The intersection probability and its properties

Fabio Cuzzolin
fabio.cuzzolin@brookes.ac.uk

Oxford Brookes University, Oxford, UK

Abstract. In this paper we discuss the properties of the intersection probability, a recent Bayesian approximation of belief functions introduced by geometric means. We propose a rationale for this approximation valid for interval probabilities, study its geometry in the probability simplex with respect to the polytope of consistent probabilities, and discuss the way it relates to important operators acting on belief functions.

1 Introduction

In the *theory of evidence* [1, 2] the best representation of chance is a *belief function* (b.f.) rather than a Bayesian probability distribution, assigning probability values to *sets* of possibilities rather than single events. The relationship between belief functions and probabilities is of course of great interest in the theory of evidence: the issue is often known as "probability transformation" [3–6]. 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* [7]. Beliefs are there represented as convex sets of probabilities or "credal sets", while decisions are made after *pignistic transformation* [8]. On his side, in his 1989 paper [9] F. Voorbraak proposed to adopt the so-called *relative plausibility* function. This is the unique probability that, given a belief function b with plausibility pl_b, assigns to each singleton its normalized plausibility.

The transformation problem can be posed in a different setting too. Belief and probability functions on finite domains can be represented as points of a large enough Cartesian space [10]. For instance, a belief function $b : 2^\Theta \rightarrow [0, 1]$ is completely specified by its belief values $\{b(A), A \subset \Theta, A \neq \emptyset, \Theta\}$ and can be seen as a point of $\mathbb{R}^{N-2}$, $N = 2^{|\Theta|}$. We can then obtain different probability transformations by minimizing different distances between the original belief function and the set of all probabilities.
In particular, we introduced a new probability $p[b]$ related to a belief function b, which we called *intersection probability*, determined by the intersection of the line joining a b.f. b and the related pl.f. pl_b with the region of all Bayesian (pseudo) b.f. [11].
In this paper we show that the intersection probability can in fact be defined for any *interval probability* system, as the unique probability obtained by assigning to all the elements of the domain the same fraction of uncertainty (Section 2). As a belief function determines an interval probability system, $p[b]$ exists for

belief functions too (for which it was originally introduced). The intersection probability can then be compared with other classical transformations like pignistic function and relative plausibility. In particular, the pignistic function has a strong credal interpretation as the barycenter of the polytope of all probabilities consistent with b. We prove that the intersection probability also possesses a credal interpretation in the probability simplex, as the "focus" of the pair of simplices embodying the interval probability system (Section 3).

In Section 4 we compare $p[b]$ with probability transformations of both the "affine" and "epistemic" family. While pignistic transformation and orthogonal projection commute with affine combination of belief functions, this is true for the intersection probability if and only if the considered interval probabilities attribute the same "weight" to the uncertainty of each element.

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

2 The intersection probability and its rationale

Belief functions and probability intervals are different but related mathematical representations of the bodies of evidence we possess on a given decision or estimation problem Q. We assume that the possible answers to Q form a finite set $\Theta = \{x_1, ..., x_n\}$ called "frame of discernment".

Given a certain amount of evidence we are allowed to describe our belief on the outcome of Q in several possible ways: the classical option is to assume a probability distribution on Θ. However, as we may need to incorporate imprecise measurements and people's opinions in our knowledge state, or cope with missing or scarce information, a more sensible approach is to assume we have no access to the "correct" probability distribution but the available evidence provides us with some sort of constraint on this true distribution. Both interval probabilities and belief functions are mathematical descriptions of such a constraint. They hence define different *credal sets* or sets of probability distributions on Θ.

An "interval probability system" is a system of constraints on the probability values of a probability measure $p : \Theta \to [0, 1]$ on a finite domain Θ of the form

$$(l, u) \doteq \{l(x) \leq p(x) \leq u(x), \forall x \in \Theta\}. \quad (1)$$

The system (1) determines an entire set of probability measures whose values are constrained to belong to a closed interval.

There are clearly many ways of selecting a single measure in order to represent a probability interval. We can point out, however, that all the intervals $[l(x), u(x)]$, $x \in \Theta$ have the same importance in the definition of the interval probability. There is no reason for the different singletons x to be treated differently.

