Rationale and properties of the intersection probability

Fabio Cuzzolin [a]

[a] *Department of Computing*
Oxford Brookes University
Wheatley campus, Oxford OX33 1HX, United Kingdom

Abstract

In this paper we analyze the properties of the intersection probability, a recent Bayesian transformation of belief functions introduced by geometric means. We propose a rationale for this approximation valid for probability intervals, and study its geometry in the probability simplex with respect to the credal sets associated with such intervals. We discuss its relations with different evidence combination rules and the affine combination operator. We conclude by running an analytical comparison between the intersection probability and other probability transformations of the so-called "affine" family.

Key words: *Belief functions, probability intervals, Bayesian transformations, intersection probability, credal sets, affine combination, combination rules.*

1 Introduction

Belief functions [24,14] and probability intervals [31,13] are different but related mathematical representations of the bodies of evidence we possess on a given decision or estimation problem Q. The possible answers to Q are often assumed to belong to a finite set $\Theta = \{x_1, ..., x_n\}$, called "frame of discernment". Given a certain amount of evidence, our belief on the outcome of Q can indeed be described in several possible ways: the classical option is to assume a probability distribution on Θ. However, in practical situations, we might need to incorporate imprecise measurements and people's opinions in

Email address: `fabio.cuzzolin@brookes.ac.uk` (Fabio Cuzzolin).
URL: `http://perception.inrialpes.fr/people/Cuzzolin` (Fabio Cuzzolin).

our knowledge state, or cope with missing or scarce information. A more cautious approach is therefore to assume that we have no access to the "correct" probability distribution, but that the available evidence provides us with some sort of constraint on this unknown distribution.
The simplest such constraint is given by a *probability interval*: the probability values $p(x)$ of the elements of Θ are assumed to belong to an interval $l(x) \leq p(x) \leq u(x)$ delimited by a lower bound $l(x)$ and an upper bound $u(x)$. A belief function is quite a more complex measure, mathematically defined as a sum function $b : 2^\Theta \rightarrow [0, 1]$ on the power set $2^\Theta = \{A \subseteq \Theta\}$ of Θ. Both probability intervals and belief functions, though, determine *credal sets*[22] or convex sets of probability distributions on Θ.

For decision making it is normal practice in the theory of evidence to approximate the evidence encoded by a belief function with a single probability distribution, usually (but not necessarily) coming from the corresponding credal set. This is, for instance, the case of the Transferable Belief Model [25], in which the so called "pignistic transformation" [27] is employed for this purpose. Many other such transformations have been proposed, according to different criteria [23,32,34,35,16,17,20,1].
An interesting approach to the problem seeks approximations which enjoy commutativity properties with respect to a specific combination rule, in particular Dempster's sum [14,15]. This is the case of the *relative plausibility of singletons* [33], the unique probability that, given a belief function b with plausibility $pl_b(A) = 1 - b(A^c)$, assigns to each singleton its normalized plausibility. Its properties have been later analyzed by Cobb and Shenoy [3,4].
Similarly to the case of belief functions, it can be useful to apply such a transformation to reduce a set of probability intervals to a single probability distribution prior to actually making a decision. However, the problem has been quite neglected so far. One could of course pick a representative from the corresponding credal set, but it makes sense to wonder whether a transformation inherently designed for probability intervals as such could be found.

An interesting probability approximation called "intersection probability" has been recently detected by geometric means, in the context of the geometric approach to uncertainty [8]. As a belief function $b : 2^\Theta \rightarrow [0, 1]$ is completely specified by its belief values $\{b(A), \emptyset \subsetneq A \subsetneq \Theta\}$, it can be represented as a point of some Cartesian space [6,10]. In this framework, the *intersection probability* $p[b]$ is the unique probability distribution determined by the intersection of the line joining a belief function b and the related plausibility function pl_b with the region of Bayesian (pseudo) belief functions [8].
Even though originally introduced as a probabilistic or "Bayesian" approximation of belief functions, the intersection probability turns out to be (as we will see here) inherently associated with probability intervals, in which context its rationale clearly emerges.

1.1 Contributions and paper outline

In this paper we show that the intersection probability can in fact be defined for any interval probability system, as the unique probability distribution obtained by assigning the same fraction of the uncertainty interval to all the elements of the domain (Section 2). As a belief function determines itself an interval probability system, the intersection probability exists for belief functions too and can therefore be compared with classical approximations like pignistic function [25] and relative plausibility and belief [9] of singletons (Section 3). Just like they are not guaranteed to be consistent in the case of belief functions, relative upper and lower bounds produce incoherent results in the case of probability intervals too. The same holds for other transformations recently proposed [30,29].

As belief functions and probability intervals possess a strong credal interpretation, we move forward to study the behavior of the intersection probability in the probability simplex (Section 4). More specifically, each interval probability system is associated with two credal sets called "lower" and "upper" simplices. We prove that, just as the pignistic function is geometrically the barycenter of the polytope of all the probabilities "consistent" with a belief function, the intersection probability is the "focus" of the pair of upper and lower simplices embodying the set of probability intervals.

Finally, after recalling the original formulation of this transformation in the space of all belief functions (Section 5), we discuss its properties with respect to several important evidence combination rules like Dempster's rule and conjunctive combination, on one side, and affine combination on the other (Section 6).
In particular we provide a justification for its name by proving that, even though $p[b]$ is *not* the actual intersection between the line joining b and pl_b and the region of Bayesian pseudo belief functions, it does behave exactly like the actual intersection when combined with a probability distribution.
We also show that, while pignistic and orthogonal transformations commute with affine combination of belief functions, this is true for the intersection probability if the considered probability intervals attribute the same "weight" to the uncertainty of each element.
At the end of the paper an analytical comparison between intersection probability and other probability transformations of the "affine" family is run.

2 Notion of intersection probability

2.1 Probability intervals and the approximation problem

A *set of probability intervals* or *interval probability system* is a system of constraints on the probability values of a probability distribution $p : \Theta \rightarrow [0, 1]$ on a finite domain Θ of the form

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

Probability intervals have been introduced as a tool for uncertain reasoning in [13], where combination and marginalization of intervals were studied in detail. As pointed out for instance in [21], a typical way in which probability intervals arise is through measurement errors. As a matter of fact, measurements can be inherently of interval nature (due to the finite resolution of the instruments). In that case the *probability* interval of interest is the class of probability measures consistent with the *measured* interval.
A set of constraints of the form (1) obviously determines an entire set of probability distributions whose values are constrained to belong to a closed interval (see again [13]). A polytope or convex set of probability distributions is usually called a *credal set*. The credal sets generated by probability interval systems are just a sub-class of all possible polytopes of probability measures. Nevertheless, aggregation operators to combine two or more sets of probability intervals can be developed without resorting to their credal interpretation.

A similar situation holds for belief functions [24]. Originally defined as sum functions subject to normalization and non-negativity constraints, they also correspond to a specific class of credal sets. Furthermore, similarly to the case of probability intervals, belief functions are normally handled and aggregated by means of operators specifically designed for this class of uncertainty measures rather than by applying, say, Bayes' rule to the vertices of the corresponding credal sets.
However, for decision making it is normal practice in the theory of evidence to approximate a belief function with a single probability distribution, usually (but not necessarily) extracted from the corresponding credal set. This is for instance the case of the Transferable Belief Model, in which a so called "pignistic transformation" is employed. Many other such transformations have been proposed, reflecting a number of different criteria.

Similarly to the case of belief functions, it can be useful to apply such a transformation to reduce a probability interval system to a single probability distribution prior to actually making a decision. Such a problem has been rather neglected so far. One could of course pick a representative from the cor-

responding credal set, but it makes sense to wonder whether a transformation inherently designed for probability intervals as such could be found.
An interesting probability approximation called "intersection probability" has been recently brought forward by geometric means, in the context of the geometric approach to uncertainty [8]. Even though originally introduced as a Bayesian transformation of belief functions, the intersection probability turns out to be inherently associated with probability intervals, in which context its rationale clearly emerges.

2.2 Rationale

There are clearly many ways of selecting a single measure to represent a collection of probability intervals (1). It is important to point out, however, that each of the intervals $[l(x), u(x)]$, $x \in \Theta$, has the same importance in the definition of the system of constraints (1). There is no reason for the different elements x of the domain to be treated differently.
It is then reasonable to request that the desired probability, candidate to represent the interval system (1), should behave homogeneously in each element x of the frame Θ. Mathematically, this translates into seeking a probability distribution $p : \Theta \to [0, 1]$ such that

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

for all the elements x of Θ, and some constant value $\alpha \in [0, 1]$ (see Figure 1). Such value needs to be between 0 and 1 in order for the sought probability distribution p to belong to the interval. It is easy to see that there is indeed

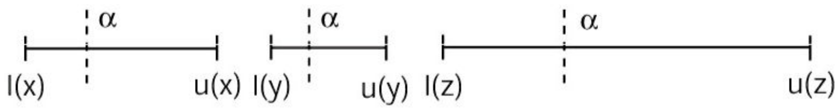

Fig. 1. An illustration of the notion of intersection probability for an interval probability system.

a *unique* solution to this problem. It suffices to enforce the normalization constraint

$$\sum_x p(x) = \sum_x \Big[l(x) + \alpha(u(x) - l(x)) \Big] = 1$$

to understand that the unique such value α is given by

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

We can define the *intersection probability* $p[(l,u)] : \Theta \to [0,1]$ associated with the interval probability system (1) as the probability measure:

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

2.3 Relative uncertainty on singletons

The ratio $\beta[(l,u)]$ (2) clearly measures the fraction of *each* probability interval which we need to add to the lower bound $l(x)$ to obtain a valid probability function (summing to one). The most interesting interpretation of $p[(l,u)]$ comes though 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 width of the probability interval for x.
The probability distribution $R[(l,u)] : \Theta \to [0,1]$ measures therefore how much the uncertainty on the probability value of each singleton "weighs" on the total width of the interval system (1). It is natural to call it *relative uncertainty* on singletons.
We can then say that $p[(l,u)]$ distributes the mass $(1 - \sum_x l(x))$ which is necessary to obtain a valid probability to each singleton $x \in \Theta$ according to the relative uncertainty $R[(l,u)](x)$ it carries in the given interval.

