Credal semantics of Bayesian transformations in terms of probability intervals

Fabio Cuzzolin

***Abstract*— In this paper we propose a credal representation of the interval probability associated with a belief function (b.f.), and show how it relates to several classical Bayesian transformations of b.f.s through the notion of "focus" of a pair of simplices. While a belief function corresponds to a polytope of probabilities consistent with it, the related interval probability is geometrically represented by a pair of upper and lower simplices. Starting from the interpretation of the pignistic function as the center of mass of the credal set of consistent probabilities, we prove that relative belief of singletons, relative plausibility of singletons and intersection probability can all be described as foci of different pairs of simplices in the region of all probability measures. The formulation of frameworks similar to the Transferable Belief Model for such Bayesian transformations appears then at hand.**

***Index Terms*— Belief function, interval probability, credal set, Bayesian transformation, upper and lower simplices, focus.**

I. INTRODUCTION

Belief functions are a popular tool to represent uncertain knowledge under scarce information, as they naturally cope with ignorance, qualitative judgements, and missing data [1]. Consider 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 that we have no access to the "correct" probability distribution. The available evidence, though, provides us with some sort of constraint on this unknown distribution. Belief functions (b.f.s) are mathematical descriptions of such a constraint.

Even though in their original definition [2] belief functions are defined as set functions $b : 2^\Theta \rightarrow [0, 1]$ on the power set 2^Θ of a finite universe Θ, their credal interpretation is at the center of a widely adopted approach to the theory of evidence, the "transferable belief model" (TBM) [3], [4]. In the TBM belief functions are indeed represented as "credal sets", i.e. convex sets of probabilities, while decisions are made by resorting to a probability called "pignistic function" $BetP[b]$. Based on a number of rationality principles, the pignistic function has a nice geometric interpretation. It coincides with the barycenter of the set of probability measures "consistent" with b, i.e. the probabilities whose values dominate that of b on all events A:

$$\mathcal{P}[b] \doteq \big\{p \in \mathcal{P} : p(A) \geq b(A)\ \forall A \subseteq \Theta\big\}.$$

Here $\doteq$ denotes "is defined as", while $\mathcal{P}$ denotes the set of all probability measures on Θ.

The relation between belief and probability measures or "Bayesian belief functions" is central in uncertainty theory and in the theory of evidence [2]. The problem has been widely studied, under different points of view [5], [6], [7], [8], [9], [10]. Some work has been directed to find efficient implementations of the rule of combination by reducing the focal elements of the functions to merge [11], [12].

A different approach seeks instead approximations which enjoy commutativity properties with respect to the combination rule, in particular the original Dempster's sum [13]. Voorbraak [14] has been probably the first to explore this direction. He has proposed to adopt the *relative plausibility* function $\tilde{pl}_b$, the unique probability that, given a belief function b with plausibility $pl_b : 2^\Theta \rightarrow [0, 1]$, $pl_b(A) = 1 - b(A^c)$, assigns to each element $x \in \Theta$ of the domain its normalized plausibility. Cobb and Shenoy [15], [16] have later analyzed the properties of the relative plausibility of singletons [17] and discussed its nature of probability function that is "equivalent" (in a very precise sense) to the original belief function. More recently, a dual *relative belief of singletons* has been investigated in terms of both its semantics [18] and its properties with respect to Dempster's rule, which mirror those exhibited by the relative plausibility.

Unlike the case of the pignistic transformation, a credal semantic is still lacking for most major Bayesian approximations of belief functions. We address this issue here in the framework of "interval probabilities". A different constraint on the true probability p which describes the given problem Q can indeed be provided by enforcing lower $l(x)$ and upper $u(x)$ bounds on its probability values on the elements of the frame Θ. Such bounds determine an *interval probability*:

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

Each belief function determines itself an interval probability, in which the lower bound to $p(x)$ is the belief value $b(x)$ of x, while its upper bound is its plausibility value $pl_b(x) = 1 - b(\{x\}^c)$:

$$\mathcal{P}[b, pl_b] \doteq \big\{p \in \mathcal{P} : b(x) \leq p(x) \leq pl_b(x), \forall x \in \Theta\big\}. \quad (2)$$

Now, interval probabilities do possess a credal representation, which for intervals (2) associated with belief functions is also strictly related to the credal set of consistent probabilities. More precisely, the probabilities consistent with a certain

F. Cuzzolin is with the Department of Computing, School of Technology, Oxford Brookes University, Oxford, UK. Email: fabio.cuzzolin@brookes.ac.uk.

interval (2) lie in the intersection of two simplices: a "lower simplex" $T^1[b]$ determined by the lower bound $b(x) \leq p(x)$, and an "upper simplex" $T^{n-1}[b]$ determined by the upper constraint $p(x) \leq pl_b(x)$:

$$\begin{array}{lll} T^1[b] & \doteq & \big\{p : p(x) \geq b(x) \ \ \forall x \in \Theta\big\}, \\ T^{n-1}[b] & \doteq & \big\{p : p(x) \leq pl_b(x) \ \ \forall x \in \Theta\big\}. \end{array}$$

A. Contribution

The credal representation of the interval probability associated with a belief function is the tool we need to provide several important Bayesian approximations with a credal semantic similar to that of the pignistic transformation. In this paper we focus our attention on relative plausibility [14] and belief [18] of singletons, and on the so-called *intersection probability*, a new Bayesian approximation introduced in [19] as the unique representative of a probability interval (1) of the form $p(x) = l(x) + \alpha(u(x) - l(x))$ for all $x \in \Theta$ and for some $\alpha \in [0, 1]$.

We prove here that each of those Bayesian transformations can be described in a homogeneous fashion as the *focus* $f(S, T)$ of a pair S, T of simplices, i.e. the unique probability which has the same coordinates w.r.t. the two simplices. A useful intuition on this notion can be given as follows. When the focus of S and T falls inside their intersection, it coincides with the unique intersection of the lines joining corresponding vertices of S and T (see Figure 1).

Here we consider the following pairs of simplices: $\{\mathcal{P}, T^1[b]\}$,

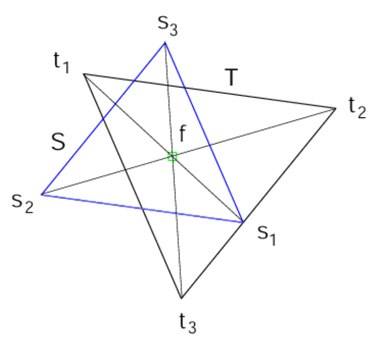

Fig. 1. The focus of a pair of simplices is the unique intersection of the lines joining their corresponding vertices.

$\{\mathcal{P}, T^{n-1}[b]\}$, $\{T^1[b], T^{n-1}[b]\}$. We prove that, while the relative belief of singletons is the focus of $\{\mathcal{P}, T^1[b]\}$, the relative plausibility of singletons is the focus of $\{\mathcal{P}, T^{n-1}[b]\}$, and the intersection probability that of $\{T^1[b], T^{n-1}[b]\}$.

Their focal coordinates encode major features of the underlying belief function: the total mass it assigns to singletons, their total plausibility, and the fraction of the related probability interval which determines the intersection probability.

This provides a consistent, comprehensive credal semantics for a wide family of Bayesian transformations in terms of geometric loci in the probability simplex, which in perspective paves the way for TBM-like frameworks based on those very transformations.

B. Paper outline

We start by recalling the credal interpretation of belief functions and interval probabilities as convex constraints on the value of the unknown probability distribution assumed to describe the problem (Section II). In particular we focus on the credal sets of probabilities consistent with a belief function and an interval probability, respectively, and what we call "lower" and "upper" simplices, i.e. the sets of probability measures which meet the lower and upper probability constraints on singletons.

While the pignistic function has a strong credal interpretation in its capacity of center of mass of the polytope of consistent probabilities, other major Bayesian transformations of belief functions (recalled in Section III) lacked so far an analogous credal interpretation.

