ respectively:

$$\tilde{pl}_b(x) \doteq \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)}, \quad \tilde{b}(x) \doteq \frac{m_b(x)}{\sum_{y\in\Theta} m_b(y)}. \tag{14}$$

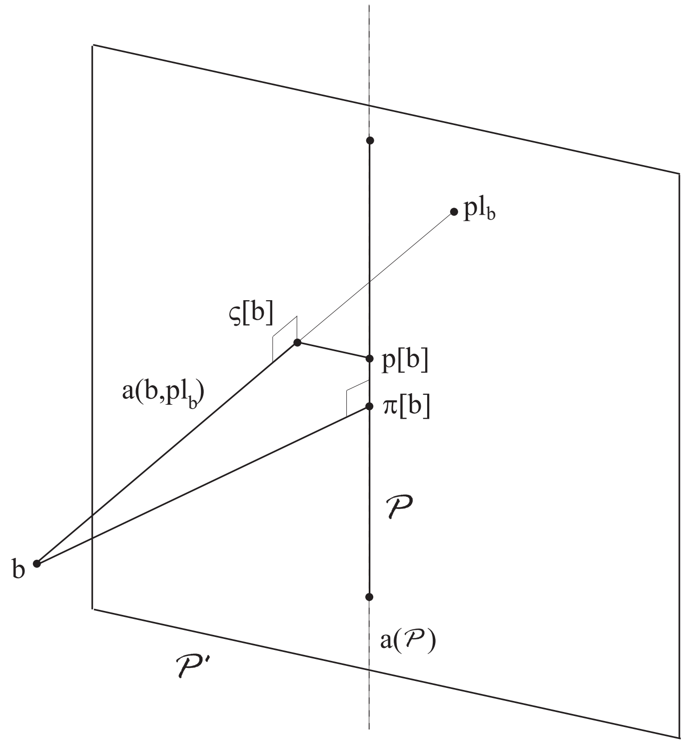

Fig. 2. The geometry of the line $a(b, pl_b)$ and the relative locations of $p[b]$, $\varsigma[b]$ and $\pi[b]$. Each b.f. b and the related pl.f. pl_b lie on opposite sides of the hyperplane $\mathcal{P}'$ of the Bayesian n.s.f. which divides $\mathbb{R}^{N-2}$ into two parts. The line $a(b, pl_b)$ 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.

$p[b]$ then becomes

$$p[b](x) = m_b(x) + (1 - k_{\tilde{b}}) \frac{pl_b(x) - m_b(x)}{k_{\tilde{pl}_b} - k_{\tilde{b}}}. \tag{15}$$

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.

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

measures then the relative contribution of each singleton x to the non Bayesianity of b. Equation (15) shows in fact that the non-Bayesian mass $1 - k_{\tilde{b}}$ of b is assigned by $p[b]$ to each singleton according to its relative contribution $R[b](x)$ to the non-Bayesianity of b.

An alternative interpretation of the intersection probability comes from comparing $p[b]$ as expressed in Equation (15) with another classical Bayesian approximation of b, the pignistic function

$$BetP[b](x) = \sum_{A \supseteq x} \frac{m_b(A)}{|A|} = m_b(x) + \sum_{A \supsetneq x} \frac{m_b(A)}{|A|}.$$

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

$$p[b](x) = m_b(x) + \beta[b](pl_b(x) - m_b(x)) \tag{16}$$

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

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

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

4.2 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.

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

The related basic plausibility assignment is, according to Equation (7),

$$\mu_b(x) = (-1)^{|x|+1} \sum_{B \supseteq \{x\}} m_b(B) =$$
$$= m_b(x) + m_b(\{x,y\}) + m_b(\{x,z\}) + m_b(\Theta) = 0.8,$$
$$\mu_b(y) = 0.6, \qquad \mu_b(z) = 0.6, \qquad \mu_b(\{x,y\}) = -0.6,$$
$$\mu_b(\{x,z\}) = -0.4, \quad \mu_b(\{y,z\}) = -0.3, \qquad \mu_b(\Theta) = +0.3.$$

As expected it sums to one, even though is not non-negative in all its components. Figure 3-top displays the events with non-zero b.p.a. (left) and b.pl.a. (right) associated with b: dashed lines indicate a negative mass.
The total mass of singletons is $k_{\tilde{b}} = 0.1 + 0 + 0.2 = 0.3$ so that the line coordinate $\beta[b]$ of the intersection $\varsigma[b]$ of the line $a(b, pl_b)$ with $\mathcal{P}'$ is equal to

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

The mass assignment of ς is therefore, by Equation (10),

$$m_{\varsigma[b]}(x) = m_b(x) + \beta[b](\mu_b(x) - m_b(x)) = 0.1 + 0.7 \cdot \tfrac{0.7}{1.7} = 0.388,$$
$$m_{\varsigma[b]}(y) = 0 + 0.6 \cdot \tfrac{0.7}{1.7} = 0.247, \qquad m_{\varsigma[b]}(z) = 0.2 + 0.4 \cdot \tfrac{0.7}{1.7} = 0.365,$$
$$m_{\varsigma[b]}(\{x,y\}) = 0.3 - 0.9 \cdot \tfrac{0.7}{1.7} = -0.071,$$
$$m_{\varsigma[b]}(\{x,z\}) = 0.1 - 0.5 \cdot \tfrac{0.7}{1.7} = -0.106,$$
$$m_{\varsigma[b]}(\{y,z\}) = 0 - 0.3 \cdot \tfrac{0.7}{1.7} = -0.123, \quad m_{\varsigma[b]}(\Theta) = 0.3 + 0 \cdot \tfrac{0.7}{1.7} = 0.3.$$

It can be noticed that all masses of singletons are indeed non negative and sum to one, while masses of non-singletons ($A \subset \Theta : |A| > 1$) *sum to zero* (-0.071 - 0.106 -0.123 + 0.3 = 0), confirming that ς is a Bayesian n.s.f. Again the mass assignment m_ς of ς has signs described by Figure 3-top-right, even though as m_ς is a weighted average of m_b and μ_b all values are closer to zero. To compare it with the intersection probability we need to recall Equation (15): 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

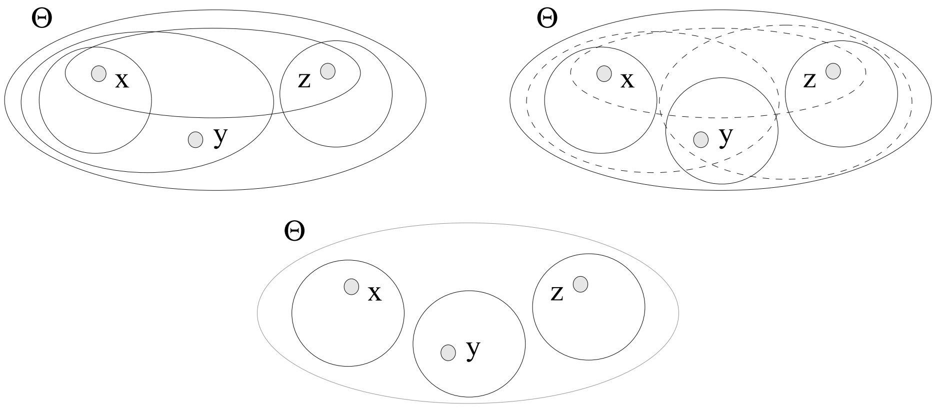

Fig. 3. Sign of non-zero masses assigned to events by the functions discussed in the example. Top left: b.p.a. associated with the belief function b of the example, with 5 focal elements. Right: the related b.pl.a. assigns positive masses to all events of size 1,3, while it assigns negative number to the events of size 2. This is the case for the mass assignment of ς too. Bottom: the intersection probability $p[b]$ retains only the masses ς gives to singletons.

share of the mass of non singletons $1 - k_{\tilde{b}} = 0.7$ proportional to the value of $R(x)$ (see Figure 3-bottom):

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

We can see that $p[b]$ coincides with ς on singletons. Equivalently, $\beta[b]$ measures the share of $pl_b(x) - m_b(x)$ assigned to each singleton:

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

5 Properties of the intersection probability with respect to Dempster's sum and convex closure

5.1 Commutativity results for Bayesian approximations

The notion of intersection probability adds another specimen to the small group of Bayesian approximations of belief functions. It is interesting to note, though, that those probability functions form in fact two families which can be

distinguished in terms of their behavior with respect to two operators acting on belief functions in the geometric approach to the ToE: namely, convex combination (5) (acting on belief functions as points of $\mathcal{B}$), and Dempster's rule of combination (3).
On one side, in fact, not only the pignistic function (9) commutes with convex combination

$$BetP[\alpha_1 b_1 + \alpha_2 b_2] = \alpha_1 BetP[b_1] + \alpha_2 BetP[b_2],$$

but a similar behavior is exhibited by the *orthogonal projection* of b onto $\mathcal{P}$ [9]

$$\pi[b](x) = \sum_{A \supseteq x} m_b(A) \Big(\frac{1 + |A^c| 2^{1-|A|}}{n} \Big) + \sum_{A \not\supset x} m_b(A) \Big(\frac{1 - |A| 2^{1-|A|}}{n} \Big),$$

namely

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

On the other side, another family of Bayesian approximations turns out to be inherently related to Dempster's combination rule. Both relative plausibility and belief of singletons (14) commute as a matter of fact with the orthogonal sum. Namely, Cobb and Shenoy [6] proved that the relative plausibility function $\tilde{pl}_b$ commutes with respect to Dempster's sum $\oplus$

$$\tilde{pl}_{b_1 \oplus b_2} = \tilde{pl}_{b_1} \oplus \tilde{pl}_{b_2}.$$

Concerning relative belief (14-right), it can proven that [10] Dempster's rule defined as in Equation (3) can be applied to a pair of normalized sum functions, yielding a new normalized sum function. As a matter of fact, as we have seen in Section 3.3, the Moebius inversion lemma (1) can be applied to a n.s.f. ς, yielding a "mass" function m_ς. Dempster's rule (3) can then be applied to a pair of such mass functions m_{ς_1}, m_{ς_2}. We can still denote the orthogonal sum of two sum functions ς_1, ς_2 with $\varsigma_1 \oplus \varsigma_2$.
Now, plausibility functions are n.s.f., since they comply with the normalization constraint $pl_b(\Theta) = 1$ for all b. Their mass function is in fact the basic plausibility assignment μ_b (7). Dempster's rule can then be formally applied to pl.f. too.
It can be proven that [11] the relative belief operator commutes with respect to Dempster's combination of plausibility functions:

$$\tilde{b}[pl_1 \oplus pl_2] = \tilde{b}[pl_1] \oplus \tilde{b}[pl_2].$$

In this part of the paper we will then study the behavior of $p[b]$ with respect to both $\oplus$ and Cl, showing that the intersection probability can be assimilated to the pair pignistic function - orthogonal projection even though it does not

commute with Cl in all cases.
We will indeed provide the conditions under which commutativity holds, and discuss their meaning in terms of degrees of belief. To deepen our understanding of the relations between those Bayesian functions which belong to the same family, their distance will be evaluated in the case study of a ternary frame as it is been done for the orthogonal projection [9].

5.2 A justification of the name "intersection probability": the representation theorem

A first application of the generalization of Dempster's rule to n.s.f. brings about a justification of the name "intersection probability" for $p[b]$, as it turns out that $p[b]$ and $\varsigma[b]$ are *equivalent* when combined with a Bayesian belief function.

We first need to recall that [10] the orthogonal sum $b\oplus(\alpha_1 b_1+\alpha_2 b_2)$, $\alpha_1+\alpha_2 = 1$ of a b.f. b and any affine combination of other belief functions reads as

$$b \oplus (\alpha_1 b_1 + \alpha_2 b_2) = \gamma_1 (b \oplus b_1) + \gamma_2 (b \oplus b_2) \tag{17}$$

where

$$\gamma_1 = \frac{\alpha_1 k(b, b_1)}{\alpha_1 k(b, b_1) + \alpha_2 k(b, b_2)} \quad \gamma_2 = \frac{\alpha_2 k(b, b_2)}{\alpha_1 k(b, b_1) + \alpha_2 k(b, b_2)}$$

and $k(b, b_i)$ is the normalization factor of the combination between b and b_i. On the other side we already know that $\varsigma[b]$ can be expressed as a convex combination of b and pl_b,

$$\varsigma[b] = \beta[b] pl_b + (1 - \beta[b]) b$$

(Equation (10)). The Bayesian b.f. $p[b]$ can be instead expressed as a convex combination of two n.s.f. By Equation (16)

$$\begin{aligned} p[b] &= \sum_{x\in\Theta} m_b(x) b_x + \beta[b] \sum_{x\in\Theta} (pl_b(x) - m_b(x)) b_x = \\ &= (1 - \beta[b]) \sum_{x\in\Theta} m_b(x) b_x + \beta[b] \sum_{x\in\Theta} pl_b(x) b_x \end{aligned}$$

and if we call the quantities

$$\bar{pl}_b \doteq \sum_{x\in\Theta} pl_b(x) b_x \qquad \bar{b} = \sum_{x\in\Theta} m_b(x) b_x \tag{18}$$

plausibility of singletons and *belief of singletons* respectively, we have

$$p[b] = \beta[b] \bar{pl}_b + (1 - \beta[b]) \bar{b} \tag{19}$$

i.e. $p[b]$ lies on the line joining $\bar{pl}_b$ and $\bar{b}$, in the same relative position of $\varsigma[b]$ on the segment $Cl(b, pl_b)$.

Plausibility and belief of singletons are both normalized sum functions, as (since $b_\Theta = \mathbf{0}$ is the nil vector in $\mathbb{R}^{N-2}$) they read as

$$\bar{pl}_b = \sum_{x \in \Theta} pl_b(x) b_x + (1 - k_{\tilde{pl}_b}) b_\Theta, \quad \bar{b} = \sum_{x \in \Theta} m_b(x) b_x + (1 - k_{\tilde{b}}) b_\Theta$$

under which form they meet the normalization constraint (trivial). More precisely, $\bar{pl}_b$ is a n.s.f. assigning mass $1 - k_{\tilde{pl}_b} \leq 0$ to the focal element Θ, while $\bar{b}$ is a b.f. assigning to Θ the mass $1 - k_{\tilde{b}} \geq 0$.
Equation (19) can then be used to prove that

Theorem 1 *The combinations of $p[b]$ and $\varsigma[b]$ with any probability function $p \in \mathcal{P}$ coincide,*

$$p[b] \oplus p = \varsigma[b] \oplus p \quad \forall p \in \mathcal{P}.$$

Theorem 1 is the formal justification of the name "intersection probability" we chose for the quantity $p[b]$. In fact, even though $p[b]$ is *not* the actual intersection between the line $a(b, pl_b)$ and the region of Bayesian n.s.f. (which is ς), as a matter of fact it behaves exactly like ς when combined with a probability.
However, $p[b]$ does not seem to have any other meaningful relations with the orthogonal sum. This is the case, though, for the convex closure operator.

5.3 Intersection probability and convex closure

On the other side, we have seen above that $p[b]$ and $BetP[b]$ are closely related Bayesian functions, the link between them being the role of the quantity $\beta[b]$. It is natural to wonder if $p[b]$ exhibits a similar behavior with respect to the convex closure operator.
Indeed, even though the situation is a bit more complex in this second case, $p[b]$ turns also out to be related to $Cl(.)$ in a rather elegant way. Let us introduce the notation $\beta[b_i] = N_i/D_i$.