2.4 Example

Consider as an example an interval probability system on a domain $\Theta = \{x, y, z\}$ of size 3:

$$0.2 \leq p(x) \leq 0.8, \quad 0.4 \leq p(y) \leq 1, \quad 0.3 \leq p(z) \leq 0.3. \tag{6}$$

Notice that there is no uncertainty at all on the value of $p(z) = 0.3$. The widths of the corresponding intervals are $\Delta(x) = 0.6$, $\Delta(y) = 0.6$, $\Delta(z) = 0$ respectively. The relative uncertainty on each singleton (5) is therefore:

$$\begin{aligned} R[(l,u)](x) &= \tfrac{\Delta(x)}{\sum_{w\in\Theta}\Delta(w)} = \tfrac{0.6}{1.2} = \tfrac{1}{2}, \\ R[(l,u)](y) &= \tfrac{1}{2}, \\ R[(l,u)](z) &= \tfrac{\Delta(z)}{\sum_{w\in\Theta}\Delta(w)} = \tfrac{0}{1.2} = 0. \end{aligned} \tag{7}$$

Computing the intersection probability is then really easy. By Equation (2) the fraction of the uncertainty $u(x)-l(x)$ on $p(x)$ we need to add to the lower bound $l(x)$ to get an admissible, normalized probability is

$$\beta = \frac{1-0.2-0.4-0.3}{0.6+0.6} = \frac{0.1}{1.2} = \frac{1}{12}.$$

The intersection probability (4) has therefore values:

$$p[(l,u)](x) = 0.2 + \tfrac{1}{12}0.6 = 0.25, \quad p[(l,u)](y) = 0.4 + \tfrac{1}{12}0.6 = 0.45,$$
$$p[(l,u)](z) = 0.3 + \tfrac{1}{12}0 = 0.3.$$

Notice that the fact of having a zero-width interval for one of the singletons does not pose a problem for the intersection probability, which falls as expected inside the probability interval for all the elements of the domain.
According to its interpretation of Equation (4), $p[(l,u)]$ is also the result of distributing the necessary mass $(1-\sum_x l(x)) = 1-0.2-0.4-0.3 = 0.1$ to each singleton in proportion to the relative uncertainty $R[(l,u)]$ (Equation (7)) of their intervals:

$$p[(l,u)](x) = 0.2 + 0.1\tfrac{1}{2} = 0.25, \quad p[(l,u)](y) = 0.4 + 0.1\tfrac{1}{2} = 0.45,$$
$$p[(l,u)](z) = 0.3 + 0.1 \cdot 0 = 0.3.$$

3 Intersection probability for belief measures

As each belief measure determines itself a set of probability intervals, the intersection probability can be defined for belief functions too.

3.1 Belief functions

A "basic probability assignment" (b.p.a.) over a finite set or "frame of discernment" Θ is a function $m : 2^\Theta \rightarrow [0,1]$ 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 of Θ associated with non-zero values of m are called "focal elements".
The *belief function* $b : 2^\Theta \rightarrow [0,1]$ associated with a basic probability assignment m on Θ is defined as:

$$b(A) = \sum_{B\subseteq A} m(B). \tag{8}$$

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.) $pl_b : 2^\Theta \rightarrow [0,1]$, where

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

(A^c denotes the complement of A in Θ). The plausibility value $pl_b(A)$ accounts for the mass that might be assigned to some element of A, and represents the evidence not against A.

Belief functions have a natural interpretation as constraints on the "true", unknown probability distribution which better describes the state of belief on the outcome of Q. According to this interpretation, the mass assigned to each focal element $A \subseteq \Theta$ can float freely among its elements $x \in A$. A probability distribution *consistent* with b emerges by re-distributing the mass of each focal element to all its singletons. Such consistent probabilities form the following polytope in the probability simplex $\mathcal{P}$ [2,19]:

$$\mathcal{P}[b] \doteq \Big\{ p \in \mathcal{P} : b(A) \leq p(A) \leq pl_b(A) \ \ \forall A \subseteq \Theta \Big\}. \tag{9}$$

3.2 Intersection probability for belief functions

By considering only the constraints (9) acting on the probability values of singletons, a pair belief-plausibility determines a probability interval system associated with the pair itself, i.e.,

$$(b, pl_b) \doteq \Big\{ p \in \mathcal{P} : b(x) \leq p(x) \leq pl_b(x), \forall x \in \Theta \Big\}. \tag{10}$$

In this case the intersection probability $p[(l,u)]$ reads as

$$p[b](x) = \beta[b] pl_b(x) + (1 - \beta[b]) m_b(x) \tag{11}$$

with

$$\beta[b] = \frac{1 - \sum_{x \in \Theta} m_b(x)}{\sum_{x \in \Theta} \Big(pl_b(x) - m_b(x) \Big)} = \frac{1 - k_b}{k_{pl_b} - k_b} \tag{12}$$

where

$$k_{pl_b} \doteq \sum_{x \in \Theta} pl_b(x) \geq 1, \qquad k_b \doteq \sum_{x \in \Theta} m_b(x) \leq 1$$

are the total plausibility and belief of singletons, respectively.

Note that, unlike stated in [18], the intersection probability is neither inherently a sub-product of the geometry of belief functions, nor it is associated in any way with a specific evidence combination rule. It is a straightforward consequence of the fact that belief functions come with a collection of probability intervals naturally associated with them.

3.3 Intersection probability, barycenter, and pignistic function

A natural probabilistic approximation of a belief function b is the center of mass of the set (10) of consistent probabilities, or "pignistic function" [27]:

$$BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}. \tag{13}$$

The naive choice of picking the barycenter of each interval $[l(x), u(x)]$ to represent an interval probability system (l, u), though, 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, for which the pignistic function (13) has a strong interpretation. As a matter of fact, $BetP[b]$ is the probability we obtain by re-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 note, however, the parallelism between the rationality principles of pignistic function and intersection probability for belief functions and probability intervals, respectively. While $BetP[b]$ re-distributes the mass of each focal element equally to all its singletons, $p[(l, u)]$ re-distributes the same fraction of each probability interval equally to all the singletons. We will come back to this later in Section 4.

3.4 Relative plausibility and belief of singletons are potential inconsistent with a probability interval system

An approach to the problem of approximating a belief function with a probability seeks approximations which enjoy commutativity properties with respect to some evidence combination rule [33,11], in particular the original Dempster's sum [14].
Voorbraak's *relative plausibility of singletons* [33] $\tilde{pl}_b$ is the unique probability that, given a belief function b with plausibility pl_b, assigns to each singleton

its normalized plausibility:

$$\tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)} = \frac{pl_b(x)}{k_{pl_b}}. \tag{14}$$

Indeed, (14) commutes with Dempster's orthogonal sum $\oplus$ [14,4].
Dually, a *relative belief of singletons* [11] can be defined. It assigns to the elements of Θ their normalized belief values:

$$\tilde{b}(x) \doteq \frac{b(x)}{\sum_{y\in\Theta} b(y)}. \tag{15}$$

Clearly $\tilde{b}$ exists iff b assigns some mass to singletons: $k_b = \sum_{x\in\Theta} m_b(x) \neq 0$.
While $\tilde{pl}_b$ is associated with the less conservative (but incoherent) scenario in which all the mass that can be assigned to a singleton is actually assigned to it, $\tilde{b}$ reflects the most conservative (but still not coherent) choice of assigning to x only the mass that the b.f. b (seen as a constraint) assures it belong to x. It can be proven that the relative belief of singletons meets a number of dual properties with respect to Dempster's sum which are the dual of those enjoyed by the relative plausibility [9].

These two approximations form a strongly linked couple. It is important to notice, though, that for the probability interval system (10) determined by a belief function, the probabilities we obtain by normalizing lower $\tilde{l}(x) = l(x)/\sum_y l(y)$ or upper bound $\tilde{u}(x) = u(x)/\sum_y u(y)$ are *not* guaranteed to be consistent with the interval itself.
For instance, if there exists an element $x \in \Theta$ such that $b(x) = pl_b(x)$ (the interval has width zero for that element) we have that

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

Therefore, both relative belief and plausibility of singletons fall outside the interval system (10). This holds for a general collection of probability intervals (1), again marking the contrast with the behavior of the intersection probability.