Drawing inspiration from the ternary case (Section IV), we prove in Section V that all the considered probability transformations (relative belief and plausibility of singletons, intersection probability) are geometrically the foci of different pairs of simplices, and discuss the meaning of the map associated with a focus in terms of mass assignment. Returning to the original analogy with the pignistic transformation, we prove that upper and lower simplices can themselves be interpreted as the sets of probabilities consistent with belief and plausibility of singletons.

Finally, in Section VI we comment on those results, and prospect alternative reasoning frameworks based on the introduced credal interpretations of upper and lower probability constraints and the associated probability transformations.

The proofs of some preliminary results are given in the Appendix.

II. Credal semantics of belief functions and interval probabilities

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. In most cases it is safe to assume that the available evidence provides us with just some sort of constraint on the true distribution on Q. Indeed, both interval probabilities and belief functions determine different *credal sets* or sets of probability distributions.

A. Credal interpretation of belief functions

A *belief function* (b.f.) $b : 2^\Theta \rightarrow [0, 1]$ on a finite set or "frame" Θ has the form

$$b(A) = \sum_{B \subseteq A} m_b(B), \tag{3}$$

where $m_b : 2^\Theta \rightarrow [0, 1]$ is a set function called "basic probability assignment" (b.p.a.) or "basic belief assignment" (b.b.a.), and is such that $m_b(A) \geq 0 \ \forall A \subseteq \Theta$ and $\sum_{A \subseteq \Theta} m_b(A) = 1$. Events $A \subseteq \Theta$ such that $m_b(A) \neq 0$ are called "focal elements" (f.e.s). Finite probabilities or *Bayesian* b.f.s are those belief functions which assign mass to singletons only: $m_b(A) = 0 \ \forall A : |A| > 1$.

In the following we will denote by b_A the unique b.f. (usually called *categorical* b.f. in the literature) which assigns unitary mass to a single event A: $m_{b_A}(A) = 1, m_{b_A}(B) = 0\ \forall B \neq A$. We can then write each belief function b with b.p.a. m_b as [20]

$$b = \sum_{A\subseteq\Theta} m_b(A) b_A. \tag{4}$$

Belief functions have a natural interpretation as constraints on the "true", unknown probability distribution 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 compatible with b emerges by redistributing the mass of each focal element to its singletons.

1) Example: Let us consider a little toy example, namely a b.f. b on a frame of cardinality three $\Theta = \{x, y, z\}$ with focal elements (Figure 2-top): $m_b(\{x, y\}) = \frac{2}{3}$, $m_b(\{y, z\}) = \frac{1}{3}$. One way of obtaining a probability consistent with b is,

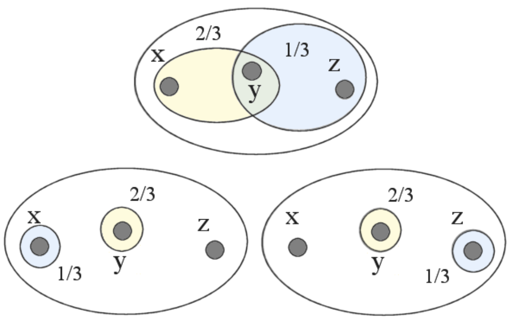

Fig. 2. Top: A simple belief function in a frame of size 3. Bottom: two probabilities consistent with it on the same frame.

for instance, equally sharing the mass of $\{x, y\}$ among its elements x and y, while attributing the entire mass of $\{y, z\}$ to y (Figure 2-bottom-left). Or, we can assign all the mass of the focal element $\{x, y\}$ to y, and give the mass of $\{y, z\}$ to z only, obtaining the Bayesian belief function of Figure 2-bottom-right.

2) Belief function as lower bound: The set of all and only the "consistent" probabilities obtained by re-assigning the mass of each f.e. to its elements turns out to be

$$\mathcal{P}[b] = \big\{p \in \mathcal{P} : p(A) \geq b(A)\ \forall A \subseteq \Theta\big\}, \tag{5}$$

i.e., the set of distributions whose values dominate that of b on all events A. These are well known to form a polytope in the space $\mathcal{P}$ of all probability measures [21], whose center of mass coincides with the pignistic transformation. More precisely, you can prove that [22]

Proposition 1: The polytope $\mathcal{P}[b]$ of all the probability measures consistent with a belief function b can be expressed as the convex closure[1] $P[b] = Cl(p^\rho[b]\ \ \forall\rho)$, where ρ is any permutation $(x_{\rho(1)}, ..., x_{\rho(n)})$ of the elements of $\Theta = \{x_1, ..., x_n\}$, and the vertex $p^\rho[b]$ is the unique Bayesian b.f. such that

$$p^\rho[b](x_{\rho(i)}) = \sum_{A\ni x_{\rho(i)}, A\not\ni x_{\rho(j)} \forall j<i} m_b(A). \tag{7}$$

Each probability function (7) attributes to each singleton $x = x_{\rho(i)}$ the mass of all the focal elements of b which contain it, without containing the elements which precede x in the ordered list $(x_{\rho(1)}, ..., x_{\rho(n)})$ generated by the permutation ρ. When compared to the classical result by Chateauneuf and Jaffray [21], Proposition 1 turns out to be a more stringent claim as it proves that the actual vertices (7) of $\mathcal{P}[b]$ are associated with permutations of elements of Θ. We are going to need this particular result in the proof of Theorem 7.

[1] Cl denotes the convex closure operator:

$$Cl(b_1, ..., b_k) = \Big\{b \in \mathcal{B} : b = \alpha_1 b_1 + \cdots + \alpha_k b_k, \sum_i \alpha_i = 1,\ \alpha_i \geq 0\ \forall i\Big\}. \tag{6}$$

B. Credal interpretation of interval probabilities

An *interval probability* provides instead lower and upper bounds for the probability values of the elements of Θ (singletons):

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

Any belief function determines itself an interval probability, in which the lower bound to $p(x)$ is the belief value $b(x)$ on x, while its upper bound is the *plausibility value* $pl_b(x)$ of x, $\big\{b(x) \leq p(x) \leq pl_b(x), \forall x \in \Theta\big\}$. The plausibility value $pl_b(A) = 1 - b(A^c)$ of an event $A \subseteq \Theta$ expresses the amount of evidence carried by b which is not against A.

Interval probabilities themselves possess a credal representation, which for intervals associated with belief functions is also strictly related to the credal set $\mathcal{P}[b]$ of all consistent probabilities.

1) Credal form: By definition (5) of $\mathcal{P}[b]$ it follows that the polygon of consistent probabilities can be decomposed into a number of polytopes

$$\mathcal{P}[b] = \bigcap_{i=1}^{n-1} \mathcal{P}^i[b], \tag{8}$$

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 distributions: $\mathcal{P}^n[b] = \mathcal{P}$.

In fact, a simple and elegant geometric description can be given if we consider instead the credal sets:

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

Here $\mathcal{P}'$ denotes the set of all *pseudo*-probabilities on Θ, 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: there may exist an element x such that $p(x) < 0$.

In particular we focus here on the set of pseudo-probability measures which meet the lower constraint *on singletons*

$$T^1[b] \doteq \big\{p \in \mathcal{P}' : p(x) \geq b(x) \quad \forall x \in \Theta\big\}, \tag{9}$$

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

$$\begin{array}{lll} T^{n-1}[b] & \doteq & \big\{p \in \mathcal{P}' : p(A) \geq b(A) \quad \forall A : |A| = n-1\big\} \\ & = & \big\{p \in \mathcal{P}' : p(\{x\}^c) \geq b(\{x\}^c) \ \ \forall x \in \Theta\big\} \\ & = & \big\{p \in \mathcal{P}' : p(x) \leq pl_b(x) \ \ \forall x \in \Theta\big\}, \end{array} \tag{10}$$

i.e., the set of pseudo-probabilities which meet the *upper bound for the elements* x of Θ.

2) Simplicial form: The extension to pseudo-probabilities allows to give the credal sets (9) and (10) the form of *simplices*. A *simplex* is the convex closure of a collection of "affinely independent" points $v_1, ..., v_k$, i.e., points which cannot be expressed as an affine combination of the others in the collection: $\nexists \ \{\alpha_j, j \neq i : \sum_{j \neq i} \alpha_j = 1\}$ such that $v_i = \sum_{j \neq i} \alpha_j v_j$. Consider Figure 3. While a triangle is a simplex in $\mathbb{R}^2$, a polygon with four vertices is not, as each vertex (e.g. v_4) can be written as an affine combination of the others: $v_4 = 1 \cdot v_2 + 1 \cdot v_3 + (-1) \cdot v_1$.