It is then reasonable to request that the desired probability, candidate to represent the interval (1), should behave homogeneously in each element x of the frame Θ. This translates into seeking a probability p such that

$$p(x) = l(x) + \alpha(u(x) - l(x))$$

homogeneously for all elements x of Θ, for some value of $\alpha \in [0,1]$ (see Figure 1). It is easy to see that there is indeed a *unique* solution to this problem. It suffices

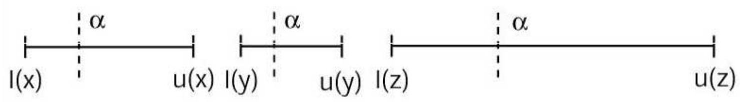

Fig. 1. An illustration of the notion of intersection probability for an upper/lower probability system.

to enforce the normalization constraint $\sum_x p(x) = \sum_x [l(x) + \alpha(u(x) - l(x))] = 1$ to understand that the unique solution is given by

$$\alpha = \beta[(l,u)] = \frac{1 - \sum_{x\in\Theta} l(x)}{\sum_{x\in\Theta}\big(u(x) - l(x)\big)}. \tag{2}$$

We can define the *intersection probability* associated with the interval probability system (1) as the probability measure with values

$$p[(l,u)](x) = \beta[(l,u)]u(x) + (1 - \beta[(l,u)])l(x). \tag{3}$$

The most interesting interpretation of $p[b]$ comes from its alternative form

$$p[(l,u)](x) = l(x) + \Big(1 - \sum_x l(x)\Big) R[(l,u)](x) \tag{4}$$

where

$$R[(l,u)](x) \doteq \frac{u(x) - l(x)}{\sum_y (u(y) - l(y))} = \frac{\Delta(x)}{\sum_y \Delta(y)}, \tag{5}$$

the quantity $\Delta(x)$ measuring the size of the probability interval on x. $R(x)$ measures how much the uncertainty on the probability value on a singletons "weights" on the total uncertainty represented by the interval probability (1). It is the natural to call it *relative uncertainty of singletons.*

Example. Consider a probability interval on a domain $\Theta = \{x, y, z\}$:

$$0.2 \leq p(x) \leq 0.8, \quad 0.4 \leq p(y) \leq 1, \quad 0 \leq p(x) \leq 0.4.$$

The widths of the corresponding intervals (the uncertainties on the probability values of the elements) are $\Delta(x) = 0.6$, $\Delta(y) = 0.6$, $\Delta(z) = 0.4$ respectively. By Equation (2) the fraction of uncertainty to add to $l(x)$ to get an admissible probability is

$$\beta = \frac{1 - 0.2 - 0.4 - 0}{0.6 + 0.6 + 0.4} = \frac{0.4}{1.6} = \frac{1}{4}.$$

The intersection probability has then values (4)

$$p[(l,u)](x) = 0.2 + \frac{1}{4}0.6 = 0.35, \qquad p[(l,u)](y) = 0.4 + \frac{1}{4}0.6 = 0.55,$$
$$p[(l,u)](z) = 0 + \frac{1}{4}0.4 = 0.1.$$

Intersection probability for belief measures. As a belief measure [1] also determines a probability interval, the intersection probability can be defined for belief functions too.
A "basic probability assignment" (b.p.a.) over a finite set or "frame of discernment" Θ is a function $m : 2^\Theta \to [0,1]$ on its power set $2^\Theta = \{A \subseteq \Theta\}$ such that 1. $m(\emptyset) = 0$; 2. $\sum_{A\subseteq\Theta} m(A) = 1$; 3. $m(A) \geq 0 \ \forall A \subseteq \Theta$. Subsets A of Θ associated with non-zero values $m(A) \neq 0$ of m are called "focal elements".
The *belief function* $b : 2^\Theta \to [0,1]$ associated with a basic probability assignment m on Θ is defined as:

$$b(A) = \sum_{B\subseteq A} m(B). \tag{6}$$

A finite probability or *Bayesian* belief function is just a special b.f. assigning non-zero masses to singletons only: $m_b(A) = 0$, $|A| > 1$.
A dual mathematical representation of the evidence encoded by a belief function b is the "plausibility function" (pl.f.) $pl_b : 2^\Theta \to [0,1]$, where $pl_b(A) \doteq 1 - b(A^c) = \sum_{B\cap A\neq\emptyset} m_b(B) \geq b(A)$ and A^c denotes the complement of A in Θ.
In the following we denote by b_A the unique "categorical" b.f. which assign unitary mass to a single event A: $m_b(A) = 1$, $m_b(B) = 0 \ \forall B \neq A$. We can then decompose each belief function b with b.p.a. $m_b(A)$ as

$$b = \sum_{A\subseteq\Theta} m_b(A) b_A. \tag{7}$$

A pair belief-plausibility determines then an interval probability system associated with a belief function, i.e.

$$(b, pl_b) \doteq \big\{p \in \mathcal{P} : b(x) \leq p(x) \leq pl_b(x), \forall x \in \Theta\big\}. \tag{8}$$

In this case the intersection probability can be written as

$$p[b](x) = \beta[b] pl_b(x) + (1 - \beta[b]) m_b(x) \tag{9}$$

with

$$\beta[b] = \frac{1 - \sum_{x\in\Theta} m_b(x)}{\sum_{x\in\Theta} \big(pl_b(x) - m_b(x)\big)} = \frac{1 - k_b}{k_{pl_b} - k_b} \tag{10}$$

where $k_{pl_b} \doteq \sum_{x\in\Theta} pl_b(x)$, $k_b \doteq \sum_{x\in\Theta} m_b(x)$ are the total plausibility and belief of singletons respectively.

3 Credal geometry in the probability simplex

Credal interpretation of belief functions and pignistic function. It is well known that a belief function determines an entire set of probabilities *consistent* with it, i.e. such that $b(A) \leq p(A) \leq pl_b(A)$ *for all events* $A \subseteq \Theta$.
Notice that this set [12, 13]:

$$\mathcal{P}[b] \doteq \{p \in \mathcal{P} : b(A) \leq p(A) \leq pl_b(A) \ \ \forall A \subseteq \Theta\} \tag{11}$$

is different from the set of probabilities (8) determined by the probability interval. A natural probabilistic approximation of b is then the center of mass of the set of consistent probabilities or "pignistic function" [8]

$$BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}. \tag{12}$$

$BetP[b]$ is the probability we obtain by assigning the mass of each focal element $A \subseteq \Theta$ of b *homogeneously* to each of its elements $x \in A$.
It is interesting to notice that the naive choice of choosing the barycenter of each interval $[l(x), u(x)]$ does not yield in general a valid probability function, for

$$\sum_x \big[l(x) + \frac{1}{2}(u(x) - l(x))\big] \neq 1.$$

This marks the difference with the case of belief functions, in which the barycenter of the set of probabilities defined by a belief function has a valid interpretation in terms of degrees of belief.

Credal interpretation of interval probabilities. However, a similar credal interpretation can be given for the intersection probability too, once we determine the credal set associated with an interval probability (1). Here we develop our argument in particular for the interval (8) determined by a belief function b. The polytope $\mathcal{P}[b]$ can be naturally decomposed as the intersection

$$\mathcal{P}[b] = \bigcap_{i=1}^{n-1} T^i[b] \tag{13}$$

of the regions $T^i[b] \doteq \big\{p \in \mathcal{P} : p(A) \geq b(A) \quad \forall A : |A| = i\big\}$ formed by all probability meeting the lower probability constraint *for size i events*. Let us consider in particular the set of probabilities which meet the lower constraint *on singletons* $T^1[b]$,

$$T^1[b] \doteq \big\{p \in \mathcal{P} : p(x) \geq b(x) \ \ \forall x \in \Theta\big\}.$$

It is also easy to see that

$$\begin{aligned} T^{n-1}[b] &\doteq \big\{p \in \mathcal{P} : p(A) \geq b(A) \ \ \forall A : |A| = n-1\big\} \\ &= \big\{p \in \mathcal{P} : p(\{x\}^c) \geq b(\{x\}^c) \ \ \forall x \in \Theta\big\} \\ &= \big\{p \in \mathcal{P} : p(x) \leq pl_b(x) \ \ \forall x \in \Theta\big\} \end{aligned}$$

expresses instead the *upper probability constraint on singletons*.
Clearly, then, the pair $(T^1[b], T^{n-1}[b])$ is the *geometric counterpart of an interval probability* in the probability simplex, exactly as the polytope of consistent probabilities $\mathcal{P}[b]$ represents there a belief function.