3.5 Relation with Sudano's proposal transformations

Other proposals have been recently brought forward by Dezert et al. [18] and Sudano [30]. The latter, in particular, has proposed the following four probability transformations [1]

$$PrPl[b](x) \doteq \sum_{A \supseteq \{x\}} m_b(A) \frac{pl_b(x)}{\sum_{y \in A} pl_b(y)} \tag{16}$$

$$PrBel[b](x) \doteq \sum_{A \supseteq \{x\}} m_b(A) \frac{b(x)}{\sum_{y \in A} b(y)} = \sum_{A \supseteq \{x\}} m_b(A) \frac{m_b(x)}{\sum_{y \in A} m_b(y)} \tag{17}$$

$$PrNPl[b](x) \doteq \frac{1}{\Delta} \sum_{A \cap \{x\} \neq \emptyset} m_b(A) = \tilde{pl}_b(x) \tag{18}$$

$$PraPl[b](x) \doteq b(x) + \epsilon \cdot pl_b(x), \quad \epsilon = \frac{1 - \sum_{y \in \Theta} b(y)}{\sum_{y \in \Theta} pl_b(y)} = \frac{1 - k_b}{k_{pl_b}} \tag{19}$$

The first two transformations are clearly inspired by the pignistic function (13). While in the latter case the mass $m_b(A)$ of each focal element is redistributed homogenously to all its elements $x \in A$, $PrPl[b]$ (Equation (16)) redistributes $m_b(A)$ proportionally to the relative plausibility of singleton x *inside* A. Similarly, $PrBel[b]$ (Equation (17)) redistributes $m_b(A)$ proportionally to the relative *belief* of singleton x within A.
It is not clear to us the rationality principle behind these transformations. We can note, however, that such a redistribution process is the rationale of another probabilistic approximation, the *orthogonal projection* $\pi[b]$ of a belief function b on to the probability simplex [8]. Instead of redistributing the mass $m_b(A)$ of each focal element to its singletons, $\pi[b]$ homogenously redistributes its mass to all the *subsets* of A. The pignistic function $BetP[b]$, $PrPl[b]$, $PrBel[b]$, and the orthogonal projection $\pi[b]$ arguably form a family of approximations inspired by the same notion, that of mass redistribution.

As proven by Equation (18), $PrNPl[b]$ is nothing but the relative plausibility of singletons (14), and suffers from the same limitations.

The fourth transformation $PraPl[b]$ is more related to the case of probability intervals, and at the same time the most related to the intersection probability. By Equation (11),

$$p[b](x) = (1 - \beta[b])\tilde{b}(x)k_b + \beta[b]\tilde{pl}_b(x)k_{pl_b} \tag{20}$$

where
$$(1 - \beta[b])k_b + \beta[b]k_{pl_b} = \frac{k_{pl_b} - 1}{k_{pl_b} - k_b} k_b + \frac{1 - k_b}{k_{pl_b} - k_b} k_{pl_b} = 1$$
i.e. $p[b]$ lies on the line joining $\tilde{pl}_b$ and $\tilde{b}$.
Like the intersection probability (20) and the relative uncertainty of singletons

[1] These transformations are here expressed in the notation of the present paper.

[8], $PraPl[b]$ can be expressed as an affine combination of relative belief and plausibility of singletons:

$$PraPl[b](x) = m_b(x) + \frac{1 - k_b}{k_{pl_b}} pl_b(x) = k_b \tilde{b}(x) + (1 - k_b)\tilde{pl}_b(x). \qquad (21)$$

More to the point, as its definition only involves belief and plausibility values *of singletons*, it is more correct to think of $PraPl[b]$ as of a probability transformation of a probability interval system (rather than an approximation of a belief function)

$$PraPl[(l,u)] \doteq l(x) + \frac{1 - \sum_y l(y)}{\sum_y u(y)} u(x)$$

just like the intersection probability.
However, it is easier to point out its weakness as a representative of probability intervals when put in the above form. Just like in the case of relative belief and plausibility of singletons, $PraPl[(l,u)]$ is not in general consistent with the original probability interval system (l,u).
If there exists an element $x \in \Theta$ such that $l(x) = u(x)$ (the interval has width $\Delta(x)$ equal to zero for that element) we have that

$$\begin{aligned} PraPl[(l,u)](x) &= l(x) + \frac{1 - \sum_y l(y)}{\sum_y u(y)} u(x) = u(x) + \frac{1 - \sum_y l(y)}{\sum_y u(y)} u(x) \\ &= u(x) \cdot \frac{\sum_y u(y) + 1 - \sum_y l(y)}{\sum_y u(y)} > u(x) \end{aligned}$$

as $\frac{\sum_y u(y) + 1 - \sum_y l(y)}{\sum_y u(y)} > 1$, and $PraPl[(l,u)]$ falls outside the interval.

Another fundamental objection against $PraPl[(l,u)]$ arises when we compare it to $p[(l,u)]$. While the latter adds to the lower bound $l(x)$ an equal fraction of the uncertainty $u(x) - l(x)$ for all singletons (4), $PraPl[(l,u)]$ adds to the lower bound $l(x)$ an equal fraction *of the upper bound* $u(x)$, effectively counting twice the evidence represented by the lower bound $l(x)$ (21).
For belief functions, this amounts to adding to the mass value $m_b(x)$ of x yet another fraction of $m_b(x)$ itself, instead of distributing only the remaining mass $pl_b(x) - m_b(x)$ allowed to be assigned to x.

3.6 Comparison on an example

Let us illustrate the above remarks with the aid of the example of Section 2.4. We have seen there that, given its rationale, the intersection probability is guaranteed to be consistent with the original probability interval system.
As stated in Section 3.4 neither relative belief (or, for intervals, the relative

lower bound) nor relative plausibility (the relative upper bound) are consistent representatives of the probability interval (6). We have indeed that

$$\tilde{l}(x) = \tfrac{0.2}{0.9} = 0.222, \quad \tilde{l}(y) = \tfrac{0.4}{0.9} = 0.444, \quad \tilde{l}(z) = \tfrac{0.3}{0.9} = \tfrac{1}{3} = 0.333 > 0.3$$

and the relative lower bound falls outside the interval system (6). Analogously,

$$PrNPl(x) = \tilde{u}(x) = \tfrac{0.8}{2.1} = 0.381, \quad PrNPl(y) = \tilde{u}(y) = \tfrac{1}{2.1} = 0.476,$$
$$PrNPl(z) = \tilde{u}(z) = \tfrac{0.3}{2.1} = \tfrac{1}{7} = 0.143 < 0.3$$

so that the relative upper bound $PrNPl = \tilde{u}$ also falls outside the original interval.
$PraPl[(l, u)]$ also fails to be consistent with the constraint (6). As

$$\frac{1 - \sum_y l(y)}{\sum_y u(y)} = \frac{1 - 0.9}{2.1} = \frac{1}{21}$$

we have again that

$$PraPl(x) = 0.2 + \tfrac{1}{21}0.8 = 0.238, \quad PraPl(y) = 0.4 + \tfrac{1}{2.1}1 = 0.448,$$
$$PraPl(z) = 0.3 + \tfrac{1}{21}0.3 = 0.314 > 0.3.$$

4 Geometry in the probability simplex

We already mentioned that the pignistic function (13) is, geometrically, the center of mass of the set of probabilities (9) consistent with b. A similar credal interpretation can be given for the intersection probability too, once we determine which credal set is associated with an interval probability system (1). For intervals (10) associated with belief functions, this credal set is also strictly related to the credal set $\mathcal{P}[b]$ of all consistent probabilities.
In the following we denote by b_A the unique "categorical" belief function which assigns unitary mass to a single event A: $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0 \ \forall B \neq A$. If we represent a b.f. as the vector

$$b = [b(A), \emptyset \subsetneq A \subsetneq \Theta]'$$

collecting its belief values on all the events of Θ, we can write each belief function b with b.p.a. $m_b(A)$ as [10]

$$b = \sum_{A \subseteq \Theta} m_b(A) b_A \qquad (22)$$

a linear (in particular, convex) combination of the vectors[2] representing categorical b.f.s, with scalar coefficients $m_b(A) \in \mathbb{R}$ determined by the associated basic probability assignment.

In particular, we can write each probability distribution $p : \Theta \to [0, 1]$ in the simplex $\mathcal{P}$ of all probabilities as the convex combination

$$p = \sum_{x \in \Theta} p(x) b_x \tag{23}$$

of the categorical probability functions b_x with $b_x(x) = 1$, $b_x(y) = 0$ for all $y \neq x$, and coefficients given by the respective probability values $p(x)$

4.1 Credal representation of intervals: upper and lower simplices

By definition (9) of $\mathcal{P}[b]$ it follows that the polygon of consistent probabilities can be decomposed into the intersection of a number of polytopes

$$\mathcal{P}[b] = \bigcap_{i=1}^{n-1} \mathcal{P}^i[b] \tag{24}$$

where $\mathcal{P}^i[b]$ is the set of probabilities meeting the lower probability constraint *for size-i events*:

$$\mathcal{P}^i[b] \doteq \Big\{ p \in \mathcal{P} : p(A) \geq b(A), \forall A : |A| = i \Big\}.$$

Note that for $i = n$ the constraint is trivially met by all probability distributions: $\mathcal{P}^n[b] = \mathcal{P}$.
A simple and elegant geometric description of the credal set associated with a probability interval system can be given if we consider instead the sets of *pseudo* probabilities meeting the analogous constraints:

$$T^i[b] \doteq \Big\{ p \in \mathcal{P}' : p(A) \geq b(A), \forall A : |A| = i \Big\}.$$

Here $\mathcal{P}'$ denotes the set of all the *pseudo*-probabilities on Θ, i.e., the functions $p : \Theta \to \mathbb{R}$ which meet the normalization constraint $\sum_{x \in \Theta} p(x) = 1$ but not necessarily the non-negativity one. In other words, there may exist an element $x \in \Theta$ such that $p(x) < 0$.
In particular we focus here on the set of pseudo-probability distributions which meet the lower constraint *on singletons*

$$T^1[b] \doteq \Big\{ p \in \mathcal{P}' : p(x) \geq b(x) \;\; \forall x \in \Theta \Big\}, \tag{25}$$

[2] Note that we will use the notation b indifferently for a belief function and the associated vector of belief values. No confusion can arise given the context.

and the set $T^{n-1}[b]$ of pseudo-probabilities which meet the lower constraint on events of size $n-1$:

$$\begin{aligned} T^{n-1}[b] &\doteq \Big\{ p \in \mathcal{P}' : p(A) \geq b(A) \quad \forall A : |A| = n-1 \Big\} \\ &= \Big\{ p \in \mathcal{P}' : p(\{x\}^c) \geq b(\{x\}^c) \ \ \forall x \in \Theta \Big\} \\ &= \Big\{ p \in \mathcal{P}' : p(x) \leq pl_b(x) \ \ \forall x \in \Theta \Big\}. \end{aligned} \tag{26}$$

The latter corresponds to the set of pseudo-probabilities which meet the *upper bound for the elements* x of Θ.