Theorem 2

$$p[\alpha_1 b_1 + \alpha_2 b_2] = \widehat{\alpha_1 D_1}(\alpha_1 p[b_1] + \alpha_2 T[b_1, b_2]) + \widehat{\alpha_2 D_2}(\alpha_1 T[b_1, b_2]) + \alpha_2 p[b_2]), \quad (20)$$

where $\widehat{\alpha_i D_i} = \frac{\alpha_i D_i}{\alpha_1 D_1 + \alpha_2 D_2}$ *and* $T[b_1, b_2]$ *is the probability with values*

$$T[b_1, b_2](x) = \hat{D}_1 p[b_2, b_1] + \hat{D}_2 p[b_1, b_2] \quad (21)$$

with $\hat{D}_i = \frac{D_i}{D_1+D_2}$,

$$p[b_2, b_1](x) = m_{b_2}(x) + \beta[b_1](pl_{b_2}(x) - m_{b_2}(x)),$$
$$p[b_1, b_2](x) = m_{b_1}(x) + \beta[b_2](pl_{b_1}(x) - m_{b_1}(x)). \quad (22)$$

Geometrically, $p[\alpha_1 b_1 + \alpha_2 b_2]$ can be constructed as in Figure 4 as a point of the simplex $Cl(T[b_1, b_2], p[b_1], p[b_2])$. The point $\alpha_1 T[b_1, b_2] + \alpha_2 p[b_2]$ is the intersection of the segment $Cl(T, p[b_2])$ with the line l_2 passing through $\alpha_1 p[b_1] + \alpha_2 p[b_2]$ and parallel to $Cl(T, p[b_1])$. Dually, $\alpha_2 T[b_1, b_2] + \alpha_1 p[b_1]$ is the intersection of the segment $Cl(T, p[b_1])$ with the line l_1 passing through $\alpha_1 p[b_1] + \alpha_2 p[b_2]$ and parallel to $Cl(T, p[b_2])$. $p[\alpha_1 b_1 + \alpha_2 b_2]$ is finally the point of the segment

$$Cl(\alpha_1 T + \alpha_2 p[b_2], \alpha_2 T + \alpha_1 p[b_1])$$

with convex coordinate $\widehat{\alpha_1 D_1}$ (or equivalently $\widehat{\alpha_2 D_2}$).

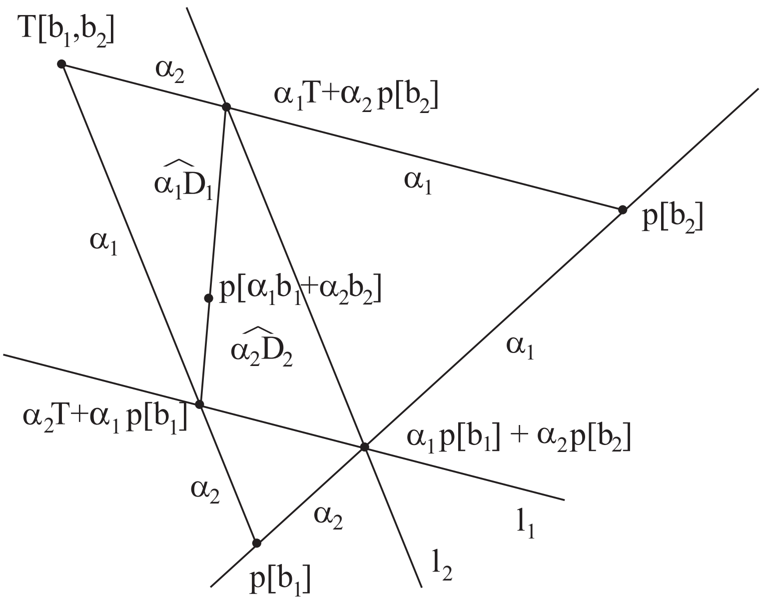

Fig. 4. Behavior of the probabilistic approximation $p[b]$ under convex combination. $\alpha_2 T + \alpha_1 p[b_1]$ and $\alpha_1 T + \alpha_2 p[b_2]$ lie on inverted locations on the segments joining $T[b_1, b_2]$ and $p[b_1]$, $p[b_2]$ respectively: $\alpha_i p[b_i] + \alpha_j T$ is the intersection of the line $Cl(T, p[b_i])$ with the parallel to $Cl(T, p[b_j])$ passing through $\alpha_1 p[b_1] + \alpha_2 p[b_2]$. $p[\alpha_1 b_1 + \alpha_2 b_2]$ is finally the point of the segment joining them with convex coordinate $\widehat{\alpha_i D_i}$.

5.3.1 *Location of $T[b_1, b_2]$ in the binary case*

As an example, let us consider the location of $T[b_1, b_2]$ in the binary belief space $\mathcal{B}_2$. In the binary case, in fact,

$$\beta[b_1] = \beta[b_2] = \frac{m_{b_i}(\Theta)}{2m_{b_i}(\Theta)} = 1/2$$

$\forall b_1, b_2 \in \mathcal{B}_2$ and $p[b]$ always commutes with the convex closure operator. Accordingly, $T[b_1, b_2](x)$ is equal to

$$\frac{m_{b_1}(\Theta)}{m_{b_1}(\Theta) + m_{b_2}(\Theta)} \left[m_{b_2}(x) + \frac{m_{b_2}(\Theta)}{2} \right] + \frac{m_{b_2}(\Theta)}{m_{b_1}(\Theta) + m_{b_2}(\Theta)} \cdot$$
$$\cdot \left[m_{b_1}(x) + \frac{m_{b_1}(\Theta)}{2} \right] = \frac{m_{b_1}(\Theta)}{m_{b_1}(\Theta) + m_{b_2}(\Theta)} p[b_2] + \frac{m_{b_2}(\Theta)}{m_{b_1}(\Theta) + m_{b_2}(\Theta)} p[b_1].$$

Looking at Figure 5, simple trigonometric considerations show that the segment $Cl(p[b_i], T[b_1, b_2])$ has length $\frac{m_i(\Theta)}{\sqrt{2}\tan\phi}$, where ϕ is the angle between the segments $Cl(b_i, T)$ and $Cl(p[b_i], T)$.
$T[b_1, b_2]$ is then the unique point of $\mathcal{P}$ such that the angles $\widehat{b_1 T p[b_1]}$ and $\widehat{b_2 T p[b_2]}$ coincide, i.e. T is the intersection of $\mathcal{P}$ with the line passing through b_i and the reflection of b_j through $\mathcal{P}$.
But now this reflection is nothing but pl_{b_j}, so that

$$T[b_1, b_2] = Cl(b_1, pl_{b_2}) \cap \mathcal{P} = Cl(b_2, pl_{b_1}) \cap \mathcal{P}.$$

5.3.2 *Equivalent condition for commutativity*

Even though the intersection probability does not commute with convex closure, $p[b]$ can still be considered affine to orthogonal projection and pignistic function. Theorem 3 states the conditions under which $p[b]$ and Cl commute.

Theorem 3 *$p[b]$ and convex closure commute iff*

$$T[b_1, b_2] = \hat{D}_1 p[b_2] + \hat{D}_2 p[b_1]$$

or, equivalently, $\beta[b_1] = \beta[b_2]$ or $R[b_1] = R[b_2]$.

As a geometric confirmation of Theorem 3, only when the two lines l_1, l_2 in Figure 4 are parallel to $a(p[b_1], p[b_2])$ (i.e. $T[b_1, b_2] \in Cl(p[b_1], p[b_2])$, compare above) the desired quantity $p[\alpha_1 b_1 + \alpha_2 b_2]$ belongs to $Cl(p[b_1], p[b_2])$ (i.e. is also a convex combination of $p[b_1]$ and $p[b_2]$).
Theorem 3 reflects the two complementary interpretations of $p[b]$ we gave in

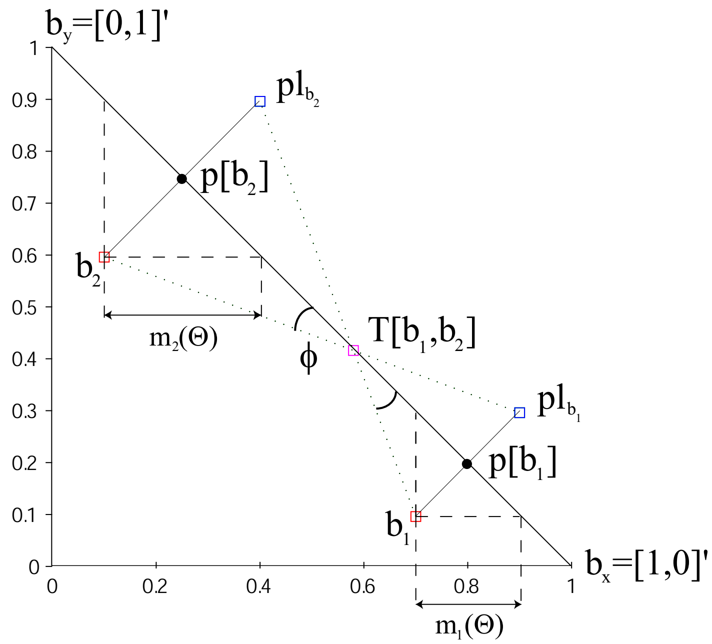

Fig. 5. Location of the probability function $T[b_1, b_2]$ in the binary belief space.

terms of $\beta[b]$ and $R[b]$ (Equations (15) and (16)):

$$p[b] = m_b(x) + (1 - k_{\tilde{b}})R[b](x), \;\; p[b] = m_b(x) + \beta[b](pl_b(x) - m_b(x)).$$

If $\beta[b_1] = \beta[b_2]$ the two b.f. attribute to each singleton the same share of its non-Bayesian contribution. If $R[b_1] = R[b_2]$ the non-Bayesian mass $1 - k_{\tilde{b}}$ is distributed in the same way to the elements of Θ.