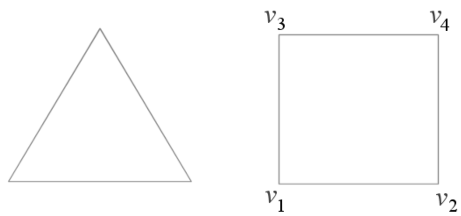

Fig. 3. While a triangle is a simplex in $\mathbb{R}^2$ (a collection of $2+1=3$ affinely independent points), a quadrangle is not as its vertices are not affinely independent.

According to Equation (4), a belief function b with basic probability assignment m_b can be written as $b = \sum_{A \subseteq \Theta} m_b(A) b_A$, where b_A is the categorical b.f. assigning mass 1 to the event A. We extensively use this notation in the following propositions, originally proven in [23].

Proposition 2: The credal set $T^1[b]$ or *lower simplex* can be written as

$$T^1[b] = Cl(t^1_x[b], x \in \Theta), \tag{11}$$

the convex closure of the vertices

$$t^1_x[b] = \sum_{y \neq x} m_b(y) b_y + \bigg(1 - \sum_{y \neq x} m_b(y)\bigg) b_x. \tag{12}$$

Dually, the *upper simplex* $T^{n-1}[b]$ reads as the convex closure

$$T^{n-1}[b] = Cl(t^{n-1}_x[b], x \in \Theta) \tag{13}$$

of the vertices

$$t^{n-1}_x[b] = \sum_{y \neq x} pl_b(y) b_y + \bigg(1 - \sum_{y \neq x} pl_b(y)\bigg) b_x. \tag{14}$$

The proof of Proposition 2 can be found in the Appendix. To further clarify those results, let us denote by

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

the total mass and plausibility of singletons, respectively. By Equation (12) each vertex $t^1_x[b]$ of the lower simplex is a probability that adds the mass $1 - k_b$ of non-singletons to the mass of the element x, leaving all the others unchanged:

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

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

$$T^1[b] = \mathcal{P}^1[b] \tag{15}$$

is *completely included* in the probability simplex.

On the other hand the vertices (14) of the upper simplex are not guaranteed to be valid probabilities, but only *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 $t^{n-1}_x[b]$ assigns to each element of Θ different from x its plausibility $pl_b(y)$, while it subtracts from $pl_b(x)$ the plausibility "in excess" $k_{pl_b} - 1$:

$$\begin{array}{lll} m_{t^{n-1}_x[b]}(x) & = & pl_b(x) + (1 - k_{pl_b}), \\ m_{t^{n-1}_x[b]}(y) & = & pl_b(y) \ \forall y \neq x. \end{array}$$

Now, as $1 - k_{pl_b}$ can be a negative quantity, $m_{t^{n-1}_x[b]}(x)$ can be negative and $t^{n-1}_x[b]$ is not guaranteed to be a "true" probability. We will see this in the example of Section IV.

In conclusion, by Equations (2), (15) and (10) the set of probabilities consistent with a probability interval is the intersection $\mathcal{P}[b, pl_b] = T^1[b] \cap T^{n-1}[b]$.

III. Bayesian transformations of a belief function

The relation between belief and probability measures or "Bayesian belief functions" is central in uncertainty theory [5], [6], [7], [8], [9], [10], and in the theory of evidence [2] in particular.

A. Pignistic function as barycenter of consistent probabilities

In Smets' "Transferable Belief Model" [24], [25], [3], [4] beliefs are represented as convex sets of probabilities, while decisions are made by resorting to a Bayesian b.f. called *pignistic function*

$$BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}. \tag{16}$$

The rationality principle behind the pignistic transformation can be explained in terms of the "floating mass" interpretation of focal elements exposed in Section II-A. If the mass of each focal element is *uniformly* distributed among all its elements, the probability we obtain is (16). The pignistic function has a strong credal interpretation, as it is known [21], [26], [27] to be the center of mass of the set of consistent probabilities: $BetP[b] = bary(\mathcal{P}[b])$.

Many other popular and significant Bayesian functions used to approximate belief functions or to represent them in a decision process, however, *have not* yet a similar credal interpretation. The aim of this paper is indeed to show that relative plausibility [14], relative belief of singletons [18], and intersection probability [19] possess such credal interpretations in terms of the probability interval associated with a belief function.

B. Relative plausibility and belief of singletons

The *relative plausibility of singletons* [14] $\tilde{pl}_b$ is the unique probability that, given a b.f. b with plausibility pl_b, assigns to each singleton its normalized plausibility

$$\tilde{pl}_b(x) \doteq \frac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)} = \frac{pl_b(x)}{k_{pl_b}}. \tag{17}$$

Voorbraak proved that $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities through Dempster's orthogonal sum $\oplus$ [13], $\tilde{pl}_b \oplus p = b \oplus p \; \forall p \in \mathcal{P}$. Cobb and Shenoy [17] later proved that $\tilde{pl}_b$ meets a number of other interesting properties with respect to $\oplus$.

Dually, a *relative belief of singletons* $\tilde{b}$ [18] can be defined. This probability function assigns to the elements of Θ their normalized belief values:

$$\tilde{b}(x) \doteq \frac{b(x)}{\sum_{y\in\Theta} b(y)}. \tag{18}$$

Even though the existence of $\tilde{b}$ is subject to quite a strong condition $k_b = \sum_{x\in\Theta} m_b(x) \neq 0$, the case in which $\tilde{b}$ does not exist is indeed pathological, as it excludes a great deal of belief and probability measures [18].

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 belongs to x. It can be proven that the relative belief of singletons meets a number of properties with respect to Dempster's sum which are the dual of those enjoyed by the relative plausibility of singletons [28]. These two approximations form a strongly linked couple: we will see what this implies in terms of their geometry in the probability simplex.

C. Intersection probability

For any interval probability (1) we can define its *intersection probability* as the unique probability of the form $p(x) = l(x) + \alpha(u(x) - l(x))$ for all $x \in \Theta$ and for some constant scalar value $\alpha \in [0,1]$ such that (Figure 4):

$$\sum_{x\in\Theta} p(x) = \sum_{x\in\Theta} \Big[l(x) + \alpha(u(x) - l(x)) \Big] = 1.$$

This corresponds to the reasonable request that the desired probability, as a candidate to represent the interval (1), should behave homogeneously for each element x of the domain. When the interval is that associated with a belief function

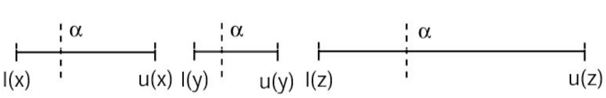

Fig. 4. An illustration of the notion of intersection probability for an upper/lower probability system.

the upper bound to the probability of each singleton is, again, given by the associated plausibility value $u(x) = pl_b(x)$, while the lower bound to $p(x)$ is the corresponding belief value: $l(x) = b(x) = m_b(x)$.

The intersection probability $p(x) = l(x) + \alpha(u(x) - l(x))$ can then be written as [19]

$$p[b](x) = \beta[b] pl_b(x) + (1 - \beta[b]) b(x), \tag{19}$$

as the quantity α in Figure 4 becomes equal to

$$\beta[b] = \frac{1 - k_b}{k_{pl_b} - k_b}. \tag{20}$$

Here again k_{pl_b}, k_b denote the total plausibility and belief of singletons, respectively.