The extension to pseudo-probabilities allows to give the credal sets (25) and (26) the form of *simplices*. A *simplex* is the convex closure

$$Cl(\mathbf{v}_1, ..., \mathbf{v}_k) = \Big\{ \mathbf{v} \in \mathbb{R}^d : \mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k, \ \ \sum_i \alpha_i = 1, \ \alpha_i \geq 0 \ \forall i \Big\} \tag{27}$$

of a collection of "affinely independent" points $v_1, ..., v_k$ of a Cartesian space, i.e., points which cannot be expressed as an affine combination of the others:

$$\nexists \Big\{ \alpha_j, j \neq i : \sum_{j \neq i} \alpha_j = 1 \Big\} \text{ such that } v_i = \sum_{j \neq i} \alpha_j v_j.$$

Consider Figure 2. While a triangle is a simplex in $\mathbb{R}^2$, a polygon with four

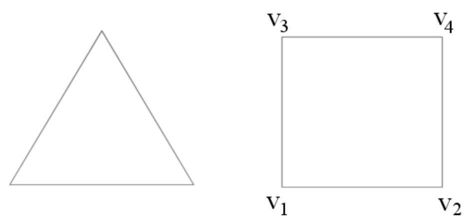

Fig. 2. While a triangle is a simplex in $\mathbb{R}^2$ (a collection of $2+1=3$ affinely independent points), a quadrangle is not, as its vertices are not affinely independent.

vertices is not, as each vertex (e.g. v_4) can be written as an affine combination of the others: $v_4 = 1 \cdot v_2 + 1 \cdot v_3 + (-1) \cdot v_1$.

Using the notation of Equation (23) we can prove the following results.

Theorem 1 *The credal set $T^1[b]$ is the simplex*

$$T^1[b] = Cl(t^1_x[b], x \in \Theta) \tag{28}$$

with vertices

$$t^1_x[b] = \sum_{y \neq x} m_b(y) b_y + \Big(1 - \sum_{y \neq x} m_b(y) \Big) b_x = \sum_{y \neq x} m_b(y) b_y + (m_b(x) + 1 - k_b) b_x. \tag{29}$$

Dually, the credal set $T^{n-1}[b]$ is the simplex

$$T^{n-1}[b] = Cl(t_x^{n-1}[b], x \in \Theta) \tag{30}$$

with vertices

$$t_x^{n-1}[b] = \sum_{y \neq x} pl_b(y) b_y + \left(1 - \sum_{y \neq x} pl_b(y)\right) b_x = \sum_{y \neq x} pl_b(y) b_y + (pl_b(x) + 1 - k_{pl_b}). \tag{31}$$

Lemma 1 *The points $\{t_x^1[b], x \in \Theta\}$ are affinely independent.*

Proof. Let us suppose against the thesis that there exists an affine decomposition of one of the points, say $t_x[b]$, in terms of the others: $t_x^1[b] = \sum_{z \neq x} \alpha_z t_z^1[b]$, $\alpha_z \geq 0 \ \forall z \neq x$, $\sum_{z \neq x} \alpha_z = 1$.
But then we would have by definition of $t_z^1[b]$

$$\begin{aligned} t_x^1[b] &= \sum_{z \neq x} \alpha_z t_z^1[b] = \sum_{z \neq x} \alpha_z \Big(\sum_{y \neq z} m_b(y) b_y \Big) + \sum_{z \neq x} \alpha_z \Big(m_b(z) + 1 - k_b \Big) b_z \\ &= m_b(x) b_x \sum_{z \neq x} \alpha_z + \sum_{z \neq x} b_z m_b(z)(1 - \alpha_z) + \\ &\quad + \sum_{z \neq x} \alpha_z m_b(z) b_z + (1 - k_b) \sum_{z \neq x} \alpha_z b_z \\ &= \sum_{z \neq x} m_b(z) b_z + m_b(x) b_x + (1 - k_b) \sum_{z \neq x} \alpha_z b_z \end{aligned}$$

which is equal to (29)

$$t_x^1[b] = \sum_{z \neq x} m_b(z) b_z + (m_b(x) + 1 - k_b) b_x$$

if and only if

$$\sum_{z \neq x} \alpha_z b_z = b_x.$$

But this is impossible, as the categorical probabilities b_x are trivially affinely independent. □

Proof of Theorem 1. Let us detail the proof for $T^1[b]$. We need to show that:

(1) all the points which belong to $Cl(t_x^1[b], x \in \Theta)$ satisfy $p(x) \geq m_b(x)$ too;
(2) all the points which *do not* belong to the above polytope do not meet the constraint either.

Concerning item (1), as

$$t_x^1[b](y) = \begin{cases} m_b(y) & x \neq y \\ 1 - \sum_{z \neq y} m_b(z) = m_b(y) + 1 - k_b & x = y, \end{cases}$$

$p \in Cl(t_x^1[b], x \in \Theta)$ is equivalent to

$$p(y) = \sum_{x \in \Theta} \alpha_x t_x^1[b](y) = m_b(y) \sum_{x \neq y} \alpha_x + (1 - k_b)\alpha_y + m_b(y)\alpha_y \quad \forall y \in \Theta,$$

where $\sum_x \alpha_x = 1$ and $\alpha_x \geq 0 \ \forall x \in \Theta$. Therefore

$$p(y) = m_b(y)(1 - \alpha_y) + (1 - k_b)\alpha_y + m_b(y)\alpha_y = m_b(y) + (1 - k_b)\alpha_y \geq m_b(y)$$

as $1 - k_b$ and α_y are both non-negative quantities.

Point (2). If $p \notin Cl(t_x^1[b], x \in \Theta)$ then $p = \sum_x \alpha_x t_x^1[b]$ where $\exists z \in \Theta$ such that $\alpha_z < 0$. But then

$$p(z) = m_b(z) + (1 - k_b)\alpha_z < m_b(z)$$

as $(1 - k_b)\alpha_z < 0$, unless $k_b = 1$ in which case b is already a probability.

By Lemma 1 the points $\{t_x^1[b], x \in \Theta\}$ are affinely independent: hence $T^1[b]$ is a simplex.

Dual proofs for Lemma 1 and Theorem 1 can be provided for the set $T^{n-1}[b]$ of pseudo probabilities which meet the upper probability constraint on singletons. We just need to replace the belief values of singletons with their plausibility values. □

We call $T^1[b]$ and $T^{n-1}[b]$ the *lower* and *upper* simplices, respectively.

By Equation (29) each vertex $t_x^1[b]$ of the lower simplex is a probability distribution that adds the mass $1 - k_b$ of non-singletons to the mass of the element x, leaving all the others unchanged:

$$m_{t_x^1[b]}(x) = m_b(x) + 1 - k_b, \quad m_{t_x^1[b]}(y) = m_b(y) \ \forall y \neq x.$$

As $m_{t_x^1[b]}(z) \geq 0 \ \forall z \in \Theta \ \forall x$ (all $t_x^1[b]$ are actual probabilities) we have that

$$T^1[b] = \mathcal{P}^1[b] \tag{32}$$

is *completely included* in the probability simplex.

On the other hand, the vertices (31) of the upper simplex are not guaranteed to be valid probabilities. They are *pseudo* probabilities in the sense that, while meeting the normalization constraint $\sum_x p(x) = 1$, they may assign negative values to some element of Θ.
Each vertex $t_x^{n-1}[b]$ of the upper simplex assigns to each element of Θ different from x its plausibility value $pl_b(y)$, while it subtracts from $pl_b(x)$ the plausibility "in excess" $k_{pl_b} - 1$:

$$m_{t_x^{n-1}[b]}(x) = pl_b(x) + (1 - k_{pl_b}), \quad m_{t_x^{n-1}[b]}(y) = pl_b(y) \ \forall y \neq x.$$

As $1 - k_{pl_b}$ can be a negative quantity, $m_{t_x^{n-1}[b]}(x)$ can be negative and $t_x^{n-1}[b]$ is not guaranteed to be a "real" probability.

In conclusion, by Equations (10), (32) and (26) the set of probabilities consistent with a probability interval system is the intersection

$$\mathcal{P}[b, pl_b] = T^1[b] \cap T^{n-1}[b].$$

4.2 Ternary example

Consider as an example the case of a belief function

$$\begin{array}{lll} m_b(x) = 0.2, & m_b(y) = 0.1, & m_b(z) = 0.3, \\ m_b(\{x, y\}) = 0.1, & m_b(\{y, z\}) = 0.2, & m_b(\Theta) = 0.1 \end{array} \tag{33}$$

defined on a ternary frame $\Theta = \{x, y, z\}$. Figure 3 illustrates the geometry of its consistent simplex $\mathcal{P}[b]$. We can notice that by Equation (24) $\mathcal{P}[b]$ (the polygon delimited by red squares) is in this case the intersection of two triangles (2-dimensional simplices) T^1 and T^2. Notice that, as argued above, not all the vertices of the upper simplex T^2 fall inside the probability simplex (i.e., some of them are pseudo-probabilities). The intersection probability

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

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

4.3 Focus of a pair of simplices

This fact, true in the general case, can be formalized by the notion of "focus" of a pair of simplices.