5.3.3 Commutativity condition in the ternary case

To better understand the meaning of Theorem 3 in terms of belief values of the involved functions b_1 and b_2, let us consider the ternary case (as we have seen above that the commutativity condition is met by all pairs of b.f. defined on a binary frame).
In this case, the condition $\beta[b_1] = \beta[b_2]$ becomes in fact

$$\frac{\sum_{|A|>1} m_1(A)}{\sum_{|A|>1} m_1(A)|A|} = \frac{\sum_{|A|>1} m_2(A)}{\sum_{|A|>1} m_2(A)|A|} \equiv$$
$$\frac{m_1(\{x,y\}) + m_1(\{x,z\}) + m_1(\{y,z\}) + m_1(\Theta)}{2\, m_1(\{x,y\}) + 2\, m_1(\{x,z\}) + 2\, m_1(\{y,z\}) + 3\, m_1(\Theta)} =$$
$$= \frac{m_2(\{x,y\}) + m_2(\{x,z\}) + m_2(\{y,z\}) + m_2(\Theta)}{2\, m_2(\{x,y\}) + 2\, m_2(\{x,z\}) + 2\, m_2(\{y,z\}) + 3\, m_2(\Theta)}.$$

This can be simplified once we denote by σ_i^j the quantity $\sum_{|A|=j} m_i(A)$. The above equation then becomes

$$\frac{\sigma_1^2 + \sigma_1^3}{2\sigma_1^2 + 3\sigma_1^3} = \frac{\sigma_2^2 + \sigma_2^3}{2\sigma_2^2 + 3\sigma_2^3}$$

which after a few passages reduces to

$$\sigma_1^2\sigma_2^3 = \sigma_1^3\sigma_2^2.$$

In nontrivial cases ($b \not\in \mathcal{P}$) either σ^2 or σ^3 are nonzero for both b_1 and b_2. However, if b_1 (b_2) has only size 2 f.e. and b_2 (b_1) only size 3 f.e. the above equality is never met. Hence there has to exist a cardinality $k \in \{2, 3\}$ such that $\sigma_1^k \neq 0 \neq \sigma_2^k$. In this case

$$\frac{\sigma_1^2}{\sigma_1^3} = \frac{\sigma_2^2}{\sigma_2^3}, \ \sigma_1^3, \sigma_2^3 \neq 0 \quad \Big| \quad \frac{\sigma_1^3}{\sigma_1^2} = \frac{\sigma_2^3}{\sigma_2^2}, \ \sigma_1^2, \sigma_2^2 \neq 0$$

i.e. the ratio between the total mass the two b.f. assign to events of size 2 and 3 is the same for the two belief functions.

5.3.4 A sufficient condition for commutativity

The discussion of the binary case of Section 5.3.3 gives as an intuition on how to provide a sufficient condition for the commutativity of $p[.]$ and $Cl(.)$, which has a meaningful interpretation in terms of belief values. Let us then consider the following decomposition of $\beta[b]$:

$$\beta[b] = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} = \frac{\sum_{k=2}^n \sum_{|B|=k} m_b(B)}{\sum_{k=2}^n k \cdot \sum_{|B|=k} m_b(B)} = \frac{\sigma_2 + \cdots + \sigma_n}{2\sigma_2 + \cdots + n\sigma_n} \tag{23}$$

where $\sigma_k \doteq \sum_{|B|=k} m_b(B)$. $\beta[b_1] = \beta[b_2]$ is then equivalent to

$$(2\sigma_2^2 + ... + n\sigma_2^n)(\sigma_1^2 + ... + \sigma_1^n) = (2\sigma_1^2 + ... + n\sigma_1^n)(\sigma_2^2 + ... + \sigma_2^n).$$

Now, as $b_1, b_2 \not\in \mathcal{P}$ in nontrivial cases, there exists at least two cardinalities k_1, k_2 for which $\sigma_1^{k_1} \neq 0$ and $\sigma_2^{k_2} \neq 0$, respectively.
Let us assume that there exists a k such that $\sigma_1^k \neq 0 \neq \sigma_2^k$.
We can then divide the two sides by σ_1^k and σ_2^k, obtaining

$$\left(2\frac{\sigma_2^2}{\sigma_2^k} + .. + k + ... + n\frac{\sigma_2^n}{\sigma_2^k}\right)\left(\frac{\sigma_1^2}{\sigma_1^k} + .. + 1 + ... + \frac{\sigma_1^n}{\sigma_1^k}\right) =$$
$$= \left(2\frac{\sigma_1^2}{\sigma_1^k} + .. + k + ... + n\frac{\sigma_1^n}{\sigma_1^{k_1}}\right)\left(\frac{\sigma_2^2}{\sigma_2^k} + .. + 1 + ... + \frac{\sigma_2^n}{\sigma_2^k}\right)$$

so that, if $\sigma_1^j/\sigma_1^k = \sigma_2^j/\sigma_2^k\ \forall j \neq k$ the condition is met.
But this is equivalent to require that

$$\frac{\sigma_1^l}{\sigma_1^m} = \frac{\sigma_2^l}{\sigma_2^m} \quad \forall l, m \geq 2\ s.t.\ \sigma_1^m, \sigma_2^m \neq 0$$

in which case $\beta[b_1]$, $\beta[b_2]$ are equal. In other words,

Theorem 4 *If the ratio between the total mass of focal elements of different size is the same for all the b.f. involved, then intersection probability (considered as an operator mapping belief functions to probabilities) and convex ccombination commute.*

On the other side, if $m_1(A) = 0$ for $|A| \neq k$ and $m_2(A) = 0$ for $|A| \neq l$, with $k \neq l$ (b_1 and b_2 are focused only on f.e. of two different cardinalities), we get

$$\frac{\sigma_1^k}{k\sigma_1^k} = \frac{\sigma_2^l}{l\sigma_2^l} \equiv \frac{1}{k} = \frac{1}{l}$$

which is impossible, so that $\beta[b_1] \neq \beta[b_2]$.

6 Comparing the intersection probability with other Bayesian approximations

We have seen that the intersection probability is a member of a family of Bayesian approximations which includes also pignistic function and orthogonal projection. The next natural step is then to discuss the difference between $p[b]$ and its "sister" functions $BetP$ and π. This is the aim of the last part of this paper.
As a matter of fact, some sufficient conditions [9] have been already worked out for the pignistic function and the orthogonal projection.

Proposition 2 *Intersection probability and orthogonal projection coincide if b is 2-additive, i.e. $m_b(A) = 0$ for all $A : |A| > 2$.*

A similar sufficient condition for the pair $p[b]$, $BetP[b]$ can be found by resorting to the decomposition of $\beta[b]$ of Equation (23).

Proposition 3 *Intersection probability and pignistic function coincide if $\exists\ k \in [2, ..., n]$ such that $\sigma^i = 0\ \forall\ i \neq k$, i.e. the focal elements of b have size 1 or k only.*

However, it is possible to go further, and formulate more stringent conditions. Also, those kind of results give only "pointwise" indications of the behavior

of each pair of approximations. Ideally it would be desirable to analyze the *distance* between those pairs of probabilities as the belief function b varies in the belief space, in order to fully understand their different behaviors in terms of belief values.
Let us then pose such a study in a significant but still simple situation, that of a belief function defined on a ternary frame.