The ratio $\beta[b]$ (20) measures the fraction of the probability interval which we need to add to the lower bound $b(x)$ to obtain a valid distribution. Another interpretation of the intersection probability comes from its alternative form

$$p[b](x) = b(x) + \Big(1 - \sum_x b(x)\Big) R[b](x), \tag{21}$$

where

$$R[b](x) \doteq \frac{pl_b(x) - b(x)}{k_{pl_b} - k_b} = \frac{pl_b(x) - b(x)}{\sum_y (pl_b(y) - b(y))} = \frac{\Delta(x)}{\sum_y \Delta(y)}. \tag{22}$$

The quantity $\Delta(x) \doteq pl_b(x) - b(x)$ measures the width of the probability interval on x, i.e., the uncertainty on the probability value on each element of Θ. Then $R[b](x)$ indicates how much the uncertainty on the probability value on x "weights" on the total uncertainty $\sum_y \Delta(y)$ associated with the interval probability (1). It is therefore natural to call (22) the *relative uncertainty* of singletons. We can observe that $p[b]$ distributes the necessary additional mass to each singleton according to the relative uncertainty it carries within the given interval.

IV. Credal interpretation of Bayesian transformations: The ternary case

Taking inspiration from the important case of the pignistic transformation, we will here be able to prove that the other Bayesian "relatives" of a b.f. also possess a credal interpretation. Let us start from the case of a frame of cardinality three: $\Theta = \{x, y, z\}$. Consider the b.f.

$$\begin{array}{lll} m_b(x) = 0.2, & m_b(y) = 0.1, & m_b(z) = 0.3, \\ m_b(\{x,y\}) = 0.1, & m_b(\{y,z\}) = 0.2, & m_b(\Theta) = 0.1. \end{array} \tag{23}$$

Figure 5 illustrates the geometry of the related consistent simplex $\mathcal{P}[b]$ in the simplex $Cl(b_x, b_y, b_z)$ of all the probability measures on Θ. By Proposition 1 $\mathcal{P}[b]$ has as vertices the probabilities $\rho^1, \rho^2, \rho^3, \rho^4, \rho^5[b]$ identified by red squares in Figure 5

$$\begin{array}{l} \rho^1 = (x,y,z), \\ \quad \rho^1[b](x) = .4, \;\; \rho^1[b](y) = .3, \;\; \rho^1[b](z) = .3; \\ \rho^2 = (x,z,y), \\ \quad \rho^2[b](x) = .4, \;\; \rho^2[b](y) = .1, \;\; \rho^2[b](z) = .5; \\ \rho^3 = (y,x,z), \\ \quad \rho^3[b](x) = .2, \;\; \rho^3[b](y) = .5, \;\; \rho^3[b](z) = .3; \\ \rho^4 = (z,x,y), \\ \quad \rho^4[b](x) = .3, \;\; \rho^4[b](y) = .1, \;\; \rho^4[b](z) = .6; \\ \rho^5 = (z,y,x), \\ \quad \rho^5[b](x) = .2, \;\; \rho^5[b](y) = .2, \;\; \rho^5[b](z) = .6; \end{array} \tag{24}$$

(as the permutations (y,x,z) and (y,z,x) yield the same distribution).

We can notice several interesting facts:

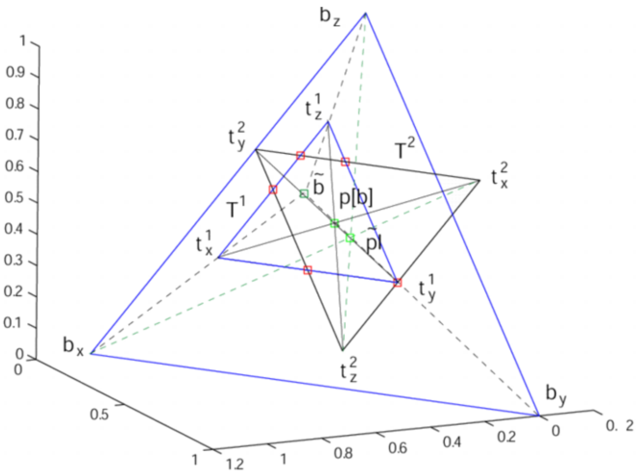

Fig. 5. The simplex of probabilities consistent with the b.f. (23) is shown. Its vertices (red squares) are given by (24). Intersection probability, relative belief and plausibility of singletons are the foci of the pairs of simplices $\{T^1[b], T^2[b]\}$, $\{T^1[b], \mathcal{P}\}$ and $\{\mathcal{P}, T^2[b]\}$, respectively. In the ternary case $T^1[b]$ and $T^2[b]$ are simply triangles. Their focus is geometrically the intersection of the lines joining their corresponding vertices (dashed lines for $\{T^1[b], \mathcal{P}\}$,$\{\mathcal{P}, T^2[b]\}$; solid lines for $\{T^1[b], T^2[b]\}$).

1) $\mathcal{P}[b]$ (the polygon delimited by the red squares) is the intersection of two triangles (2-dimensional simplices) $T^1[b]$ and $T^2[b]$;
2) the relative belief of singletons,
$$\tilde{b}(x) = \frac{.2}{.6} = \frac{1}{3}, \quad \tilde{b}(y) = \frac{.1}{.6} = \frac{1}{6}, \quad \tilde{b}(z) = \frac{.3}{.6} = \frac{1}{2},$$
is the *intersection of the lines joining the corresponding vertices of probability simplex $\mathcal{P}$ and lower simplex $T^1[b]$*;
3) the relative plausibility of singletons,
$$\begin{array}{rcl} \tilde{pl}_b(x) & = & \frac{m_b(x)+m_b(\{x,y\})+m_b(\Theta)}{k_{pl_b}-k_b} \\ & = & \frac{.4}{.4+.5+.6} = \frac{4}{15}, \\ \tilde{pl}_b(y) & = & \frac{.5}{.4+.5+.6} = \frac{1}{3}, \quad \tilde{pl}_b(z) = \frac{2}{5}, \end{array}$$
is the *intersection of the lines joining the corresponding vertices of $\mathcal{P}$ and upper simplex $T^2[b]$*;
4) finally, the intersection probability
$$\begin{array}{rcl} p[b](x) & = & m_b(x) + \beta[b](m_b(\{x,y\}) + m_b(\Theta)) \\ & = & .2 + \frac{.4}{1.5-0.4} 0.2 = .27; \\ p[b](y) & = & .1 + \frac{.4}{1.1} 0.4 = .245; \quad p[b](z) = .485, \end{array}$$
is the unique intersection of the lines joining the corresponding vertices of upper $T^2[b]$ and lower $T^1[b]$ simplices.

Point 1) can be explained by noticing that in the ternary case, by Equation (8), $\mathcal{P}[b] = T^1[b] \cap T^2[b]$. Figure 5 suggests that $\tilde{b}$, $\tilde{pl}_b$ and $p[b]$ might be consistent with b, i.e. they could lie inside the consistent simplex $\mathcal{P}[b]$. This, though, is an artifact of the ternary case.

Theorem 1: The relative belief of singletons is not always consistent.

Proof. Consider a belief function $b : 2^\Theta \rightarrow [0,1]$, $\Theta = \{x_1, x_2, ..., x_n\}$ such that $m_b(x_1) = k_b/n$, $m_b(\{x_1,x_2\}) = 1 - k_b$. Then
$$b(\{x_1,x_2\}) = 2 \cdot \frac{k_b}{n} + 1 - k_b = 1 - k_b\Big(\frac{n-2}{n}\Big),$$
$$\tilde{b}(x_1) = \tilde{b}(x_2) = \frac{1}{n} \Rightarrow \tilde{b}(\{x_1,x_2\}) = \frac{2}{n}.$$
For $\tilde{b}$ to be consistent with b it is necessary that $\tilde{b}(\{x_1,x_2\}) \geq b(\{x_1,x_2\})$. This translates as
$$\frac{2}{n} \geq 1 - k_b \frac{n-2}{n} \equiv k_b \geq 1,$$
i.e., $k_b = 1$. Therefore if $k_b < 1$ (b is not a probability) its relative belief of singletons is not consistent. □

A similar counterexample can be found for $\tilde{pl}_b$. The inconsistency of relative belief and plausibility is due to the fact that those functions only constrain the probabilities of singletons, not considering higher size events as full belief functions do. Indeed these approximations $\tilde{b}$, $\tilde{pl}_b$, $p[b]$ are *consistent with the interval probability $\mathcal{P}[b, pl_b]$ associated with b*:
$$\tilde{b}, \tilde{pl}_b, p[b] \in \mathcal{P}[b, pl_b] = T^1[b] \cap T^{n-1}[b].$$
Their geometric behavior described by points 2), 3) and 4) holds instead in the general case.