Definition 1 *Consider a pair of simplices $S = Cl(s_1, ..., s_n)$, $T = Cl(t_1, ..., t_n)$ in $\mathbb{R}^{n-1}$. We call* focus *of the pair (S, T) the unique point $f(S, T)$ of $\mathbb{R}^{n-1}$ which has the same affine coordinates in both simplices:*

$$f(S, T) = \sum_{i=1}^{n} \alpha_i s_i = \sum_{j=1}^{n} \alpha_j t_j, \quad \sum_{i=1}^{n} \alpha_i = 1. \tag{34}$$

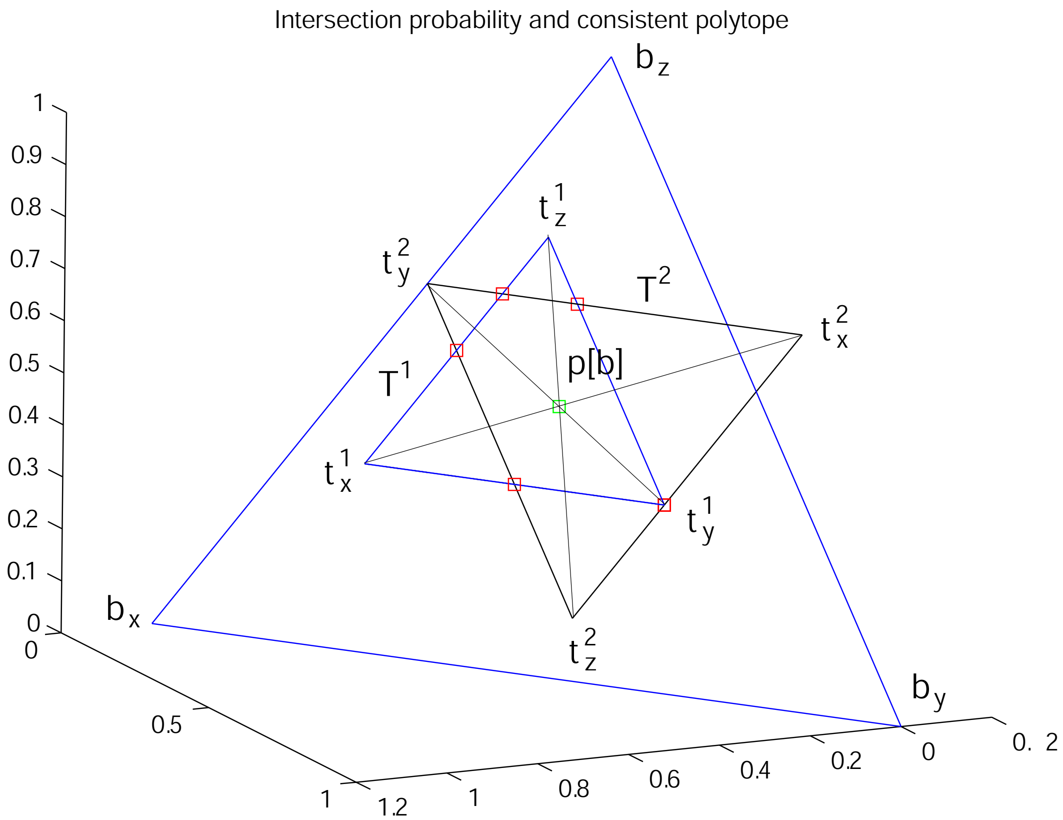

Fig. 3. The polytope of all the probabilities consistent with the belief function (33) is shown here in the simplex $\mathcal{P} = Cl(b_x, b_y, b_z)$ of all probability distributions on $\Theta = \{x, y, z\}$, as the polygon with little squares as vertices. The intersection probability $p[b]$ is the focus of the two simplices $T^1[b]$ and $T^{n-1}[b]$. In the ternary case these reduce to the triangles T^1 and T^2. Their focus is geometrically the intersection of the lines joining their corresponding vertices.

Such a point always exists. As a matter of fact condition (34) can be written as

$$\sum_{i=1}^{n} \alpha_i (s_i - t_i) = 0.$$

As the vectors $\{s_i - t_i, i = 1, ..., n\}$ cannot be linearly independent in $\mathbb{R}^{n-1}$ (since there are n of them) there exists a set of real numbers $\{\alpha'_i, i = 1, ..., n\}$ which meet the above condition. By normalizing these real numbers in order for them to sum to 1, we have the coordinates of the focus.

The focus of two simplices does not always fall in their intersection $S \cap T$ (i.e., α_i is not necessarily non-negative for all i). However, if this is the case, the focus coincides with the unique intersection of the lines $a(s_i, t_i)$ joining corresponding vertices of S and T (see Figure 4-left):

$$f(S, T) = \bigcap_{i=1}^{n} a(s_i, t_i).$$

Suppose indeed that a point p is such that $p = \alpha s_i + (1-\alpha) t_i \ \forall \ i = 1, ..., n$ (i.e.

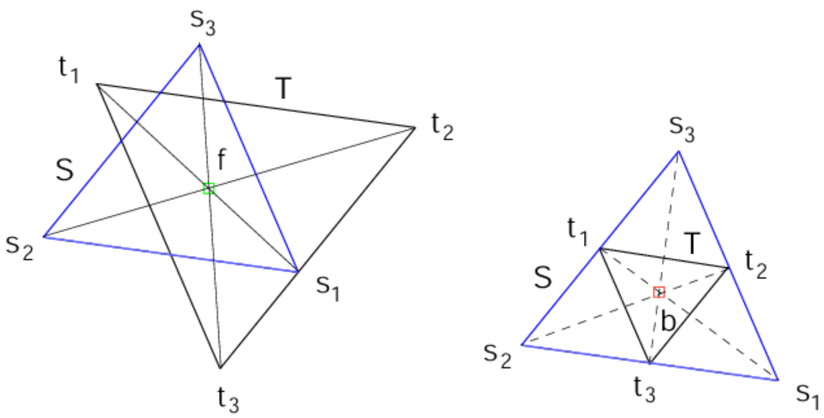

Fig. 4. When internal to both, the focus f of a pair of simplices S, T is the unique intersection of the lines joining corresponding vertices of the two simplices (left). The center of mass b of a simplex S is the focus of the simplex itself and the simplex T formed by the centers of mass of all its $n-1$-dimensional faces. Here a 2-dimensional example is shown (right).

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

$$p = \sum_{i=1}^n \alpha_i t_i = \frac{1}{1-\alpha}\Big[p - \alpha \sum_i \alpha_i s_i\Big]$$

which implies $p = \sum_i \alpha_i s_i$, i.e. p is the focus of (S, T).

The barycenter itself of a simplex is a special case of focus. The center of mass of a d-dimensional simplex S is the intersection of the medians of S, i.e. the lines joining each vertex with the barycenter of the opposite ($d-1$ dimensional) face (see Figure 4-right). But the barycenters of all the $d-1$ dimensional faces of a simplex S form themselves the vertices of another simplex T.

4.4 Intersection probability as focus of upper and lower simplices

Theorem 2 *For each belief function b, the intersection probability $p[b]$ 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 function (5) for b:

$$R[b](x) = \frac{pl_b(x) - m_b(x)}{k_{pl_b} - k_b}. \quad (35)$$

Recalling the expression (29) of the vertices of $T^1[b]$, the point of the simplex $T^1[b]$ with coordinates (35) is

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

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

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

$= p[b]$ by Equation (35). □

4.5 Semantic of foci and a rationality principle

The pignistic function adhere to sensible rationality principles, and as a consequence it has a clean geometrical interpretation as center of mass of the credal set associated with a belief function b. Similarly, the intersection probability has an elegant geometric behavior with respect to the credal set associated with an interval probability system, being the focus of the related upper and lower simplices.
It is quite straightforward to notice that the geometric notion of focus turns out to possess a simple semantic in terms of probability constraints. Selecting the focus of two simplices representing two different constraints (i.e., the point with the same convex coordinates in the two simplices) means adopting the single probability distribution which meets both constraints *in exactly the same way*.
If we assume homogeneous behavior in the two sets of constraints $\{p(x) \geq b(x) \ \forall x\}$, $\{p(x) \leq pl_b(x) \ \forall x\}$ as a rationality principle for the probability transformation of an interval probability system, then the intersection probability necessarily follows as the unique solution to the problem. This provides a rationale for this probability transformation which mirrors the requirement

of homogeneous behavior in the interval constraints for all elements of the frame we gave in Section 2.