6.1 A case study: the ternary frame

In the ternary frame $\pi[b] = BetP[b]$ [9], so that checking whether $p[b] = BetP[b]$ is equivalent to check the analogous condition for the pignistic function. We will consider three different norms when measuring distances in the probability simplex, namely

$$\|p - p'\|_1 \doteq \sum_{x\in\Theta} |p(x) - p'(x)|, \qquad \|p - p'\|_2 \doteq \sqrt{\sum_{x\in\Theta} |p(x) - p'(x)|^2}, \\ \|p - p'\|_\infty \doteq \max_{x\in\Theta} |p(x) - p'(x)|, \tag{24}$$

i.e. the classical L_1, L_2, and L_∞ norms. Let us then recall the expressions of $p[b]$ and $BetP[b]$ in the ternary frame:

$$\begin{aligned} p[b](x) &= m_b(x) + \frac{\sigma^2+\sigma^3}{2\sigma^2+3\sigma^3}(m_b(\{x,y\}) + m_b(\{x,z\}) + m_b(\Theta)) = \\ &= m_b(x) + \frac{\sigma^2+\sigma^3}{2\sigma^2+3\sigma^3}(\sigma^2 - m_b(x^c) + \sigma^3); \\ BetP[b](x) &= m_b(x) + \frac{1}{2}(m_b(\{x,y\}) + m_b(\{x,z\})) + \frac{1}{3}m_b(\Theta) = \\ &= m_b(x) + \frac{1}{2}(\sigma^2 - m_b(x^c)) + \frac{1}{3}\sigma^3 \end{aligned}$$

$\forall x \in \Theta$, once defined as above $\sigma^k \doteq \sum_{|A|=k} m_b(A)$. Immediately

$$\begin{aligned} p[b](x) - BetP[b](x) &= (\sigma^2 - m_b(x^c))\Big(\frac{\sigma^2+\sigma^3}{2\sigma^2+3\sigma^3} - \frac{1}{2}\Big) + \\ +\sigma^3\Big(\frac{\sigma^2+\sigma^3}{2\sigma^2+3\sigma^3} - \frac{1}{3}\Big) &= (\sigma^2 - m_b(x^c))\frac{-\sigma^3/2}{2\sigma^2+3\sigma^3} + \sigma^3\frac{\sigma^2/3}{2\sigma^2+3\sigma^3} = \\ &= \frac{\sigma^3}{2\sigma^2+3\sigma^3}\Big(\frac{\sigma^2}{3} - \frac{\sigma^2 - m_b(x^c)}{2}\Big) \end{aligned}$$

so that the difference between $p[b](x)$ and $\pi[b] = BetP[b](x)$ in the ternary frame is

$$p[b](x) - \pi[b](x) = \frac{\sigma^3}{6}\frac{1}{2\sigma^2+3\sigma^3}(3m_b(x^c) - \sigma^2). \tag{25}$$

Clearly, the minima of all norms correspond to belief functions b for which the difference (25) goes to zero for all singletons $x \in \Theta$, i.e.

$$\sigma^3 = m_b(\Theta) = 0 \quad | \quad 3\, m_b(x^c) = \sigma^2 \;\; \forall x$$

or equivalently

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

In other words, the difference is minimal when either b is 2-additive or it assigns equal mass to non-singleton events of the same cardinality.

To find the maxima of the norms of the difference, instead, we need to explicitly compute them.
It is easy to see that, for each norm L_p, $p = 1, 2, \infty$, the norm of the difference between intersection probability and orthogonal projection $\|p[b] - \pi[b]\|_p$ has the following form

$$\|p[b] - \pi[b]\|_p = f(\sigma^2, \sigma^3) \cdot g_p(b)$$

where (for $\sigma^2 = const$) $g_p(b) = g(u, v, w)$ is a function of

$$u = m_b(\{x,y\}), \qquad v = m_b(\{x,z\}), \qquad w = m_b(\{y,z\})$$

and has maximum when the mass is concentrated on one of the three events:

$$\begin{array}{r}
\|p[b] - \pi[b]\|_1 = \frac{2}{3}\frac{\sigma^2\sigma^3}{2\sigma^2+3\sigma^3} \cdot (|3w - \sigma^2| + |3v - \sigma^2| + |3u - \sigma^2|); \\
\|p[b] - \pi[b]\|_2 = \frac{2}{3}\frac{\sigma^2\sigma^3}{2\sigma^2+3\sigma^3} \cdot \sqrt{(3u - \sigma^2)^2 + (3v - \sigma^2)^2 + (3w - \sigma^2)^2}; \\
\|p[b] - \pi[b]\|_\infty = \frac{1}{3}\frac{\sigma^2\sigma^3}{2\sigma^2+3\sigma^3} \cdot \max\{|3w - \sigma^2|, |3v - \sigma^2|, |3u - \sigma^2|\}.
\end{array}$$

Let us then consider the behavior of the factor $g(u, v, w)$ for each norm.

For L_1 the factor $g_1(u, v, w) = (|3w - \sigma^2| + |3v - \sigma^2| + |3u - \sigma^2|)$ can be rewritten after changing variables as

$$g_1(U, V) = |U| + |V| + |U + V|$$

where $U \doteq 3u - \sigma^2$, $V \doteq 3v - \sigma^2$, with constraints $U, V \geq -\sigma^2$, $-2\sigma^2 \leq U + V \leq \sigma^2$. As it can be seen from the plot of Figure 6-left, $g_1(U, V)$ has three maxima in

$$(U, V) = (-\sigma^2, 2\sigma^2), (-\sigma^2, -\sigma^2), (2\sigma^2, -\sigma^2) \tag{26}$$

which correspond to

$$(u, v, w) = (0, \sigma^2, 0), (\sigma^2, 0, 0), (0, 0, \sigma^2) \tag{27}$$

with maximum value $\max g_1 = 4\sigma^2$.
A similar substitution for the L_2 case yields (Figure 6-right)

Fig. 6. Left: plot of the factor $g_1(U, V)$ in the norm L_1 of the difference between intersection probability and orthogonal projection, ternary case. Right: plot of $g_2(U, V)$.

$$g_2(U, V) = \sqrt{U^2 + V^2 + (U + V)^2}$$

with maxima in (26) again, $\max g_2(b) = \sqrt{2}\sigma^2$.
Finally, for L_∞, it is easy to see that, again,

$$\arg\max_b g_\infty(b) = b : \exists x \in \Theta \ s.t. \ m_b(x^c) = \sigma^2, \qquad \max g_\infty(b) = 2\sigma^2$$

The three norms *have the same maxima and minima*, no matter the different form of the three functions, giving a consistent picture of the distance between intersection probability and pignistic function.

We are then left with the study of the factor

$$f(\sigma^2, \sigma^3) \doteq \frac{\sigma^2\sigma^3}{2\sigma^2 + 3\sigma^3}, \qquad \sigma^2 + \sigma^3 \leq 1, \qquad \sigma^2, \sigma^3 \geq 0. \tag{28}$$

The derivatives with respect to the two variables σ^2, σ^3 are

$$\frac{\partial f}{\partial \sigma^2} = \frac{3(\sigma^3)^2}{(2\sigma^2 + 3\sigma^3)^2}, \qquad \frac{\partial f}{\partial \sigma^2} = 2\Big(\frac{\sigma^2}{2\sigma^2 + 3\sigma^3}\Big)$$

so that the gradient

$$\nabla f = \frac{1}{(2\sigma^2 + 3\sigma^3)^2}[3(\sigma^3)^2, 2(\sigma^2)^2]'$$

is non-decreasing in both σ^2 and σ^3. $\nabla f = 0$ when $\sigma^2 = \sigma^3 = 0$ where f is not defined, so that maxima have to be found on the border of the triangular

domain of f.
For $\sigma^2 = 0$, $0 \leq \sigma^3 \leq 1$ we have that $f = 0$, and the same is true for the other segment $\sigma^3 = 0$, $0 \leq \sigma^2 \leq 1$ (see Figure 7-left). This is obviously consistent with Proposition 3 which states that, if b gives nonzero mass to events of a fixed size only, then $p[b] = BetP[b]$.
In other words, Figure 7-left is exactly the extension of Proposition 3 we were looking for.
It follows that maxima have to belong to the segment $\sigma^2 + \sigma^3 = 1$, where $\sigma^3 = 1 - \sigma^2$: Let us then study the function

$$f'(\sigma^2) = \frac{\sigma^2(1-\sigma^2)}{2\sigma^2 + 3(1-\sigma^2)} = \sigma^2 \frac{1-\sigma^2}{3-\sigma^2}.$$

By computing its derivative with respect to σ^2

$$\frac{\partial f'}{\partial \sigma^2} = 1 - \frac{6}{(\sigma^2 - 3)^2} = 0$$

we get that $f'(\sigma^2)$ has a maximum in $\sigma^2 = 3 - \sqrt{6}$ (see Figure 7-right). In