V. Credal geometry of Bayesian approximations

We appreciated in the ternary case that relative belief, plausibility and intersection probability lie in the intersection of the lines joining corresponding vertices of pairs formed by the upper simplex, the lower simplex, and the probability simplex. This remark can be formalized through the notion of *focus* of a pair of simplices, laying the foundations of a credal interpretation of these three Bayesian functions.

A. 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:
$$f(S,T) = \sum_{i=1}^n \alpha_i s_i = \sum_{j=1}^n \alpha_j t_j, \quad \sum_{i=1}^n \alpha_i = 1. \qquad (25)$$
Such a point always exists. As a matter of fact condition (25) 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 $S \cap T$. 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 1 again):

$$f(S,T) = \bigcap_{i=1}^{n} a(s_i, t_i). \qquad (26)$$

Suppose indeed that a point p is such that

$$p = \alpha s_i + (1-\alpha)t_i, \qquad \forall\, i = 1, ..., n \qquad (27)$$

(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

$$\begin{array}{lll} p & = & \sum_{i=1}^n \alpha_i t_i = \frac{1}{1-\alpha}\sum_i \alpha_i [p - \alpha s_i] = \\ & = & \frac{1}{1-\alpha}\big[p \sum_i \alpha_i - \alpha \sum_i \alpha_i s_i\big] = \frac{1}{1-\alpha}\big[p - \alpha \sum_i \alpha_i s_i\big]. \end{array}$$

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

Notice that the barycenter itself of a simplex is a special case of focus. Indeed, 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 6). But those barycenters for all $d-1$ dimensional faces form themselves a simplex T.

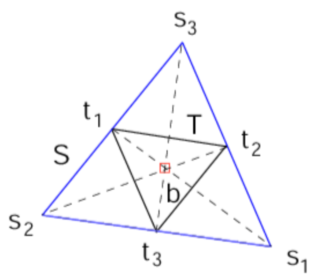

Fig. 6. The barycenter of a simplex is a special case of focus.

B. Relative belief and plausibility of singletons

The notion of focus of a pair of simplices provides a unified geometric interpretation of the coherent family of Bayesian approximations formed by relative belief and plausibility of singletons, and intersection probability.

1) Bayesian approximations as foci:

Theorem 2: The relative belief of singletons is the focus of the pair of simplices $\{\mathcal{P}, T^1[b]\}$.

Proof. We need to prove that $\tilde{b}$ has the same simplicial coordinates in $\mathcal{P}$ and $T^1[b]$. By definition (18) $\tilde{b}$ can be expressed in terms of the vertices of the probability simplex $\mathcal{P}$ as

$$\tilde{b} = \sum_{x\in\Theta} \frac{m_b(x)}{k_b} b_x.$$

We then need to prove that $\tilde{b}$ can be written as the same affine combination

$$\tilde{b} = \sum_{x\in\Theta} \frac{m_b(x)}{k_b} t_x^1[b]$$

in terms of the vertices $t_x^1[b]$ of $T^1[b]$. Replacing (12) in the above equation yields $\sum_{x\in\Theta} \frac{m_b(x)}{k_b} t_x^1[b] =$

$$\begin{array}{l} = \displaystyle\sum_{x\in\Theta} \frac{m_b(x)}{k_b}\Bigg[\sum_{y\neq x} m_b(y) b_y + \Big(1 - \sum_{y\neq x} m_b(y)\Big) b_x\Bigg] = \\ \quad = \displaystyle\sum_{x\in\Theta} b_x \Big(\frac{m_b(x)}{k_b}\sum_{y\neq x} m_b(y)\Big) + \sum_{x\in\Theta}\frac{m_b(x)}{k_b} b_x + \\ \quad - \displaystyle\sum_{x\in\Theta} b_x \Big(\frac{m_b(x)}{k_b}\sum_{y\neq x} m_b(y)\Big) = \sum_{x\in\Theta}\frac{m_b(x)}{k_b} b_x = \tilde{b}. \end{array}$$

□

A dual result can be proven for the relative plausibility of singletons.

Theorem 3: The relative plausibility of singletons is the focus of the pair $\{\mathcal{P}, T^{n-1}[b]\}$.

Proof. You just need to replace $m_b(x)$ with $pl_b(x)$ in the proof of Theorem 2. □

The notion of focus of upper and lower simplices provides indeed the desired credal semantics for the family of Bayesian approximations linked to Dempster's rule of combination, in terms of the credal set associated with the related interval probability.

2) Line coordinates of $\tilde{b}$, $\tilde{pl}_b$: It is interesting to note that the affine coordinate of both belief and plausibility of singletons as foci on the respective intersecting lines (26) has a meaning in terms of degrees of belief.

Theorem 4: The affine coordinate of $\tilde{b}$ as the focus of $\{\mathcal{P}, T^1[b]\}$ on the corresponding intersecting lines is the inverse of the total mass of singletons.

Proof. In the case of the pair $\{\mathcal{P}, T^1[b]\}$ we can compute the (affine) line coordinate α of $\tilde{b} = f(\mathcal{P}, T^1[b])$ by imposing condition (27). The latter assumes the following form (being $s_i = b_x$, $t_i = t_x^1[b]$): $\sum_{x\in\Theta} \frac{m_b(x)}{k_b} b_x =$

$$\begin{array}{l} = t_x^1[b] + \alpha(b_x - t_x^1[b]) = (1-\alpha)t_x^1[b] + \alpha b_x \\ = (1-\alpha)\Bigg[\displaystyle\sum_{y\neq x} m_b(y) b_y + \big(1 - k_b + m_b(x)\big) b_x\Bigg] + \alpha b_x \\ = b_x\Big[(1-\alpha)\big(1 - k_b + m_b(x)\big) + \alpha\Big] + \\ \quad + \displaystyle\sum_{y\neq x} m_b(y)(1-\alpha) b_y, \end{array}$$

and for $1 - \alpha = \frac{1}{k_b}$, $\alpha = \frac{k_b - 1}{k_b}$ the condition is met. □.

A similar result holds for the relative plausibility of singletons.

Theorem 5: The affine coordinate of $\tilde{pl}_b$ as focus of $\{\mathcal{P}, T^{n-1}[b]\}$ on the corresponding intersecting lines is the inverse of the total plausibility of singletons.

Proof. Again we can compute the line coordinate α of $\tilde{pl}_b = f(\mathcal{P}, T^{n-1}[b])$ by imposing condition (27). The latter assumes the form (being $s_i = b_x$, $t_i = t_x^{n-1}[b]$): $\sum_{x\in\Theta} \frac{pl_b(x)}{k_{pl_b}} b_x =$

$$\begin{array}{l} = t_x^{n-1}[b] + \alpha(b_x - t_x^{n-1}[b]) = (1-\alpha)t_x^{n-1}[b] + \alpha b_x \\ = (1-\alpha)\Bigg[\displaystyle\sum_{y\neq x} pl_b(y) b_y + \big(1 - k_{pl_b} + pl_b(x)\big) b_x\Bigg] + \alpha b_x \\ = b_x\Big[(1-\alpha)\big(1 - k_{pl_b} + pl_b(x)\big) + \alpha\Big] + \\ \quad + \displaystyle\sum_{y\neq x} pl_b(y)(1-\alpha) b_y, \end{array}$$

and for $1-\alpha = \frac{1}{k_{pl_b}}$, $\alpha = \frac{k_{pl_b}-1}{k_{pl_b}}$ the condition is met. □.

C. Intersection probability

An analogous result has recently been proven [23] for the intersection probability (19).

Proposition 3: The intersection probability is the focus of the pair of simplices $\{T^{n-1}[b], T^1[b]\}$.

For sake of completeness the proof of Proposition 3 can be found in the Appendix too.

The line coordinate of the intersection probability as a focus also corresponds to a basic feature of the underlying belief function (or better, the associated interval probability).