5 Original formulation in the belief space

We studied in Section 4 the behavior of the intersection probability in the probability simplex $\mathcal{P}$, starting from the credal interpretations of belief functions and probability intervals. The intersection probability, in fact, had originally been proposed in the framework of the geometric approach to belief measures [8], in which belief functions themselves are represented as points of a convex space [10].
Given a frame of discernment Θ, a belief function $b : 2^\Theta \rightarrow [0, 1]$ is completely specified by its $N - 2$ belief values $\{b(A), \emptyset \subsetneq A \subsetneq \Theta\}$, $N \doteq 2^{|\Theta|}$ (as $b(\emptyset) = 0$, $b(\Theta) = 1$ for all b.f.s), and can then be seen as a point of $\mathbb{R}^{N-2}$.
The "belief space" associated with Θ is the set of points $\mathcal{B} \subset \mathbb{R}^{N-2}$ which correspond to b.f.s. This turns out to be the simplex determined by the convex closure of all the categorical belief functions b_A

$$\mathcal{B} = Cl(b_A,\ \emptyset \subsetneq A \subseteq \Theta)$$

(b_Θ included). The *faces* of a simplex are all the simplices generated by a subset of its vertices. The set of all the Bayesian b.f.s on Θ, $\mathcal{P} = Cl(b_x, x \in \Theta)$, is then a face of $\mathcal{B}$.
Plausibility functions, also determined by their $N - 2$ values $\{pl_b(A), \emptyset \subsetneq A \subsetneq \Theta\}$, can too be seen as points of $\mathbb{R}^{N-2}$. We call "plausibility space" the corresponding region $\mathcal{PL}$ of $\mathbb{R}^{N-2}$, again, a simplex [6].

5.1 *Intersection probability in the binary case*

The geometry of the intersection probability as a belief measure can be appreciated in a simple example. Figure 5 shows the geometry of belief and plausibility spaces for a binary frame $\Theta_2 = \{x, y\}$. Belief and plausibility vectors can be described as points of a plane with coordinates

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

respectively. The two simplices are symmetric with respect to the Bayesian region $\mathcal{P}$, and each pair (b, pl_b) determines a line orthogonal to $\mathcal{P}$, so that b and pl_b lie on symmetric positions on the two sides of the Bayesian region. In the simple binary case the set $\mathcal{P}[b]$ of probabilities compatible with b (9) form

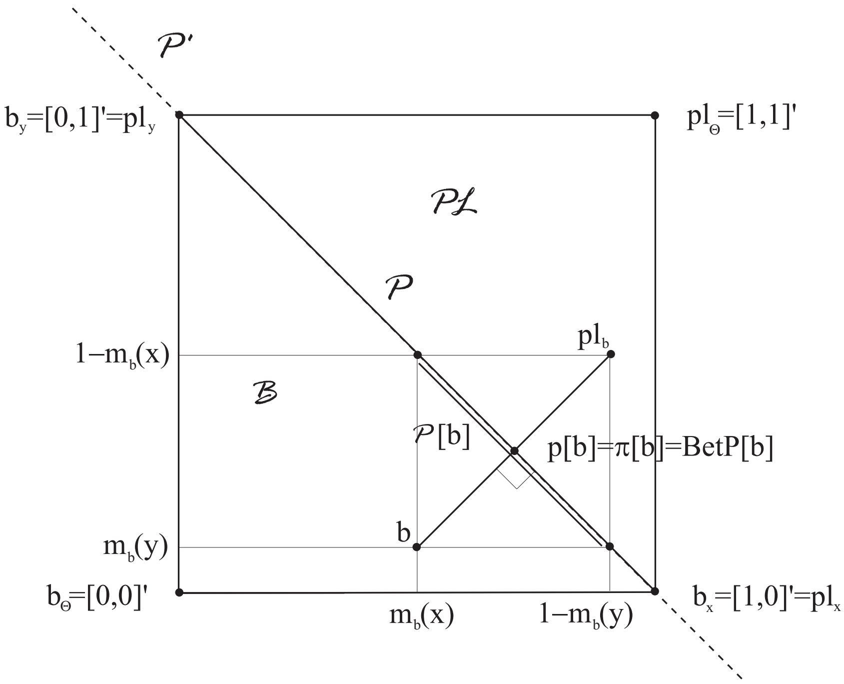

Fig. 5. In a binary frame $\Theta_2 = \{x, y\}$ both belief $\mathcal{B}$ and plausibility $\mathcal{PL}$ spaces are simplices, whose vertices are the categorical belief/plausibility functions $\{b_\Theta = [0,0]', b_x = [1,0]', b_y = [0,1]'\}$ and $\{pl_\Theta = [1,1]', pl_x = b_x, pl_y = b_y\}$, respectively. 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 segment (b, pl_b) with $\mathcal{P}$. The line $\mathcal{P}'$ of all pseudo-probabilities is also drawn.

a segment (1-dimensional simplex) in $\mathcal{P}$ (see Figure 1 again). The intersection probability

$$
\begin{aligned}
p[b](x) &= m_b(x) + \frac{1 - m_b(x) - m_b(y)}{(pl_b(x) - m_b(x)) + (pl_b(y) - m_b(y))} \\
&= m_b(x) + \frac{m_b(\Theta)}{2m_b(\Theta)} = m_b(x) + \frac{m_b(\Theta)}{2}
\end{aligned}
$$

is nothing but the intersection of the segment (b, pl_b) with the Bayesian simplex $\mathcal{P}$, and it coincides (again, in the binary case) with the pignistic function $p[b] = BetP[b]$.

5.2 General case: the geometry of the dual line

In the case of a general frame Θ the situation is a bit more complex. The "dual" line $a(b, pl_b)$ is orthogonal to the affine space $\mathcal{P}'$ of all pseudo-probabilities,

$b - pl_b \perp \mathcal{P}'$ [8], and intersects the latter in the pseudo probability:

$$\varsigma[b] \doteq a(b, pl_b) \cap \mathcal{P}' = b + \beta[b](pl_b - b), \tag{36}$$

where $\beta[b]$ is given by Equation (12). However, such intersection does not fall, in general, inside the set $\mathcal{P}$ of *proper* probabilities. The two functions $\varsigma[b]$ and $p[b]$ are then distinct in the general case. However, the masses they assign to singletons coincide:

$$m_{\varsigma[b]}(x) = m_b(x) + \beta[b] \sum_{A \supsetneq \{x\}} m_b(A) = p[b](x). \tag{37}$$

Indeed, the intersection probability was originally introduced as the unique probability satisfying such condition (37).

We mentioned that important Bayesian transformations (such as relative plausibility and belief of singletons) are inherently related to Dempster's sum (and disjunctive combination [28]) as they commute with $\oplus$:

$$\tilde{pl}_{b_1 \oplus b_2} = \tilde{pl}_{b_1} \oplus \tilde{pl}_{b_2}, \;\; \tilde{b}_{pl_1 \oplus pl_2} = \tilde{b}_{pl_1} \oplus \tilde{b}_{pl_2}.$$

Others however, like the pignistic function and the orthogonal projection $\pi[b]$ of b onto the probability simplex $\mathcal{P}$, are inherently related to *affine combination* of belief functions (as vectors of the belief space). Indeed, they commute with it [8]. Namely, if $\alpha_1 + \alpha_2 = 1$ then

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

On one side, we can make use of the original formulation of the intersection probability in the geometric framework in order to justify its name, by proving that $p[b]$ really behaves like the actual intersection $\varsigma[b]$ (36) when combined with a probability. On the other hand, it makes sense to explore its behavior with respect to affine combination, to be able to relate it to the family of "affine" probability transformations.

6 Intersection probability and operators

6.1 Relation with combination rules: On the name "intersection probability"

Several different combination rules have been proposed to merge the evidence carried by different belief functions. Historically, the first to be formulated has been Dempster's rule [14].

Definition 2 *The* orthogonal sum *or* Dempster's sum *of two belief functions* $b_1, b_2 : 2^\Theta \rightarrow [0,1]$ *is a new belief function* $b_1 \oplus b_2 : 2^\Theta \rightarrow [0,1]$ *with b.p.a.*

$$m_{b_1 \oplus b_2}(A) = \frac{\sum_{B \cap C = A} m_{b_1}(B) m_{b_2}(C)}{\sum_{B \cap C \neq \emptyset} m_{b_1}(B) m_{b_2}(C)} \tag{38}$$

whenever $A \subseteq \Theta$. *We denote by* $k(b_1, b_2)$ *the denominator of (38).*

Other operators have been later proposed, notably in the context of the Transferable Belief Model [26].

Definition 3 *The* conjunctive combination *of two belief functions* $b_1, b_2 : 2^\Theta \rightarrow [0,1]$ *is a new belief function* $b_1 \cap b_2 : 2^\Theta \rightarrow [0,1]$ *with b.p.a.*

$$m_{b_1 \cap b_2}(A) = \sum_{B \cap C = A} m_{b_1}(B) m_{b_2}(C). \tag{39}$$

Their disjunctive combination *is the belief function* $b_1 \cup b_2$ *with b.p.a.*

$$m_{b_1 \cup b_2}(A) = \sum_{B \cup C = A} m_{b_1}(B) m_{b_2}(C). \tag{40}$$

Dempster's, disjunctive and conjunctive rules can be applied to a pair ς_1, ς_2 of pseudo belief functions too, i.e. b.f.s $\varsigma : 2^\Theta \rightarrow [0,1]$, $\varsigma(A) = \sum_{B \subseteq A} m_\varsigma(B)$ whose Moebius inverse

$$m_\varsigma(A) = \sum_{B \subseteq A} (-1)^{|A-B|} \varsigma(B) \tag{41}$$

is not necessarily non-negative. It suffices to apply (38), (39), or (40) to their Moebius inverses m_{ς_1}, m_{ς_2}.

Their application to the pseudo probability $\varsigma[b]$ (37) provides a justification for the name "intersection probability" we gave to $p[b]$. It turns out indeed that $p[b]$ and $\varsigma[b]$ are *equivalent* when combined with a Bayesian belief function. We first need to recall that [7]

Proposition 1 *The orthogonal sum* $b \oplus (\alpha_1 b_1 + \alpha_2 b_2)$ *of a belief function* b *and any affine combination* $\alpha_1 b_1 + \alpha_2 b_2$, $\alpha_1 + \alpha_2 = 1$ *of other two belief functions* b_1, b_2 *on the same frame reads as*

$$b \oplus (\alpha_1 b_1 + \alpha_2 b_2) = \gamma_1 (b \oplus b_1) + \gamma_2 (b \oplus b_2), \tag{42}$$

where

$$\gamma_i = \frac{\alpha_i k(b, b_i)}{\alpha_1 k(b, b_1) + \alpha_2 k(b, b_2)}$$

and $k(b, b_i)$ *is the normalization factor of the orthogonal sum* $b \oplus b_i$.