They form a higher dimensional triangle or *simplex*, i.e. the convex closure

$$Cl(\mathbf{v}_1, ..., \mathbf{v}_k) = \left\{ \mathbf{v} \in \mathbb{R}^d : \mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \sum_i \alpha_i = 1, \alpha_i \geq 0 \ \forall i \right\} \quad (14)$$

of a collection $\mathbf{v}_1, ..., \mathbf{v}_k$ of *affinely independent* points. The points $\mathbf{v}_1, ..., \mathbf{v}_k$ are said to be affinely independent if none of them can be expressed as an affine combination of the others: $\nexists \mathbf{v}_i, \ \{\alpha_j, j \neq i : \sum_{j \neq i} \alpha_j = 1\}$ such that $\mathbf{v}_i = \sum_{j \neq i} \alpha_j \mathbf{v}_j$. A triangle in the plane is a simplex, while a square is not.

We then call $T^1[b]$ and $T^{n-1}[b]$ *lower* and *upper simplices* respectively. They have very simple expressions in terms of the basic probability assignment of b. Using the notation of Equation (7) it can be proven that [14]:

Proposition 1. *The set $T^1[b]$ of all probabilities meeting the lower probability constraint on singletons is the simplex $T^1[b] = Cl(t^1_x[b], x \in \Theta)$ with vertices*

$$t^1_x[b] = \sum_{y \neq x} m_b(y) b_y + \big(1 - \sum_{y \neq x} m_b(y)\big) b_x. \quad (15)$$

A dual proof can be provided for the set $T^{n-1}[b]$ of probabilities which meet the upper probability constraint on singletons [14]. We just need to replace belief with plausibility values on singletons.

Proposition 2. *$T^{n-1}[b] = Cl(t^{n-1}_x[b], x \in \Theta)$ is a simplex with vertices*

$$t^{n-1}_x[b] = \sum_{y \neq x} pl_b(y) b_y + \big(1 - \sum_{y \neq x} pl_b(y)\big) b_x. \quad (16)$$

Consider as an example the case of a belief function

$$\begin{array}{lll} m_b(x) = 0.2, & m_b(y) = 0.1, & m_b(z) = 0.3, \\ m_b(\{x, y\}) = 0.1, & m_b(\{y, z\}) = 0.2, & m_b(\Theta) = 0.1 \end{array} \quad (17)$$

defined on a ternary frame $\Theta = \{x, y, z\}$. Figure 2 illustrates the geometry of its consistent simplex $\mathcal{P}[b]$. We can notice that by Equation (13) $\mathcal{P}[b]$ (the polygon delimited by tiny squares) is in this case the intersection of two triangles (2-dimensional simplices) $T^1[b]$ and $T^2[b]$. The intersection probability

$$\begin{array}{l} p[b](x) = m_b(x) + \beta[b](m_b(\{x, y\}) + m_b(\Theta)) = .2 + \frac{.4}{1.5 - 0.4} 0.2 = .27; \\ p[b](y) = .1 + \frac{.4}{1.1} 0.4 = .245; \quad p[b](z) = .485, \end{array}$$

is the unique intersection of the lines joining the corresponding vertices of the upper $T^2[b]$ and lower $T^1[b]$ simplices.

Intersection probability as focus of upper and lower simplices. This fact, true in the general case, can be formalized by the notion of "focus" of a pair of simplex.