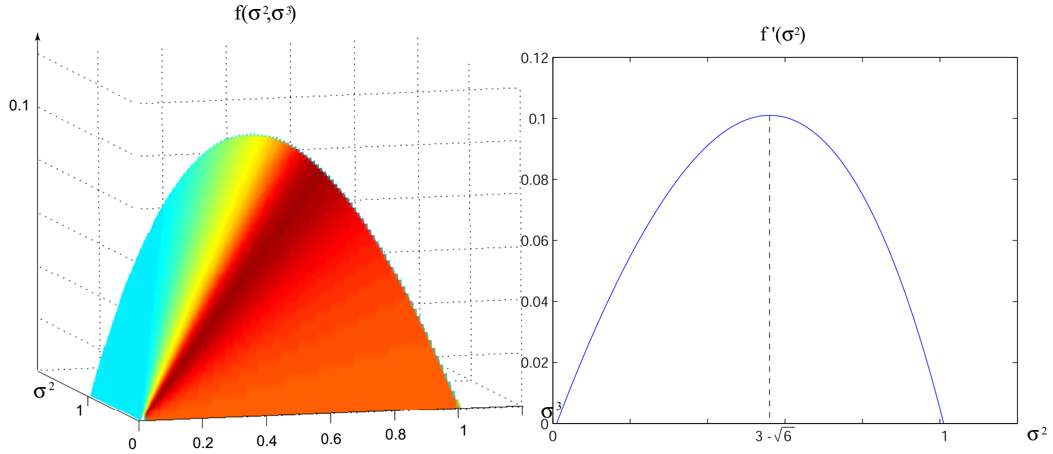

Fig. 7. Shape of the factor (28) in the norm $\|p[b] - \pi[b]\|_p$ for a belief function b in the ternary frame. Left: plot of the function $f(\sigma^2, \sigma^3)$ on the domain $\sigma^2 + \sigma^3 \leq 1$, $\sigma^2, \sigma^3 \geq 0$. Right: plot of its section $f'(\sigma^2)$.

conclusion, the factor (28) has a maximum in $\sigma^2 = 3 - \sqrt{6}$, $\sigma^3 = \sqrt{6} - 2$, so that by combining this result with Equation (27) we can claim that

Theorem 5 $\|p[b] - BetP[b]\|_p$ *in the ternary belief space* $\mathcal{B}_3$ *has three maxima, corresponding to the three belief functions with b.p.a.*

$$m_{b_1} = [0, 0, 0, 3 - \sqrt{6}, 0, 0, \sqrt{6} - 2]',$$
$$m_{b_2} = [0, 0, 0, 0, 3 - \sqrt{6}, 0, \sqrt{6} - 2]',$$
$$m_{b_3} = [0, 0, 0, 0, 0, 3 - \sqrt{6}, \sqrt{6} - 2]'$$

no matter the norm $p = 1, 2, \infty$ chosen.

The results of this case study can be seen as a first step towards a more complete quantitative analysis of the differences between the intersection probability and the other Bayesian approximations of the same family. Another sensible direction to follow, in this pursuit, is to look for less stringent conditions for their equivalence, by relaxing the hypotheses of Propositions 2 and 3. This is the aim of the rest of this Section.

6.2 Equidistribution as sufficient condition for $p = BetP$

The analysis of the ternary case, in which pignistic function and orthogonal projection coincide, invites us to assess the possibility that equidistribution

$$m_b(A) = const \ \forall A : |A| = k, \ \forall k = 2, ..., n. \tag{29}$$

is a sufficient condition for the equality of intersection and pignistic probability. Indeed, this turns out to be the case.

Theorem 6 *If a belief function b meets condition (29) (mass is equally distributed among focal elements of the same size) then the related pignistic and intersection probabilities coincide.*

Proof. If b meets (29), then the expression (12) for the probability values of the intersection probability gives, for each $x \in \Theta$,

$$p[b](x) = m_b(x) + \beta[b] \sum_{A \supsetneq x} m_b(A) = m_b(x) + \beta[b] \sum_{k=2}^{n} \sigma^k \frac{\binom{n-1}{k-1}}{\binom{n}{k}} =$$

(as there are $\binom{n-1}{k-1}$ events of size k containing x, and $\binom{n}{k}$ events of size k)

$$= m_b(x) + \beta[b] \sum_{k=2}^{n} \sigma^k \frac{k}{n} = m_b(x) + \frac{1}{n} \frac{\sigma^2 + ... + \sigma^n}{2\sigma^2 + ... + n\sigma^n} (2\sigma^2 + ... + n\sigma^n) =$$
$$= m_b(x) + \frac{1}{n}(\sigma^2 + ... + \sigma^n)$$

after recalling the decomposition (23) of $\beta[b]$.
On the other side, under the hypothesis the pignistic function reads as

$$\begin{aligned} BetP[b](x) = m_b(x) + \sum_{k=2}^{n} \sum_{A \supsetneq x, |A| = k} \frac{m_b(A)}{k} = \\ = m_b(x) + \sum_{k=2}^{n} \frac{\sigma^k}{k} \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = m_b(x) + \sum_{k=2}^{n} \frac{\sigma^k}{k} \frac{k}{k} = m_b(x) + \sum_{k=2}^{n} \frac{\sigma^k}{n} \end{aligned} \tag{30}$$

and the two functions coincide. □

6.3 *Equidistribution as sufficient condition for* $p[b] = \pi[b]$

In the same way, condition (29) is sufficient to guarantee the equality of intersection probability and orthogonal projection.

Theorem 7 *If a belief function b meets condition (29) (mass is equally distributed among focal elements of the same size) then the related orthogonal projection and intersection probability coincide.*

Proof. The orthogonal projection of a belief function b on the probability simplex $\mathcal{P}$ has the following expression [9]:

$$\pi[b](x) = \sum_{A \supseteq x} m_b(A)\Big(\frac{1+|A^c|2^{1-|A|}}{n}\Big) + \sum_{A \not\supseteq x} m_b(A)\Big(\frac{1-|A|2^{1-|A|}}{n}\Big). \quad (31)$$

Under condition (29) it becomes

$$\begin{array}{c} \pi[b](x) = m_b(x) + \displaystyle\sum_{k=2}^{n}\left(\frac{1+(n-k)2^{1-k}}{n}\right) \sum_{A\supseteq x, |A|=k} m_b(A) + \\ + \displaystyle\sum_{k=2}^{n}\left(\frac{1-(n-k)2^{1-k}}{n}\right) \sum_{A\not\supseteq x, |A|=k} m_b(A) \end{array} \quad (32)$$

where again $\sum_{A\supseteq x, |A|=k} m_b(A) = \sigma^k k/n$, while

$$\sum_{A\not\supseteq x, |A|=k} m_b(A) = \sigma^k \frac{\binom{n-1}{k}}{\binom{n}{k}} = \sigma^k \frac{(n-1)!}{k!(n-k-1)!} \frac{k!(n-k)!}{n!} = \sigma^k \frac{n-k}{n}.$$

Replacing those expressions in Equation (32) yields

$$\begin{array}{l} m_b(x) + \displaystyle\sum_{k=2}^{n}\left(\frac{1+(n-k)2^{1-k}}{n}\right)\sigma^k \frac{k}{n} + \sum_{k=2}^{n}\left(\frac{1-(n-k)2^{1-k}}{n}\right)\sigma^k \frac{n-k}{n} = \\ \quad = m_b(x) + \displaystyle\sum_{k=2}^{n}\left(\sigma^k \frac{k}{n^2} + \sigma^k \frac{n-k}{n^2}\right) = m_b(x) + \frac{1}{n}\sum_{k=2}^{n}\sigma^k \end{array}$$

which is exactly the value (30) of the intersection probability under the same assumption. □

7 Conclusions

In this paper we provided an interpretation in terms of degrees of belief of a Bayesian approximation of belief functions derived from purely geometric considerations, the intersection probability. Also, after noticing that Bayesian approximations can be classified in two groups according to the operator they commute with, we analyzed its behavior with respect to Dempster's sum and convex combination (of b.f. as points of the belief space). It turns out that $p[b]$, even though does not always commute with Cl can be assimilated to pignistic function and orthogonal projection. The conditions under which commutativity holds are given in terms of non-Bayesian contributions and degrees of belief. Finally, we investigated the case study of a ternary frame as a first step towards a fully quantitative assessment of the distance between different approximations, in particular between $p[b]$ and the couple $BetP, \pi$, showing that results are consistent no matter the particular norm chosen.
The latter, in particular, is promising direction of research which is worth pursuing. In the near future we plan to present an extensive and exhaustive discussion of the quantitative distance between all Bayesian approximations in a coherent picture.