Similar results can be proven for both conjunctive and disjunctive rules.

Lemma 2 *Affine combination commutes with both conjunctive and disjunctive rules:*

$$b \cap (\alpha_1 b_1 + \alpha_2 b_2) = \alpha_1 (b \cap b_1) + \alpha_2 (b \cap b_2),$$
$$b \cup (\alpha_1 b_1 + \alpha_2 b_2) = \alpha_1 (b \cup b_1) + \alpha_2 (b \cup b_2)$$

whenever $\alpha_1 + \alpha_2 = 1$.

Proof. By definition (39), we have that $b \cap (\alpha_1 b_1 + \alpha_2 b_2)$ has b.p.a.:

$$\begin{aligned} m_{b\cap(\alpha_1 b_1+\alpha_2 b_2)}(A) &= \sum_{B\cap C=A} m_b(B) m_{\alpha_1 b_1+\alpha_2 b_2}(C) \\ &= \sum_{B\cap C=A} m_b(B)\Big(\alpha_1 m_{b_1}(C) + \alpha_2 m_{b_2}(C)\Big) \\ &= \alpha_1 \sum_{B\cap C=A} m_b(B) m_{b_1}(C) + \alpha_2 \sum_{B\cap C=A} m_b(B) m_{b_2}(C) \\ &= \alpha_1 m_{b\cap b_1}(A) + \alpha_2 m_{b\cap b_2}(A). \end{aligned}$$

An analogous proof holds for (40). □

6.1.1 Dempster's and conjunctive rules

Theorem 3 *The combinations of $p[b]$ and $\varsigma[b]$ with any probability function $p \in \mathcal{P}$ coincide under Dempster's (38) and conjunctive (39) rules:*

$$p[b] \oplus p = \varsigma[b] \oplus p, \qquad p[b] \cap p = \varsigma[b] \cap p, \ \forall p \in \mathcal{P}. \tag{43}$$

Proof. Let us define by $\mu(A) = \sum_{B\subseteq A}(-1)^{|A-B|} pl_b(B)$ the Moebius inverse of a plausibility function pl_b. It can be proven that [12]

$$\sum_{A\supseteq\{x\}} \mu_b(A) = m_b(x). \tag{44}$$

Now,applying Equation (42) to $\varsigma \oplus p$ yields:

$$\begin{aligned} \varsigma \oplus p &= \Big[\beta[b] pl_b + (1-\beta[b]) b\Big] \oplus p \\ &= \frac{\beta[b] k(p, pl_b) pl_b \oplus p + (1-\beta[b]) k(p,b) b \oplus p}{\beta[b] k(p, pl_b) + (1-\beta[b]) k(p,b)} \end{aligned} \tag{45}$$

where

$$\begin{aligned} k(p, pl_b) &= \sum_{x\in\Theta} p(x)\Big(\sum_{A\supseteq\{x\}} \mu_b(A)\Big) = \sum_{x\in\Theta} p(x) m_b(x), \\ k(p, b) &= \sum_{x\in\Theta} p(x)\Big(\sum_{A\supseteq\{x\}} m_b(A)\Big) = \sum_{x\in\Theta} p(x) pl_b(x). \end{aligned}$$

by Equation (44), and by definition of plausibility of singletons $\sum_{A\supseteq\{x\}} m_b(A) = pl_b(x)$.

On the other hand, recalling Equation (20), we can write

$$p[b] = \beta[b]\bar{pl}_b + (1-\beta[b])\bar{b} \tag{46}$$

where we call the quantities

$$\bar{pl}_b \doteq \sum_{x\in\Theta} pl_b(x) b_x \qquad \bar{b} = \sum_{x\in\Theta} m_b(x) b_x \tag{47}$$

plausibility of singletons and *belief of singletons* respectively [5].
Therefore when we apply (42) to $p[b]\oplus p$, instead, we get (by Equation (46)):

$$\begin{aligned} p[b]\oplus p &= \left[\beta[b]\bar{pl}_b + (1-\beta[b])\bar{b}\right]\oplus p \\ &= \frac{\beta[b]k(p,\bar{pl}_b)\bar{pl}_b\oplus p + (1-\beta[b])k(p,\bar{b})\bar{b}\oplus p}{\beta[b]k(p,\bar{pl}_b) + (1-\beta[b])k(p,\bar{b})}. \end{aligned} \tag{48}$$

By definition of Dempster's combination (38):

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

Hence

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

as:

(1) $\sum_{x\in\Theta} b_x p(x) = p$;
(2) in the calculation of $b\oplus p$, each singleton x is assigned mass $p(x)\cdot\sum_{A\supseteq\{x\}} m_b(A) = p(x)pl_b(x)$;
(3) in the calculation of $pl_b\oplus p$, each singleton x is assigned mass $p(x)\cdot\sum_{A\supseteq\{x\}} \mu(A) = p(x)m_b(x)$ (by Equation (44) again).

After replacing these expressions in the numerator of Equation (48) we can notice that, as

$$\beta[b] = \frac{1-k_b}{k_{pl_b}-k_b}, \quad 1-\beta[b] = \frac{k_{pl_b}-1}{k_{pl_b}-k_b},$$

the contributions of p vanish, leaving expression (45) for $\varsigma\oplus p$.

As conjunctive rule and affine combination commute, and $k(b_1, b_2) = 1$ for each pair of pseudo belief functions b_1, b_2 under conjunctive combination, the proof holds for $\cap$ too. $\square$

Even though $p[b]$ is *not* the actual intersection $\varsigma[b]$ of the line $a(b, pl_b)$ with the region of pseudo probabilities, it behaves exactly like it when aggregated to a probability distribution.

Notice that Theorem 3 is *not* a simple consequence of Voorbraak's representation theorem:

$$b \oplus p = \tilde{pl}_b \oplus p.$$

In fact, a few passages are sufficient to prove that the relative plausibility or "contour function" of $\varsigma[b]$ *is not* $p[b]$, but the probability with values

$$\tilde{pl}_{\varsigma[b]}(x) = \beta[b]m_b(x) + (1 - \beta[b])pl_b(x) \neq p[b](x) = \beta[b]pl_b(x) + (1 - \beta[b])m_b(x).$$

6.1.2 Disjunctive rule

Dempster's and conjunctive rules are characterized by the fact that $b \oplus p \in \mathcal{P}$, $b \cap p \in \mathcal{P}$ for any probability function p. Theorem 3 confirms this fact. The (dual) disjunctive rule, on the other hand, do not possess the same property. In fact, it can be proven that

$$b \cup cs \in \mathcal{CS}$$

whenever $cs \in \mathcal{CS}$ is a *consistent* belief function, i.e., a b.f. whose focal elements have non-empty intersection.
This is reflected in its relation with the intersection probability. Even though the representation property (43) does not hold for disjunctive combinations, a weaker result can be proven. We first need the following Lemma.

Lemma 3 *Whenever b is a normalized belief function ($m_b(\emptyset) = 0$) and $p \in \mathcal{P}$ a probability measure, we have that $\widetilde{p \cup b} = \widetilde{p \cup \tilde{b}}$.*

Proof. By definition of disjunctive combination, the belief function $p \cup \tilde{b}$ assigns to each singleton $x \in \Theta$ the mass

$$p(x) \cdot \tilde{b}(x) = p(x)\frac{m_b(x)}{k_b},$$

and by normalizing we get

$$\widetilde{p \cup \tilde{b}}(x) = \frac{p(x)m_b(x)}{k_b} \cdot \frac{k_b}{\sum_{y \in \Theta} p(y)m_b(y)} = \frac{p(x)m_b(x)}{\sum_{y \in \Theta} p(y)m_b(y)}.$$

On the other hand, $p \cup b$ assigns to each singleton x the mass $p(x)m_b(x)$ (as $A = \{x\}$ is the only event such that $A \cup \{x\} = \{x\}$, since b is normalized). By normalization,

$$\widetilde{p \cup b}(x) = \frac{p(x)m_b(x)}{\sum_{y \in \Theta} p(y)m_b(y)}.$$

□

Theorem 4 *The relative belief of singletons of the disjunctive combinations of* $\varsigma[b]$ *and the intersection probability* $p[b]$ *with any probability function coincide:*

$$\widetilde{\varsigma[b] \cup p} = \widetilde{p[b] \cup p}$$

whenever $p \in \mathcal{P}$*.*

Proof. Recalling the definition $\varsigma = \beta pl_b + (1-\beta)b$ of $\varsigma[b]$ we get

$$\widetilde{\varsigma[b] \cup p} = \beta\widetilde{pl_b \cup p} + (1-\beta)\widetilde{b \cup p} = \beta\widetilde{\tilde{pl}_b \cup p} + (1-\beta)\widetilde{\tilde{b} \cup p},$$

by Lemma 2 and Lemma 3.
Similarly by applying Lemma 2 and Lemma 3 to Equation (46) $p[b] = \beta\bar{pl}_b + (1-\beta)\bar{b}$ we obtain:

$$\begin{aligned}\widetilde{p[b] \cup p} &= \beta\widetilde{\bar{pl}_b \cup p} + (1-\beta)\widetilde{\bar{b} \cup p} = \beta\widetilde{\tilde{\bar{pl}}_b \cup p} + (1-\beta)\widetilde{\tilde{\bar{b}} \cup p} \\ &= \beta\widetilde{\tilde{pl}_b \cup p} + (1-\beta)\widetilde{\tilde{b} \cup p},\end{aligned}$$

as clearly $\tilde{\bar{pl}}_b(x) = \tilde{pl}_b(x)$, $\tilde{\bar{b}}(x) = \tilde{b}(x)$ for all $x \in \Theta$. □

Even though the combinations of $\varsigma[b]$ and $p[b]$ with probability measures do not necessarily coincide under disjunctive rule, the corresponding beliefs of singleton (which are probability distributions) do. A more thorough investigation of the properties of the intersection probability under disjunctive combination is worth to be pursued in the near future.