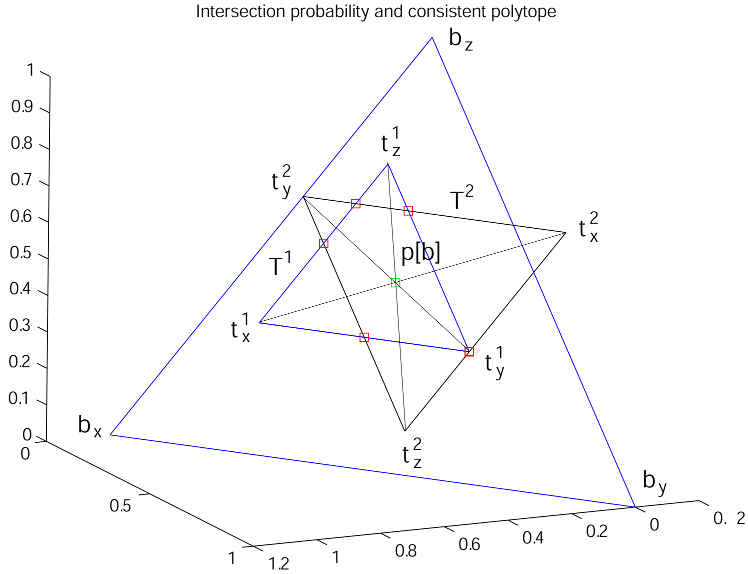

Fig. 2. The intersection probability is the focus of the two simplices $T^1[b]$ and $T^{n-1}[b]$. In the ternary case the latter reduce to the triangles $T^1[b]$ and $T^2[b]$. Their focus is geometrically the intersection of the lines joining their corresponding vertices.

Definition 1. *Consider a pair of simplices $S = Cl(s_1, ..., s_n)$, $T = Cl(t_1, ..., t_n)$. We call* focus *of the pair (S, T) the unique point $f(S, T)$ which has the same simplicial coordinates in both simplices:*

$$f = \sum_{i=1}^{n} \alpha_i s_i = \sum_{j=1}^{n} \beta_j t_j, \qquad \sum_{i=1}^{n} \alpha_i = \sum_{j=1}^{n} \beta_j = 1. \tag{18}$$

It is easy to see that such point always exists, even though it does not always fall in the intersection of the two simplices. In this case, though, the focus coincides with the unique intersection of the lines $a(s_i, t_i)$ joining corresponding vertices of S and T (see Figure 3): $f(S, T) = \bigcap_{i=1}^{n} a(s_i, t_i)$. Suppose indeed that a point p is such that $p = \alpha s_i + (1-\alpha) t_i \ \forall\ i = 1, ..., n$ (i.e. p lies on the line passing through s_i and t_i $\forall i$). Then necessarily $t_i = \frac{1}{1-\alpha}[p - \alpha s_i] \ \forall\ i = 1, ..., n$. If p has coordinates $\{\alpha_i, i = 1, ..., n\}$ in T, $p = \sum_{i=1}^{n} \alpha_i t_i$, then $p = \displaystyle\sum_{i=1}^{n} \alpha_i t_i = \frac{1}{1-\alpha}[p - \alpha \sum_i \alpha_i s_i]$ which implies $p = \sum_i \alpha_i s_i$, i.e. p is the focus of (S, T).

Theorem 1. *The intersection probability is the focus of the pair of upper and lower simplices $(T^{n-1}[b], T^1[b])$.*

Proof. We need to show that $p[b]$ has the same simplicial coordinates in $T^1[b]$ and $T^{n-1}[b]$. These coordinates turn out to be the values of the relative uncertainty

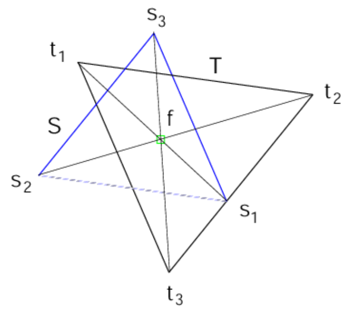

Fig. 3. If the focus of a pair of simplices belongs to their intersection, it is the unique intersection of the lines joining corresponding vertices of the two simplices.