Appendix

Proof of Theorem 1

Applying Equation (17) to $\varsigma \oplus p$ yields $\varsigma \oplus p =$

$$= \Big[\beta[b]pl_b + (1-\beta[b])b\Big] \oplus p = \frac{\beta[b]k(p, pl_b)pl_b \oplus p + (1-\beta[b])k(p,b)b \oplus p}{\beta[b]k(p, pl_b) + (1-\beta[b])k(p,b)} \quad (33)$$

where

$$k(p, pl_b) = \sum_{x \in \Theta} p(x)\Big(\sum_{A \supset x} \mu_b(A)\Big) = \sum_{x \in \Theta} p(x) m_b(x),$$
$$k(p, b) = \sum_{x \in \Theta} p(x)\Big(\sum_{A \supset x} m_b(A)\Big) = \sum_{x \in \Theta} p(x) pl_b(x).$$

When we apply (17) to $p[b] \oplus p$, instead, we get (recalling Equation (19)) $p[b] \oplus p =$

$$= \Big[\beta[b]\bar{pl}_b + (1-\beta[b])\bar{b}\Big] \oplus p = \frac{\beta[b]k(p, \bar{pl}_b)\bar{pl}_b \oplus p + (1-\beta[b])k(p,\bar{b})\bar{b} \oplus p}{\beta[b]k(p, \bar{pl}_b) + (1-\beta[b])k(p,\bar{b})}. \quad (34)$$

But now, by definition of Dempster's combination (3),

$$\bar{pl}_b \oplus p = \frac{\sum_{x\in\Theta} b_x p(x)(pl_b(x) + 1 - k_{\tilde{pl}_b})}{\sum_{x\in\Theta} p(x) pl_b(x) + 1 - k_{\tilde{pl}_b}}, \quad \bar{b} \oplus p = \frac{\sum_{x\in\Theta} b_x p(x)(m_b(x) + 1 - k_{\tilde{b}})}{\sum_{x\in\Theta} p(x) m_b(x) + 1 - k_{\tilde{b}}}$$

so that

$$\begin{aligned}
k(p,\bar{pl}_b)\bar{pl}_b \oplus p &= \sum_{x\in\Theta} b_x p(x) pl_b(x) + (1 - k_{\tilde{pl}_b}) \sum_{x\in\Theta} b_x p(x) = \\
&= k(p,\tilde{pl}_b) p \oplus \tilde{pl}_b + (1 - k_{\tilde{pl}_b})p = k(b,p) b \oplus p + (1 - k_{\tilde{pl}_b})p \\
k(p,\bar{b})\bar{b} \oplus p &= \sum_{x\in\Theta} b_x p(x) m_b(x) + (1 - k_{\tilde{b}}) \sum_{x\in\Theta} b_x p(x) = \\
&= k(p,\tilde{b}) p \oplus \tilde{b} + (1 - k_{\tilde{b}})p = k(pl_b, p) pl_b \oplus p + (1 - k_{\tilde{b}})p.
\end{aligned}$$

as from [30] $b \oplus p = \tilde{pl}_b \oplus p$, while [11] $pl_b \oplus p = \tilde{b} \oplus p$. After replacing these expressions in the numerator of Equation (34) we can notice that, as

$$\beta[b] = \frac{1 - k_{\tilde{b}}}{k_{\tilde{pl}_b} - k_{\tilde{b}}}, \quad 1 - \beta[b] = \frac{k_{\tilde{pl}_b} - 1}{k_{\tilde{pl}_b} - k_{\tilde{b}}},$$

the contributions of p vanish, leaving expression (33) for $\varsigma \oplus p$.

Proof of Theorem 2

$p[\alpha_1 b_1 + \alpha_2 b_2](x)$ can be written, by definition, as

$$p[\alpha_1 b_1 + \alpha_2 b_2](x) = m_{\alpha_1 b_1 + \alpha_2 b_2}(x) + \beta[\alpha_1 b_1 + \alpha_2 b_2] \sum_{A \supsetneq x} m_{\alpha_1 b_1 + \alpha_2 b_2}(A). \quad (35)$$

But now $\beta[\alpha_1 b_1 + \alpha_2 b_2] =$

$$\begin{aligned}
&\frac{\sum_{|A|>1} m_{\alpha_1 b_1 + \alpha_2 b_2}(A)}{\sum_{|A|>1} m_{\alpha_1 b_1 + \alpha_2 b_2}(A)|A|} = \frac{\alpha_1 \sum_{|A|>1} m_{b_1}(A) + \alpha_2 \sum_{|A|>1} m_{b_2}(A)}{\alpha_1 \sum_{|A|>1} m_{b_1}(A)|A| + \alpha_2 \sum_{|A|>1} m_{b_2}(A)|A|} \\
&= \frac{\alpha_1 N_1 + \alpha_2 N_2}{\alpha_1 D_1 + \alpha_2 D_2} = \frac{\alpha_1 D_1 \beta[b_1] + \alpha_2 D_2 \beta[b_2]}{\alpha_1 D_1 + \alpha_2 D_2} = \widehat{\alpha_1 D_1} \beta[b_1] + \widehat{\alpha_2 D_2} \beta[b_2]
\end{aligned}$$

once introduced the notation $\beta[b_i] = N_i/D_i$. Replacing this decomposition for $\beta[\alpha_1 b_1 + \alpha_2 b_2]$ into Equation (35) yields

$$\alpha_1 m_{b_1}(x) + \alpha_2 m_{b_2}(x) + (\widehat{\alpha_1 D_1}\beta[b_1] + \widehat{\alpha_2 D_2}\beta[b_2])\Big(\alpha_1 \sum_{A \supsetneq x} m_{b_1}(A) + \alpha_2 \cdot$$

$$\cdot \sum_{A \supsetneq x} m_{b_2}(A)\Big) = \frac{1}{\alpha_1 D_1 + \alpha_2 D_2}\Big\{(\alpha_1 D_1 + \alpha_2 D_2)(\alpha_1 m_{b_1}(x) + \alpha_2 m_{b_2}(x)) +$$

$$+ (\alpha_1 D_1 \beta[b_1] + \alpha_2 D_2 \beta[b_2])\Big(\alpha_1 \sum_{A \supsetneq x} m_{b_1}(A) + \alpha_2 \sum_{A \supsetneq x} m_{b_2}(A)\Big)\Big\} =$$

$$= \frac{\alpha_1 D_1}{\alpha_1 D_1 + \alpha_2 D_2}\Big[\alpha_1 m_{b_1}(x) + \alpha_2 m_{b_2}(x) + \beta[b_1]\Big(\alpha_1 \sum_{A \supsetneq x} m_{b_1}(A) +$$

$$+ \alpha_2 \sum_{A \supsetneq x} m_{b_2}(A)\Big)\Big] + \frac{\alpha_2 D_2}{\alpha_1 D_1 + \alpha_2 D_2}\Big[\alpha_1 m_{b_1}(x) + \alpha_2 m_{b_2}(x) + \beta[b_2] \cdot$$

$$\cdot \Big(\alpha_1 \sum_{A \supsetneq x} m_{b_1}(A) + \alpha_2 \sum_{A \supsetneq x} m_{b_2}(A)\Big)\Big] = \frac{\alpha_1 D_1}{\alpha_1 D_1 + \alpha_2 D_2} \cdot$$

$$\cdot \Big[\alpha_1\Big(m_{b_1}(x) + \beta[b_1] \sum_{A \supsetneq x} m_{b_1}(A)\Big) + \alpha_2\Big(m_{b_2}(x) + \beta[b_1] \sum_{A \supsetneq x} m_{b_2}(A)\Big)\Big] +$$

$$+ \frac{\alpha_2 D_2}{\alpha_1 D_1 + \alpha_2 D_2}\Big[\alpha_1\Big(m_{b_1}(x) + \beta[b_2] \sum_{A \supsetneq x} m_{b_1}(A)\Big) +$$

$$+ \alpha_2\Big(m_{b_2}(x) + \beta[b_2] \sum_{A \supsetneq x} m_{b_2}(A)\Big)\Big] = \frac{\alpha_1^2 D_1}{\alpha_1 D_1 + \alpha_2 D_2} p[b_1] +$$

$$+ \frac{\alpha_2^2 D_2}{\alpha_1 D_1 + \alpha_2 D_2} p[b_2] + \frac{\alpha_1 \alpha_2}{\alpha_1 D_1 + \alpha_2 D_2}\Big[D_1 p[b_2, b_1](x) + D_2 p[b_1, b_2](x)\Big].$$

after recalling Equation (22). We can then notice that the function

$$F(x) \doteq D_1 p[b_2, b_1](x) + D_2 p[b_1, b_2](x)$$

is such that $\sum_{x \in \Theta} F(x) =$

$$= \sum_{x \in \Theta}\Big[D_1 m_{b_2}(x) + N_1(pl_{b_2}(x) - m_{b_2}(x)) + D_2 m_{b_1}(x) + N_2(pl_{b_1}(x) +$$

$$- m_{b_1}(x))\Big] = D_1(1 - N_2) + N_1 D_2 + D_2(1 - N_1) + N_2 D_1 = D_1 + D_2$$