Theorem 6: The coordinate of the intersection probability as focus of $\{T^1[b], T^{n-1}[b]\}$ on the corresponding intersecting lines coincides with the ratio $\beta[b]$ (20).

Proof. Again, we need to impose condition (27) on the pair $\{T^1[b], T^{n-1}[b]\}$, or

$$p[b] = t^1_x[b] + \alpha(t^{n-1}_x[b] - t^1_x[b]) = (1-\alpha)t^1_x[b] + \alpha t^{n-1}_x[b]$$

for all the elements $x \in \Theta$ of the frame, α being some constant. This is equivalent to (after replacing the expressions (12), (14) of $t^1_x[b]$ and $t^{n-1}_x[b]$)

$$\begin{array}{l}\displaystyle\sum_{x\in\Theta} b_x\big[m_b(x) + \beta[b](pl_b(x) - m_b(x))\big] = \\ \displaystyle= (1-\alpha)\Big[\sum_{y\neq x} m_b(y)b_y + (1-k_b+m_b(x))b_x\Big] + \\ \displaystyle\quad +\alpha\Big[\sum_{y\neq x} pl_b(y)b_y + \Big(1-\sum_{y\neq x} pl_b(y)\Big)b_x\Big] \\ \displaystyle= (1-\alpha)\Big[\sum_{y\in\Theta} m_b(y)b_y + (1-k_b)b_x\Big] + \\ \displaystyle\quad +\alpha\Big[\sum_{y\in\Theta} pl_b(y)b_y + (1-k_{pl_b})b_x\Big] \\ \displaystyle= b_x\Big[(1-\alpha)(1-k_b) + (1-\alpha)m_b(x) + \alpha pl_b(x) + \\ \displaystyle\quad +\alpha(1-k_{pl_b})\Big] + \sum_{y\neq x} b_y\big[(1-\alpha)m_b(y) + \alpha pl_b(y)\big] \\ \displaystyle= b_x\Big\{(1-k_b) + m_b(x) + \\ \displaystyle\quad +\alpha\big[pl_b(x) + (1-k_{pl_b}) - m_b(x) - (1-k_b)\big]\Big\} + \\ \displaystyle\quad +\sum_{y\neq x} b_y\big[m_b(y) + \alpha(pl_b(y) - m_b(y))\big].\end{array}$$

If we set $\alpha = \beta[b] = \frac{1-k_b}{k_{pl_b}-k_b}$ we get for the coefficient of b_x (the probability value of x): $\frac{1-k_b}{k_{pl_b}-k_b}[pl_b(x) + (1-k_{pl_b}) - m_b(x) - (1-k_b)] + (1-k_b) + m_b(x) = \beta[b][pl_b(x) - m_b(x)] + (1-k_b) + m_b(x) - (1-k_b) = p[b](x)$. On the other hand, $m_b(y) + \alpha(pl_b(y) - m_b(y)) = m_b(y) + \beta[b](pl_b(y) - m_b(y)) = p[b](y)$ for all $y \neq x$, no matter the choice of x. □

The fraction $\alpha = \beta[b]$ of the width of the probability interval that generates the intersection probability can be read in the probability simplex as its coordinate on any of the lines determining the focus of $\{T^1[b], T^{n-1}[b]\}$.

1) Mapping associated with a Bayesian transformation: Each pair of simplices $S = Cl(s_1, ..., s_n)$, $T = Cl(t_1, ..., t_n)$ in $\mathbb{R}^{n-1}$ is naturally associated with a mapping, which maps each point of $\mathbb{R}^{n-1}$ with simplicial coordinates α_i in S to the point of $\mathbb{R}^{n-1}$ with the same simplicial coordinates α_i in T:

$$\begin{array}{rcl} F_{S,T} : \mathbb{R}^{n-1} & \to & \mathbb{R}^{n-1} \\ v = \displaystyle\sum_{i=1}^n \alpha_i s_i & \mapsto & F_{S,T}(v) = \displaystyle\sum_{i=1}^n \alpha_i t_i. \end{array} \tag{28}$$

Clearly the focus is the (unique) fixed point of this transformation: $F_{S,T}(f(S,T)) = f(S,T)$. Each Bayesian transformation in 1-1 correspondence with a pair of simplices (relative plausibility, relative belief, and intersection probability) determines therefore a mapping of probabilities to probabilities.

The mapping (28) induced by the relative belief of singletons is actually quite interesting.

Any distribution $p = \sum_x p(x)b_x$ is mapped by $F_{\mathcal{P},T^1[b]}$ to the probability $F_{\mathcal{P},T^1[b]}(p) = \sum_x p(x)t^1_x[b] =$,

$$\begin{array}{rl} = & \displaystyle\sum_{x\in\Theta} p(x)\Big[\sum_{y\neq x} m_b(y)b_y + \Big(1-\sum_{y\neq x} m_b(y)\Big)b_x\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\Big[\Big(1-\sum_{y\neq x} m_b(y)\Big)p(x) + m_b(x)(1-p(x))\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\Big[p(x) - p(x)\sum_{y\in\Theta} m_b(y) + m_b(x)\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\big[m_b(x) + p(x)(1-k_b)\big]. \end{array} \tag{29}$$

The map (29) generates a probability by adding to the belief value of each singleton x a fraction $p(x)$ of mass $(1-k_b)$ of non-singletons. In particular, according to Equation (19), (29) maps the relative uncertainty of singletons $R[b]$ to the intersection probability $p[b]$:

$$\begin{array}{rcl} F_{\mathcal{P},T^1[b]}(R[b]) & = & \displaystyle\sum_{x\in\Theta} b_x\big[m_b(x) + R[b](x)(1-k_b)\big] \\ & = & \displaystyle\sum_{x\in\Theta} b_x p[b](x) = p[b]. \end{array}$$

In a similar fashion, the relative plausibility of singletons is associated with the mapping $F_{\mathcal{P},T^{n-1}[b]}(p) = \sum_{x\in\Theta} p(x)t^{n-1}_x[b] =$

$$\begin{array}{rl} = & \displaystyle\sum_{x\in\Theta} p(x)\Big[\sum_{y\neq x} pl_b(y)b_y + \Big(1-\sum_{y\neq x} pl_b(y)\Big)b_x\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\Big[\Big(1-\sum_{y\neq x} pl_b(y)\Big)p(x) + pl_b(x)(1-p(x))\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\Big[p(x) - p(x)\sum_{y\in\Theta} pl_b(y) + pl_b(x)\Big] \\ = & \displaystyle\sum_{x\in\Theta} b_x\big[pl_b(x) + p(x)(1-k_{pl_b})\big], \end{array} \tag{30}$$

which generates a probability by subtracting to the plausibility of each singleton x a fraction $p(x)$ of plausibility $k_{pl_b} - 1$ in "excess".

It is curious to note that the map associated with $\tilde{pl}_b$ also maps

$R[b]$ to $p[b]$. Indeed, $R[b] \mapsto F_{\mathcal{P},T^{n-1}[b]}(R[b]) =$

$$
\begin{aligned}
&= \sum_{x\in\Theta} b_x \big[pl_b(x) + R[b](x)(1-k_{pl_b})\big] = \\
&= \sum_{x\in\Theta} \left[pl_b(x) + \frac{1-k_{pl_b}}{k_{pl_b}-k_b}(pl_b(x)-m_b(x))\right] \\
&= \sum_{x\in\Theta} b_x \big[pl_b(x) + (\beta[b]-1)(pl_b(x)-m_b(x))\big] \\
&= \sum_{x\in\Theta} \big[\beta[b]pl_b(x) + (1-\beta[b])m_b(x)\big] \\
&= \sum_{x\in\Theta} b_x p[b](x) = p[b].
\end{aligned}
$$

A similar mapping exists for the intersection probability too.