function (5) for b:

$$R[b](x) = \frac{pl_b(x) - m_b(x)}{k_{pl_b} - k_b}. \tag{19}$$

Recalling the expression (15) of the vertices of $T^1[b]$, the point of the simplex $T^1[b]$ with coordinates (19) is

$$\sum_x R[b](x) t_x^1[b] = \sum_x R[b](x) \big[\sum_{y \neq x} m_b(y) b_y + \big(1 - \sum_{y \neq x} m_b(y)\big) b_x \big]$$
$$= \sum_x R[b](x) \big[\sum_{y \in \Theta} m_b(y) b_y + (1 - k_b) b_x \big]$$
$$= \sum_x b_x \big[(1 - k_b) R[b](x) + m_b(x) \sum_y R[b](y) \big] = \sum_x b_x \big[(1 - k_b) R[b](x) + m_b(x) \big]$$

as $R[b]$ is a probability ($\sum_y R[b](y) = 1$).
By Equation (4) the above quantity coincides with $p[b]$.
The point of $T^{n-1}[b]$ with the same coordinates $\{R[b](x), x \in \Theta\}$ is again

$$\sum_x R[b](x) t_x^{n-1}[b] = \sum_x R[b](x) \big[\sum_{y \neq x} pl_b(y) b_y + \big(1 - \sum_{y \neq x} pl_b(y)\big) b_x \big]$$
$$= \sum_x R[b](x) \big[\sum_{y \in \Theta} pl_b(y) b_y + (1 - k_{pl_b}) b_x \big] =$$
$$= \sum_x b_x \big[(1 - k_{pl_b}) R[b](x) + pl_b(x) \sum_y R[b](y) \big] =$$
$$= \sum_x b_x \big[(1 - k_{pl_b}) R[b](x) + pl_b(x) \big] = \sum_x b_x \big[pl_b(x) \frac{1 - k_b}{k_{pl_b} - k_b} - m_b(x) \frac{1 - k_b}{k_{pl_b} - k_b} \big]$$

$= p[b]$ by Equation (19). □

Pignistic function and intersection probability both adhere to rationality principles for belief functions and interval probabilities respectively. Geometrically, this translates into a similar behavior in the probability simplex, in which they are the center of mass of the consistent polytope and the focus of the pair of lower and upper probability simplices.

4 Intersection probability and other transformations

The epistemic family. An approach to the problem of approximating a b.f. with a probability seeks transformations which enjoy commutativity properties with respect to some combination rule [9, 15], in particular the original Dempster's sum [2]. Voorbraak's *relative plausibility of singletons* [9] (rel.plaus.) $\tilde{pl}_b$ is the unique probability that, given a belief function b with plausibility pl_b, assigns to each singleton its normalized plausibility

$$\tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)} = \frac{pl_b(x)}{k_{pl_b}} \tag{20}$$

and commutes with Dempster's orthogonal sum $\oplus$ [2, 16].
Dually, a *relative belief of singletons* [15] (rel.bel.) can be defined which assigns to the elements of Θ their normalized belief values:

$$\tilde{b}(x) \doteq \frac{b(x)}{\sum_{y\in\Theta} b(y)}. \tag{21}$$

Clearly $\tilde{b}$ exists iff b assigns some mass to singletons: $k_b = \sum_{x\in\Theta} m_b(x) \neq 0$.

These two approximations form a strongly linked couple. They are sometimes called the *epistemic* family of transformations [15].
It is important to notice, though, that in the interpretation of a belief function as a probability interval (8), the probabilities we obtain by normalizing the lower bound $\tilde{l}(x) = l(x)/\sum_y l(y)$ or the upper bound $\tilde{u}(x) = u(x)/\sum_y u(y)$ of the interval are *not* consistent with the interval itself. If there exists an $x \in \Theta$ such that $b(x) = pl_b(x)$ (the interval has width zero for that element) we have that

$$\tilde{b}(x) = \frac{m_b(x)}{\sum_y m_b(y)} > pl_b(x), \qquad \tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_y pl_b(y)} < b(x)$$

both the relative belief and plausibility of singletons fall outside the interval (8). This holds for a general probability interval (1), and supports the argument in favor of the interval probability.