(recalling Equation (11)), so that $T[b_1, b_2](x) = F(x)/(D_1 + D_2)$ is a probability (as $T[b_1, b_2](x)$ is always non negative), expressed by Equation (21). $p[\alpha_1 b_1 + \alpha_2 b_2](x)$ then reduces to

$$p[\alpha_1 b_1 + \alpha_2 b_2](x) = \frac{\alpha_1 D_1}{\alpha_1 D_1 + \alpha_2 D_2}\alpha_1 p[b_1](x) +$$

$$+ \frac{\alpha_2 D_2}{\alpha_1 D_1 + \alpha_2 D_2}\alpha_2 p[b_2](x) + \frac{\alpha_1 \alpha_2}{\alpha_1 D_1 + \alpha_2 D_2}(D_1 + D_2) T[b_1, b_2](x)$$

i.e. Equation (20).

Proof of Theorem 3

By Equation (20) we have that

$$\begin{gathered} p[\alpha_1 b_1 + \alpha_2 b_2] - \alpha_1 p[b_1] - \alpha_2 p[b_2] = \widehat{\alpha_1 D_1} \alpha_1 p[b_1] + \widehat{\alpha_1 D_1} \alpha_2 T + \\ + \widehat{\alpha_2 D_2} \alpha_1 T + \widehat{\alpha_2 D_2} \alpha_2 p[b_2] - \alpha_1 p[b_1] - \alpha_2 p[b_2] = \alpha_1 p[b_1](\widehat{\alpha_1 D_1} - 1) + \\ + \widehat{\alpha_1 D_1} \alpha_2 T + \widehat{\alpha_2 D_2} \alpha_1 T + \alpha_2 p[b_2](\widehat{\alpha_2 D_2} - 1) = -\alpha_1 p[b_1] \widehat{\alpha_2 D_2} + \\ + \widehat{\alpha_1 D_1} \alpha_2 T + \widehat{\alpha_2 D_2} \alpha_1 T - \alpha_2 p[b_2] \widehat{\alpha_1 D_1} = \widehat{\alpha_1 D_1}(\alpha_2 T - \alpha_2 p[b_2]) + \\ + \widehat{\alpha_2 D_2}(\alpha_1 T - \alpha_1 p[b_1]) = \frac{\alpha_1 \alpha_2}{\alpha_1 D_1 + \alpha_2 D_2} [D_1(T - p[b_2]) + D_2(T - p[b_1])] \end{gathered}$$

which is nil iff

$$T[b_1, b_2](D_1 + D_2) = p[b_1] D_2 + p[b_2] D_1 \equiv T[b_1, b_2] = \hat{D}_1 p[b_2] + \hat{D}_2 p[b_1]$$

as $\frac{\alpha_1 \alpha_2}{\alpha_1 D_1 + \alpha_2 D_2}$ is always non-zero in non-trivial cases. This is equivalent to (after replacing the expressions for $p[b]$ (16) and $T[b_1, b_2]$ (21))

$$D_1(pl_{b_2}(x) - m_{b_2}(x))(\beta[b_2] - \beta[b_1]) + D_2(pl_{b_1}(x) - m_{b_1}(x))(\beta[b_1] - \beta[b_2]) = 0$$

in turn equivalent to

$$(\beta[b_2] - \beta[b_1]) \Big[D_1(pl_{b_2}(x) - m_{b_2}(x)) - D_2(pl_{b_1}(x) - m_{b_1}(x)) \Big] = 0.$$

Obviously this is true iff $\beta[b_1] = \beta[b_2]$ *or* the second factor is zero, i.e.

$$\begin{gathered} D_1 D_2 \frac{pl_{b_2}(x) - m_{b_2}(x)}{D_2} - D_1 D_2 \frac{pl_{b_1}(x) - m_{b_1}(x)}{D_1} = \\ = D_1 D_2 (R[b_2](x) - R[b_1](x)) = 0 \end{gathered}$$

$= 0$ for all $x \in \Theta$, i.e. $R[b_1] = R[b_2]$.

References

[1] Martin Aigner, *Combinatorial theory*, Classics in Mathematics, Springer, New York, 1979.

[2] M. Bauer, *Approximations for decision making in the Dempster-Shafer theory of evidence*, Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence (F. Horvitz, E.; Jensen, ed.), Portland, 1-4 August 1996, pp. 73–80.

[3] Mathias Bauer, *Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study*, International Journal of Approximate Reasoning **17** (1997), 217–237.

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

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

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

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

[8] F. Cuzzolin, *A geometric approach to the theory of evidence*, to appear on the IEEE Transactions on Systems, Man and Cybernetics part C (2007).

[9] ———, *Two new Bayesian approximations of belief functions based on convex geometry*, IEEE Transactions on Systems, Man, and Cybernetics - Part B, to appear (2007).

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

[11] ———, *Dual properties of the relative belief of singletons*, submitted to the IEEE Transactions on Fuzzy Systems (February 2007).

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

[13] Fabio Cuzzolin and Ruggero Frezza, *Geometric analysis of belief space and conditional subspaces*, Proceedings of the 2^{nd} International Symposium on Imprecise Probabilities and their Applications (ISIPTA2001), 26-29 June 2001.

[14] A. P. Dempster, *Upper and lower probability inferences based on a sample from a finite univariate population*, Biometrika **54** (1967), 515–528.

[15] A.P. Dempster, *Upper and lower probabilities generated by a random closed interval*, Annals of Mathematical Statistics **39** (1968), 957–966.

[16] ———, *Upper and lower probabilities inferences for families of hypothesis with monotone density ratios*, Annals of Mathematical Statistics **40** (1969), 953–969.

[17] T. Denoeux, *Inner and outer approximation of belief structures using a hierarchical clustering approach*, Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **9(4)** (2001), 437–460.

[18] T. Denoeux and A. Ben Yaghlane, *Approximating the combination of belief functions using the Fast Moebius Transform in a coarsened frame*, International Journal of Approximate Reasoning **31(1-2)** (October 2002), 77–101.

[19] D. Dubois and H. Prade, *Consonant approximations of belief functions*, International Journal of Approximate Reasoning **4** (1990), 419–449.

[20] Didier Dubois, Henri Prade, and Philippe Smets, *New semantics for quantitative possibility theory.*, ISIPTA, 2001, pp. 152–161.

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

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

[23] Glenn Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

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

[25] Philippe Smets, *Belief functions versus probability functions*, Uncertainty and Intelligent Systems (Saitta L. Bouchon B. and Yager R., eds.), Springer Verlag, Berlin, 1988, pp. 17–24.

[26] ———, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in Artificial Intelligence, 5 (M. Henrion, R.D. Shachter, L.N. Kanal, and J.F. Lemmer, eds.), Elsevier Science, 1990, pp. 29–39.

[27] ———, *The canonical decomposition of a weighted belief*, Proceedings of the International Joint Conference on AI, IJCAI95, Montréal, Canada, 1995, pp. 1896–1901.

[28] Philippe Smets and Robert Kennes, *The transferable belief model*, Artificial Intelligence **66** (1994), 191–234.

[29] Bjornar Tessem, *Approximations for efficient computation in the theory of evidence*, Artificial Intelligence **61:2** (1993), 315–329.

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

[31] A. Ben Yaghlane, T. Denoeux, and K. Mellouli, *Coarsening approximations of belief functions*, Proceedings of ECSQARU'2001 (S. Benferhat and P. Besnard, eds.), 2001, pp. 362–373.