D. Upper and lower simplices as consistent probabilities

Relative belief and plausibility are then the foci associated with lower $T^1[b]$ and upper $T^{n-1}[b]$ simplices, the incarnations of lower and upper constraints on singletons. We can close the circle opened by the analogy with the pignistic transformation by showing that those two simplices can in fact also be interpreted as the sets of probabilities consistent with two important quantities related to b. Let us call the set functions

$$
\begin{aligned}
\bar{pl}_b &\doteq \sum_{x\in\Theta} pl_b(x)b_x + (1-k_b)b_\Theta, \\
\bar{b} &\doteq \sum_{x\in\Theta} m_b(x)b_x + (1-k_{pl_b})b_\Theta
\end{aligned}
\tag{31}
$$

plausibility of singletons and *belief of singletons*, respectively[2] (for obvious reasons). While $\bar{b}$ is a "discounted" probability which assigns the total mass of non-singletons $1-k_b$ to Θ, $\bar{pl}_b$ is a "pseudo" belief function which assigns to Θ the non-positive quantity $1-\sum_x pl_b(x) = 1-k_{pl_b}$. The set of pseudo probabilities consistent with a pseudo b.f. ς can be defined as

$$\mathcal{P}[\varsigma] \doteq \{p \in \mathcal{P}' : p(A) \geq \varsigma(A)\ \forall A \subseteq \Theta\},$$

just as we did for "standard" b.f.s. We can then prove the following result.

Theorem 7: The simplex $T^1[b] = \mathcal{P}^1[b]$ of the lower probability constraint for singletons (9) is the set of probabilities consistent with the belief of singletons $\bar{b}$:

$$T^1[b] = \mathcal{P}[\bar{b}].$$

The simplex $T^{n-1}[b]$ of the upper probability constraint for singletons (10) is the set of pseudo probabilities consistent with the plausibility of singletons $\bar{pl}_b$:

$$T^{n-1}[b] = \mathcal{P}[\bar{pl}_b].$$

Proof. For each belief function b, the vertices of the consistent polytope $\mathcal{P}[b]$ are generated by a permutation ρ of the elements of Θ (7). This is true for the b.f. $\bar{b}$ too, i.e., the vertices of $\mathcal{P}[\bar{b}]$ are also generated by permutations of singletons. In this case though:

- given such a permutation $\rho = (x_{\rho(1)}, ..., x_{\rho(n)})$ the mass of Θ (the only non-singleton focal element of $\bar{b}$) is assigned according to the mechanism of Proposition 1 to $x_{\rho(1)}$, while all the other elements receive only their original mass $m_b(x_{\rho(j)})$, $j > 1$;
- therefore all the permutations ρ putting the same element in the first place yield the same vertex of $\mathcal{P}[\bar{b}]$;
- hence there are just n such vertices, one for each choice of the first element $x_{\rho(1)} = x$;
- but this vertex, a probability distribution, has masses (simplicial coordinates in $\mathcal{P}$)

  $$m(x) = m_b(x) + (1-k_b), \qquad m(y) = m_b(y)\ \forall y \neq x,$$

  as $(1-k_b)$ is the mass $\bar{b}$ assigns to Θ;
- the latter clearly corresponds to $t^1_x[b]$ (12).

A similar proof holds for the case of $\bar{pl}_b$, as Proposition 1 remains valid for pseudo b.f.s too. □

A straightforward consequence is that

Corollary 1: The barycenter $t^1[b]$ of the lower simplex $T^1[b]$ is the pignistic transform of $\bar{b}$:

$$t^1[b] \doteq bary(T^1[b]) = BetP[\bar{b}].$$

The barycenter $t^{n-1}[b]$ of the upper simplex $T^{n-1}[b]$ is the pignistic transform of $\bar{pl}_b$:

$$t^{n-1}[b] \doteq bary(T^{n-1}[b]) = BetP[\bar{pl}_b].$$

Proof. As the pignistic function is the center of mass of the simplex of consistent probabilities, and upper and lower simplices are the sets of probabilities consistent with $\bar{b}$, $\bar{pl}_b$ respectively (by Theorem 7) the thesis follows. □

Another corollary stems from the fact that pignistic function and affine combination commute:

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

whenever $\alpha_1 + \alpha_2 = 1$.

Corollary 2: $p[b] = \beta[b]t^{n-1}[b] + (1-\beta[b])t^1[b]$.

Proof. By Equation (19) the intersection probability $p[b]$ lies on the line joining $\bar{pl}_b$ and $\bar{b}$, with coordinate $\beta[b]$:

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

If we apply the pignistic transformation we get directly

$$
\begin{aligned}
BetP[p[b]] &= p[b] = BetP[\beta[b]\bar{b} + (1-\beta[b])\bar{pl}_b] \\
&= \beta[b]BetP[\bar{b}] + (1-\beta[b])BetP[\bar{pl}_b] \\
&= \beta[b]t^{n-1}[b] + (1-\beta[b])t^1[b]
\end{aligned}
$$

by Corollary 1. □

The intersection probability lies on the segment delimited by the barycenters of the upper and lower simplices, with coordinate $\beta[b]$.

VI. Comments and perspectives

A. A bird's eye view

In summary, all the considered Bayesian transformations of a belief function (pignistic function, relative plausibility, relative belief, and intersection probability) possess a simple credal interpretation in the probability simplex.

Such interpretations have a common denominator, in the sense that they can all be linked to different sets of probabilities which meet the lower probability constraint, in this way

[2]These quantities are clearly related to relative belief and plausibility of singletons, as $\tilde{pl}_b = \bar{pl}_b/k_{pl_b}$, $\tilde{b} = \bar{b}/k_b$.

extending the classical interpretation of the pignistic transformation as barycenter of the polygon of consistent probabilities. As $\mathcal{P}[b]$ is the geometric incarnation of a belief function b, upper and lower simplices geometrically embody the probability interval associated with b:

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

By applying the notion of focus to all the possible pairs of simplices in the triad $\{\mathcal{P}, T^1[b], T^{n-1}[b]\}$ we obtain in turn all the different Bayesian transformations of the considered family:

$$\begin{array}{lll} \big\{\mathcal{P}, T^1[b]\big\}, & f(\mathcal{P}, T^1[b]) & = \tilde{b}, \\ \big\{\mathcal{P}, T^{n-1}[b]\big\}, & f(\mathcal{P}, T^{n-1}[b]) & = \tilde{pl}_b, \\ \big\{T^1[b], T^{n-1}[b]\big\}, & f(T^1[b], T^{n-1}[b]) & = p[b]. \end{array} \quad (32)$$

Their coordinates as foci encode major features of the underlying belief function: the total mass it assigns to singletons, their total plausibility, and the fraction β of the related probability interval which yields the intersection probability.

B. Alternative versions of the transferable belief model

The credal interpretation of upper, lower, and interval probability constraints on singletons lays in perspective the foundations for the formulation of TBM-like frameworks for such systems. In the Transferable Belief Model belief functions b are represented by their credal sets, while decisions are made through the corresponding barycenter, the pignistic function $BetP[b]$.
We can think of the TBM as of a pair $\{\mathcal{P}[b], BetP[b]\}$ formed by a credal set linked to each belief function b (in this case the polytope of consistent probabilities) and a probability transformation (the pignistic function).

If we recall that the barycenter of a simplex is a special case of focus, we realize that the pignistic transformation is a probability transformation induced by the focus of two simplices. The results of this paper, summarized by Equation (32), suggest therefore similar frameworks $\Big\{\big\{\mathcal{P}, T^1[b]\big\}, \tilde{b}\Big\}$, $\Big\{\big\{\mathcal{P}, T^{n-1}[b]\big\}, \tilde{pl}_b\Big\}$, $\Big\{\big\{T^1[b], T^{n-1}[b]\big\}, \tilde{pl}_b\Big\}$ in which lower, upper, and interval constraints on probability distributions on $\mathcal{P}$ are represented by similar pairs, formed by the associated credal set (in the form of a pair of simplices) and by the probability transformation determined by their focus. In such frameworks decisions would be made based on the appropriate focus probability: relative belief, plausibility, or interval probability respectively.

In the TBM [29] disjunctive/conjunctive combination rules are applied to belief functions to update or revise our state of belief according to new evidence. The formulation of similar alternative frameworks for lower, upper, and interval probability systems would then need to design specific evidence elicitation/revision operators for such credal sets. We plan to elaborate on this issue in the near future.