The affine family. Many uncertainty measures can be represented as points of a Cartesian space [10]. In that context, affine combination is the geometric counterpart of the normalization constraint imposed on basic probability assignments. As a result, all significant entities form convex regions of the Cartesian space.
A different family of probability transformations commute indeed with affine combination of belief functions (as points of such a space). This is the case of pignistic function $BetP[b]$ and orthogonal projection $\pi[b]$ [11]. If $\alpha_1 + \alpha_2 = 1$

$$\begin{aligned} BetP[\alpha_1 b_1 + \alpha_2 b_2] &= \alpha_1 BetP[b_1] + \alpha_2 BetP[b_2] \\ \pi[\alpha_1 b_1 + \alpha_2 b_2] &= \alpha_1 \pi[b_1] + \alpha_2 \pi[b_2]. \end{aligned}$$

The condition under which the intersection probability commutes with convex combination is quite interesting.

Theorem 2. *$p[b]$ and affine combination commute, $p[\alpha_1 b_1 + \alpha_2 b_2] = \alpha_1 p[b_1] + \alpha_2 p[b_2]$ for $\alpha_1 + \alpha_2 = 1$, if and only if the relative uncertainty of the singletons is the same for both intervals: $R[b_1] = R[b_2]$.*

Proof. By definition (9) $p[\alpha_1 b_1 + \alpha_2 b_2] =$

$$= \alpha_1 m_1(x) + \alpha_2 m_2(x) + (1 - k_{\alpha_1 b_1 + \alpha_2 b_2}) \frac{\alpha_1 \Delta_1(x) + \alpha_2 \Delta_2(x)}{\sum_{y\in\Theta}(\alpha_1 \Delta_1(y) + \alpha_2 \Delta_2(y))}$$

that after defining $R(x) \doteq \frac{\alpha_1 \Delta_1(x) + \alpha_2 \Delta_2(x)}{\sum_{y\in\Theta}(\alpha_1 \Delta_1(y) + \alpha_2 \Delta_2(y))}$ becomes

$$\begin{aligned} &\alpha_1 m_1(x) + \alpha_2 m_2(x) + [1 - (\alpha_1 k_{b_1} + \alpha_2 k_{b_2})] R(x) = \\ &\qquad = \alpha_1 \big(m_1(x) + (1 - k_{b_1}) R(x)\big) + \alpha_2 \big(m_2(x) + (1 - k_{b_2}) R(x)\big) \end{aligned}$$

which is equal to $\alpha_1 p[b_1] + \alpha_2 p[b_2]$ iff

$$\alpha_1 (1 - k_{b_1})(R(x) - R[b_1](x)) + \alpha_2 (1 - k_{b_2})(R(x) - R[b_2](x)) = 0$$

which happens if and only if $R(x) = R[b_1](x) = R[b_2](x)$ $\forall x$, as $1 - k_{b_i} \neq 0$ unless b_i is a probability, and the thesis is trivially true for $\alpha_i = 1, \alpha_j = 0$. □

The intersection probability does not meet the nice relation with affine combination which characterizes pignistic function and orthogonal projection. However, Theorem 2 states that they commute exactly when each uncertainty interval $l(x) \leq p(x) \leq u(x)$ has the same "weight" in the two interval probabilities.

Comparison with the members of the affine family. It is natural to wonder what are the other differences between $p[b]$ and its "sister" functions $BetP[b]$ and $\pi[b]$. Some sufficient conditions [11] have been already worked out in the past. More stringent conditions can be formulated.

Theorem 3. *If a belief function b is such that its mass is equally distributed among focal elements of the same size*

$$m_b(A) = const \ \forall A : |A| = k, \forall k = 2, ..., n. \tag{22}$$

then its pignistic and intersection probabilities coincide: $BetP[b] = p[b]$.

Proof. If b meets (22), then the expression 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)

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

after recalling the decomposition of $\beta[b]$:

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

On the other side, under the hypothesis, the pignistic function reads as

$$\begin{aligned} BetP[b](x) &= m_b(x) + \sum_{k=2}^n \sum_{A\supseteq x, |A|=k} \frac{m_b(A)}{k} = m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{k} \frac{\binom{n-1}{k-1}}{\binom{n}{k}} \\ &= m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{k}\frac{k}{k} = m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{n} \end{aligned} \quad (24)$$

and the two functions coincide. □

Condition (22) is sufficient to guarantee the equality of intersection probability and orthogonal projection [11] too.