6.2 Intersection probability and affine combination

We have seen in Section 5 that, in the geometric interpretation of belief functions, all significant entities like the sets of all probabilities, or the set of consistent probabilities, form convex regions of $\mathbb{R}^{N-2}$. Besides, important probability transformations like pignistic function and orthogonal projection commute with the affine combination of b.f.s.
As a matter of fact, the condition under which the intersection probability commutes with affine (and therefore convex) combination is quite interesting.

Theorem 5 *The intersection probability $p[b]$ commutes with the affine combination of two belief functions*

$$p[\alpha_1 b_1 + \alpha_2 b_2] = \alpha_1 p[b_1] + \alpha_2 p[b_2], \quad \alpha_1 + \alpha_2 = 1,$$

if the relative uncertainty of the singletons is the same for the probability interval systems associated with the two b.f.s:

$$R[b_1] = R[b_2].$$

Proof. By definition (11), $p[\alpha_1 b_1 + \alpha_2 b_2] =$

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

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

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

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

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

The condition is clearly met if $R(x) = R[b_1](x) = R[b_2](x)$ $\forall x$. □

6.3 Comparison with the members of the affine family

In virtue of its elegant relation with affine combination, the intersection probability can be considered a member of a family of Bayesian transformations which includes also pignistic function and orthogonal projection: the "affine" family. It is then natural to discuss the difference between $p[b]$ and its "sister" functions $BetP[b]$ and $\pi[b]$.
As a matter of fact, preliminary sufficient conditions have been already worked out in the recent past [8].

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 following decomposition of $\beta[b]$:

$$\beta[b] = \frac{\sum_{|B|>1} m_b(B)}{\sum_{|B|>1} m_b(B)|B|} = \frac{\sum_{k=2}^n \sum_{|B|=k} m_b(B)}{\sum_{k=2}^n k \cdot \sum_{|B|=k} m_b(B)} = \frac{\sigma_2 + \cdots + \sigma_n}{2\sigma_2 + \cdots + n\sigma_n} \tag{49}$$

where $\sigma_k \doteq \sum_{|B|=k} m_b(B)$.

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.*

This is the case for binary frames, in which all belief functions meet the conditions of both Proposition 2 and Proposition 3. As a result, $p[b] = BetP[b] = \pi[b]$ for all the b.f.s defined on $\Theta = \{x, y\}$ (see Figure 5 again).

More stringent conditions can however be formulated in terms of equal distribution of masses among focal elements.

Theorem 6 *If a belief function* b *is such that its mass is equally distributed among focal elements of the same size*

$$m_b(A) = const \; \forall A : |A| = k, \; \forall k = 2, ..., n. \tag{50}$$

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

Proof. If b meets (50), then the expression (37) 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 (49) of $\beta[b]$.
On the other hand, under the hypothesis, the pignistic function reads as

$$\begin{aligned} BetP[b](x) &= m_b(x) + \sum_{k=2}^n \sum_{A \supsetneq \{x\}, |A|=k} \frac{m_b(A)}{k} = m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{k} \frac{\binom{n-1}{k-1}}{\binom{n}{k}} \\ &= m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{k} \frac{k}{k} = m_b(x) + \sum_{k=2}^n \frac{\sigma^k}{n}, \end{aligned} \tag{51}$$

and the two functions coincide. □

Condition (50) is sufficient to guarantee the equality of intersection probability and orthogonal projection too.

Theorem 7 *If a belief function b meets condition (50) (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 the probability simplex $\mathcal{P}$ has the following expression [8]:

$$\pi[b](x) = \sum_{A \supseteq \{x\}} m_b(A) \Big(\frac{1 + |A^c| 2^{1-|A|}}{n} \Big) + \sum_{A \not\supseteq \{x\}} m_b(A) \Big(\frac{1 - |A| 2^{1-|A|}}{n} \Big). \quad (52)$$

Under condition (50) it becomes

$$\begin{aligned} \pi[b](x) = m_b(x) &+ \sum_{k=2}^{n} \Big(\frac{1 + (n-k)2^{1-k}}{n} \Big) \sum_{A \supseteq \{x\}, |A|=k} m_b(A) \\ &+ \sum_{k=2}^{n} \Big(\frac{1 - (n-k)2^{1-k}}{n} \Big) \sum_{A \not\supseteq \{x\}, |A|=k} m_b(A) \end{aligned} \quad (53)$$

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 (53) yields

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

i.e., the value (51) of the intersection probability under the same assumptions. □

7 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 a probability interval system, and studied the condition under which it commutes with convex combination.

References

[1] M. Bauer, *Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study*, International Journal of Approximate Reasoning **17** (1997), no. 2–3, 217–237.

[2] 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.

[3] B.R. Cobb and P.P. Shenoy, *A comparison of Bayesian and belief function reasoning*, Information Systems Frontiers **5** (2003), no. 4, 345–358.

[4] ——, *On the plausibility transformation method for translating belief function models to probability models*, International Journal of Approximate Reasoning **41** (April 2006), no. 3, 314–330.

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

[6] ——, *Geometry of upper probabilities*, Proc. of the 3^{rd} International Symposium on Imprecise Probabilities and Their Applications, 2003.

[7] ——, *Geometry of Dempster's rule of combination*, IEEE Transactions on Systems, Man and Cybernetics part B **34** (2004), no. 2, 961–977.

[8] ——, *Two new Bayesian approximations of belief functions based on convex geometry*, IEEE Transactions on Systems, Man, and Cybernetics - Part B **37** (2007), no. 4, 993–1008.

[9] ——, *Dual properties of the relative belief of singletons*, Proceedings of the Tenth Pacific Rim Conference on Artificial Intelligence (PRICAI'08), Hanoi, Vietnam, December 15-19 2008, 2008.

[10] ——, *A geometric approach to the theory of evidence*, IEEE Transactions on Systems, Man, and Cybernetics - Part C **38** (2008), no. 4, 522–534.

[11] ——, *Semantics of the relative belief of singletons*, Workshop on Uncertainty and Logic, Kanazawa, Japan, March 25-28 2008, 2008.

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

[13] L. de Campos, J. Huete, and S. Moral, *Probability intervals: a tool for uncertain reasoning*, Int. J. Uncertainty Fuzziness Knowledge-Based Syst. **1** (1994), 167–196.

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

[15] A.P. Dempster, *A generalization of Bayesian inference*, Journal of the Royal Statistical Society, Series B **30** (1968), 205–247.

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

[17] 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** (2002), no. 1-2, 77–101.

[18] J. Dezert and F. Smarandache, *A new probabilistic transformation of belief mass assignment*, 2007.

[19] D. Dubois, H. Prade, and Ph. Smets, *New semantics for quantitative possibility theory.*, ISIPTA, 2001, pp. 152–161.

[20] R. Haenni and N. Lehmann, *Resource bounded and anytime approximation of belief function computations*, International Journal of Approximate Reasoning **31** (2002), no. 1-2, 103–154.

[21] V.-N. Huynh, Y. Nakamori, H. Ono, J. Lawry, V. Kreinovich, and H.T. Nguyen (eds.), *Interval / probabilistic uncertainty and non-classical logics*, Springer, 2008.

[22] I. Levi, *The enterprise of knowledge*, MIT Press, 1980.

[23] J.D. Lowrance, T.D. Garvey, and T.M. Strat, *A framework for evidential reasoning systems*, Readings in uncertain reasoning (Shafer and Pearl, eds.), Morgan Kaufman, 1990, pp. 611–618.

[24] G. Shafer, *A mathematical theory of evidence*, Princeton Univ. Press, 1976.

[25] Ph. 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] ______, *Belief functions : The disjunctive rule of combination and the generalized Bayesian theorem*, International Journal of Approximate Reasoning **9** (1993), 1–35.

[27] ______, *Decision making in the TBM: the necessity of the pignistic transformation*, International Journal of Approximate Reasoning **38** (2005), no. 2, 133–147.

[28] Ph. Smets and R. Kennes, *The transferable belief model*, Artificial Intelligence **66** (1994), no. 2, 191–234.

[29] J.J. Sudano, *Pignistic probability transforms for mixes of low- and high-probability events*, Proceedings of the Fourth International Conference on Information Fusion (ISIF'01), Montreal, Canada, 2001, pp. 23–27.

[30] ———, *Equivalence between belief theories and nave bayesian fusion for systems with independent evidential data*, Proceedings of the Sixth International Conference on Information Fusion (ISIF'03), 2003.

[31] B. Tessem, *Interval probability propagation*, IJAR **7** (1992), 95–120.

[32] ———, *Approximations for efficient computation in the theory of evidence*, Artificial Intelligence **61** (1993), no. 2, 315–329.

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

[34] T. Weiler, *Approximation of belief functions*, IJUFKS **11** (2003), no. 6, 749–777.

[35] 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.