The TMB has been successfully applied to problems as diverse as camera motion classification [30], valuation-based systems [31], climate sensitivity analysis [32]. The role of the pignistic transform in the final decision is crucial in the TBM, and heavily influences the final results. It will be interesting to run experimental comparisons on the classification and decision performances obtained in such real-world scenarios by the application of the different approaches outlined above, in order to empirically compare the different focus-based probability transformations presented here in a coherent theoretical framework.

VII. Conclusions

In this paper we gave a rather comprehensive picture of the behavior of the most common Bayesian approximations of belief functions in the probability simplex, starting from the classical interpretation of the pignistic transformation as barycenter of the polytope of consistent probabilities. We proved that most Bayesian transformations possess a credal interpretation in terms of the notion of focus of a pair of simplices, formed in the different cases by the probability simplex and the upper/lower simplices associated with the upper/lower probability constraint on singletons. Upper and lower simplices incarnate the interval probability associated with a belief function, but can also be interpreted as the polytopes of probabilities consistent with plausibility and belief of singletons. We discussed the possibility that the credal interpretation of upper, lower, and interval probability constraints on singletons might lay the foundations of the formulation of TBM-like frameworks for such systems.

Appendix

Lemma

Lemma 1: The points $\{t^1_x[b], x \in \Theta\}$ 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: $t^1_x[b] = \sum_{z\neq x} \alpha_z t^1_z[b]$, $\alpha_z \geq 0$ $\forall z \neq x$, $\sum_{z\neq x} \alpha_z = 1$.
But then we would have by definition of $t^1_z[b]$

$$\begin{array}{rcl} t^1_x[b] & = & \displaystyle\sum_{z\neq x} \alpha_z t^1_z[b] = \sum_{z\neq x} \alpha_z \Big(\sum_{y\neq z} m_b(y) b_y\Big) + \\ & & \displaystyle + \sum_{z\neq x} \alpha_z \big(m_b(z) + 1 - k_b\big) b_z \\ & = & \displaystyle m_b(x) b_x \sum_{z\neq x} \alpha_z + \sum_{z\neq x} m_b(z)(1-\alpha_z) b_z + \\ & & \displaystyle + \sum_{z\neq x} \alpha_z m_b(z) b_z + (1-k_b) \sum_{z\neq x} \alpha_z b_z \\ & = & \displaystyle \sum_{z\neq x} m_b(z) b_z + m_b(x) b_x + (1-k_b) \sum_{z\neq x} \alpha_z b_z, \end{array}$$

which is equal to (12)

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

if and only if

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

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

Proof of Proposition 2

We need to show that:

1) all the points which belong to $Cl(t^1_x[b], x \in \Theta)$ 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

$$\begin{array}{rcll} t^1_x[b](y) & = & m_b(y) & \text{if } y \neq x, \\ t^1_x[b](x) & = & 1 - \sum_{z \neq x} m_b(z) = m_b(x) + 1 - k_b, & \end{array}$$

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

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

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

$$\begin{array}{rcl} p(y) & = & m_b(y)(1-\alpha_y) + (1-k_b)\alpha_y + m_b(y)\alpha_y = \\ & = & m_b(y) + (1-k_b)\alpha_y \geq m_b(y), \end{array}$$

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

Point 2). If $p \notin Cl(t^1_x[b], x \in \Theta)$ then $p = \sum_x \alpha_x t^1_x[b]$ 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 $\{t^1_x[b], x \in \Theta\}$ are affinely independent: hence $T^1[b]$ is a simplex.

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

Proof of Proposition 3

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 (22) for b:

$$R[b](x) = \frac{pl_b(x) - m_b(x)}{k_{pl_b} - k_b}. \quad (33)$$

Recalling the expression (12) of the vertices of $T^1[b]$, the point of the simplex $T^1[b]$ with simplicial coordinates (33) is $\sum_x R[b](x) t^1_x[b] =$

$$\begin{aligned} &= \sum_x R[b](x) \Big[\sum_{y \neq x} m_b(y) b_y + \Big(1 - \sum_{y \neq x} m_b(y)\Big) b_x \Big] \\ &= \sum_x R[b](x) \Big[\sum_{y \in \Theta} m_b(y) b_y + (1-k_b) b_x \Big] \\ &= \sum_x b_x \Big[(1-k_b) R[b](x) + m_b(x) \sum_y R[b](y) \Big] \\ &= \sum_x b_x \big[(1-k_b) R[b](x) + m_b(x) \big], \end{aligned}$$

as $R[b]$ is a probability ($\sum_y R[b](y) = 1$).

By Equation (21) 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 $\sum_x R[b](x) t^{n-1}_x[b] =$

$$\begin{aligned} &= \sum_x R[b](x) \Big[\sum_{y \neq x} pl_b(y) b_y + \Big(1 - \sum_{y \neq x} pl_b(y)\Big) b_x \Big] \\ &= \sum_x R[b](x) \Big[\sum_{y \in \Theta} pl_b(y) b_y + (1-k_{pl_b}) b_x \Big] = \\ &= \sum_x b_x \Big[(1-k_{pl_b}) R[b](x) + pl_b(x) \sum_y R[b](y) \Big] = \\ &= \sum_x b_x \big[(1-k_{pl_b}) R[b](x) + pl_b(x) \big] \\ &= \sum_x b_x \Big[pl_b(x) \frac{1-k_b}{k_{pl_b} - k_b} - m_b(x) \frac{1-k_b}{k_{pl_b} - k_b} \Big] \end{aligned}$$

$= p[b]$ by Equation (33).

References

[1] G. de Cooman and M. Zaffalon, "Updating beliefs with incomplete observations," *Artif. Intell.*, vol. 159, no. 1-2, pp. 75–125, 2004.

[2] G. Shafer, *A mathematical theory of evidence*. Princeton University Press, 1976.

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

[4] P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation," *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, February 2005.

[5] A. B. Yaghlane, T. Denoeux, and K. Mellouli, "Coarsening approximations of belief functions," in *Proceedings of ECSQARU'2001*, S. Benferhat and P. Besnard, Eds., 2001, pp. 362–373.

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

[7] T. Denoeux and A. B. Yaghlane, "Approximating the combination of belief functions using the Fast Moebius Transform in a coarsened frame," *International Journal of Approximate Reasoning*, vol. 31, no. 1-2, pp. 77–101, October 2002.

[8] R. Haenni and N. Lehmann, "Resource bounded and anytime approximation of belief function computations," *International Journal of Approximate Reasoning*, vol. 31, no. 1-2, pp. 103–154, October 2002.

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

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

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

[12] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61, no. 2, pp. 315–329, 1993.

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

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

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

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

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

[18] F. Cuzzolin, "Semantics of the relative belief of singletons," in *International Workshop on Uncertainty and Logic UNCLOG'08, Kanazawa, Japan*, 2008.

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

[20] ——, "A geometric approach to the theory of evidence," *IEEE Transactions on Systems, Man and Cybernetics part C*, vol. 38, no. 3, May 2008.

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

[22] F. Cuzzolin, "On the structure of simplex of inner Bayesian approximations," in *Proceedings of the European Conference on Logics in Artificial Intelligence (JELIA'08), Dresden, Germany*, 2008.

[23] ——, "Rationale and properties of the intersection probability," *submitted to Artificial Intelligence Journal*, 2008.

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

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

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

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

[28] ——, "Dual properties of the relative belief of singletons," in *Proceedings of the Pacific Rim International Conference on Artificial Intelligence (PRICAI'08), Hanoi, Vietnam*, 2008.

[29] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of Approximate reasoning*, vol. 9, pp. 1–35, 1993.

[30] M. Guironnet, D. Pellerin, and M. Rombaut, "Camera motion classification based on the transferable belief model," in *Proceedings of EUSIPCO'06, Florence, Italy*, 2006.

[31] H. Xu, Y.-T. Hsia, and P. Smets, "The transferable belief model for decision making in the valuation-based systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 26A, pp. 698–707, 1996.

[32] M. Ha-Duong, "Hierarchical fusion of expert opinion in the transferable belief model, application on climate sensivity," HAL, Working Papers halshs-00112129-v3, 2006.