Theorem 4. *If a belief function b meets condition (22) (i.e., its 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 to the probability simplex $\mathcal{P}$ has the following expression [11]:

$$\pi[b](x) = \sum_{A\supseteq x} m_b(A)\Big(\frac{1 + |A^c|2^{1-|A|}}{n}\Big) + \sum_{A\not\supseteq x} m_b(A)\Big(\frac{1 - |A|2^{1-|A|}}{n}\Big). \quad (25)$$

Under condition (22) it becomes

$$\begin{aligned} \pi[b](x) = m_b(x) &+ \sum_{k=2}^n \Big(\frac{1 + (n-k)2^{1-k}}{n}\Big) \sum_{A\supseteq x, |A|=k} m_b(A) \\ &+ \sum_{k=2}^n \Big(\frac{1 - (n-k)2^{1-k}}{n}\Big) \sum_{A\not\supseteq x, |A|=k} m_b(A) \end{aligned} \quad (26)$$

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 (26) yields

$$\begin{aligned} &m_b(x) + \sum_{k=2}^n \Big(\frac{1 + (n-k)2^{1-k}}{n}\Big)\sigma^k \frac{k}{n} + \sum_{k=2}^n \Big(\frac{1 - (n-k)2^{1-k}}{n}\Big)\sigma^k \frac{n-k}{n} = \\ &= m_b(x) + \sum_{k=2}^n \Big(\sigma^k \frac{k}{n^2} + \sigma^k \frac{n-k}{n^2}\Big) = m_b(x) + \frac{1}{n}\sum_{k=2}^n \sigma^k \end{aligned}$$

which is exactly the value (24) of the intersection probability under the same assumption. □

5 Conclusions

In this paper we studied the intersection probability, a Bayesian transformation of belief functions originally derived from purely geometric arguments, from the more abstract point of view of interval probabilities, providing a rationality principle for it. We studied its behavior in the probability simplex, proving that it can be described as the focus of the upper and lower simplices which geometrically embody an interval probability. We compared it to transformations of both the affine and epistemic families, and studied the condition under which it commutes with convex combination.

References

1. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
2. Dempster, A.: Upper and lower probabilities generated by a random closed interval. Annals of Mathematical Statistics **39** (1968) 957–966
3. Denoeux, T.: 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
4. Denoeux, T., Yaghlane, A.B.: 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
5. Haenni, R., Lehmann, N.: Resource bounded and anytime approximation of belief function computations. IJAR **31**(1-2) (2002) 103–154
6. Bauer, M.: Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study. IJAR **17**(2-3) (1997) 217–237
7. Smets, P.: Belief functions versus probability functions. In Bouchon B., S.L., R., Y., eds.: Uncertainty and Intelligent Systems. Springer Verlag, Berlin (1988) 17–24
8. Smets, P.: Decision making in the TBM: the necessity of the pignistic transformation. IJAR **38**(2) (2005) 133–147
9. Voorbraak, F.: A computationally efficient approximation of Dempster-Shafer theory. International Journal on Man-Machine Studies **30** (1989) 525–536
10. Cuzzolin, F.: A geometric approach to the theory of evidence. IEEE Transactions on Systems, Man, and Cybernetics - Part C **38**(4) (2007) 522–534
11. Cuzzolin, F.: Two new Bayesian approximations of belief functions based on convex geometry. IEEE Trans. on Systems, Man, and Cybernetics B **37**(4) (2007) 993–1008
12. Chateauneuf, A., Jaffray, J.Y.: Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. Mathematical Social Sciences **17** (1989) 263–283
13. Dubois, D., Prade, H., Smets, P.: New semantics for quantitative possibility theory. In: ISIPTA. (2001) 152–161
14. Cuzzolin, F.: Credal semantics of Bayesian approximations in terms of probability intervals. submitted to the IEEE Trans. on Systems, Man, and Cybernetics B (2008)
15. Cuzzolin, F.: Semantics of the relative belief of singletons. In: International Workshop on Uncertainty and Logic UNCLOG'08, Kanazawa, Japan. (2008)
16. Cobb, B., Shenoy, P.: On the plausibility transformation method for translating belief function models to probability models. IJAR **41**(3) (2006) 314–330