Credal semantics of Bayesian transformations

Fabio Cuzzolin
Oxford Brookes University, Oxford, UK
fabio.cuzzolin@brookes.ac.uk

Abstract

In this paper we propose a credal representation of the set of interval probabilities associated with a belief function, and show how it relates to several classical Bayesian transformations of belief functions through the notion of "focus" of a pair of 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 and plausibility of singletons and intersection probability can be described as foci of different pairs of simplices in the simplex of all probability measures. Such simplices are associated with the lower and upper probability constraints, respectively. This paves the way for the formulation of frameworks similar to the transferable belief model for lower, upper, and interval constraints.

Keywords. Belief functions, credal sets, Bayesian transformations, upper and lower simplices, focus.

1 Introduction

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 true distribution.
Such a constraint is often given in the form of a *credal set*, i.e., the convex set of probability distributions maintained by the agent [14]. A specific class of credal sets is formed by *belief functions* [16]. Even though in their original definition [16] belief functions are defined as set functions $b : 2^\Theta \to [0, 1]$ on the power set 2^Θ of a finite universe Θ, they are equivalent to a set of linear constraints determining a credal set. Belief functions are a popular tool for representing uncertain knowledge under scarce information, as they can naturally cope with ignorance, qualitative judgements, and missing data.
Their credal interpretation is at the core of a widely adopted approach to the theory of evidence, the "Transferable Belief Model" (TBM) [20, 21]. In the TBM, decisions are made by resorting to a probability called "pignistic function". Based on a number of rationality principles, the pignistic function has a nice geometric interpretation as the center of mass of the credal set of probability measures "consistent" with b, i.e. the probabilities that dominate b on all the events A: $\mathcal{P}[b] \doteq \{p \in \mathcal{P} : p(A) \geq b(A) \ \forall A \subseteq \Theta\}$ (here $\mathcal{P}$ denotes the set of all the probability measures on Θ).

The relation between belief and probability measures or "Bayesian belief functions" has been widely studied in the context of the theory of evidence [1, 10, 11, 13, 26, 27], often with different goals. While some authors have looked for efficient implementations of the rule of combination [15, 23], others have argued that Bayesian and belief calculi have the same expressive power as each model can be transformed into the other.
An approach to the Bayesian transformation problem seeks approximations which enjoy commutativity properties with respect to some evidence combination rule, in particular the original Dempster's sum [9]. Voorbraak [24] was probably the first to explore this direction. He proposed to adopt the *relative plausibility of singletons*, i.e., the unique probability that, given a belief function b with plausibility $pl_b : 2^\Theta \to [0, 1]$, $pl_b(A) = 1 - b(A^c)$, assigns to each element $x \in \Theta$ of the domain its normalized plausibility. Cobb and Shenoy later analyzed its properties in detail [3]. More recently, a dual *relative belief of singletons* has been investigated in terms of both its

semantics [7] and its properties with respect to Dempster's rule. The condition under which some of those transformations coincide has been studied in [4].
Unlike the case of the pignistic transformation, a credal semantic is still lacking for most other major Bayesian approximations of belief functions. Moreover, not all such transformations are consistent with the original belief function, i.e., they do not necessarily fall into the corresponding credal set. We address this issue here in the framework of "probability intervals". An admissible constraint on the true probability p which describes the given problem Q can be provided by enforcing lower and upper bounds on its probability values on the elements of the frame Θ. What we get is a set of *probability intervals* [8, 22, 25]:

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

Probability intervals are themselves a special class of credal sets. Besides, each belief function determines itself such a set of intervals, in which the lower bound to $p(x)$ is the belief value $b(x)$ on x, while its upper bound is the plausibility value $pl_b(x) = 1 - b(\{x\}^c)$:

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

The credal set (2) determined by the set of probability intervals associated with a belief function is strictly related to the credal set of consistent probabilities. More precisely, it is the intersection of two higher-dimensional triangles or "simplices": A "lower simplex" $T^1[b]$ determined by the lower bound constraint $b(x) \leq p(x)$, and an "upper simplex" $T^{n-1}[b]$ determined by the upper bound constraint $p(x) \leq pl_b(x)$.

1.1 Contribution

We can exploit the different credal sets associated with a belief function to provide several important Bayesian transformations with a credal semantic similar to that of the pignistic transformation. In this paper we focus on relative plausibility [24] and belief [7] of singletons, and on the so-called *intersection probability*, a new Bayesian approximation introduced in [5]. We prove that each of the above transformations can be geometrically described in a homogeneous fashion as the *focus* $f(S,T)$ of a pair S, T of simplices, i.e., the unique point which has the same coordinates w.r.t. the two simplices. When the focus of two simplices falls into their intersection, it is the unique intersection of the lines joining corresponding vertices of S and T (see Figure 1).

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

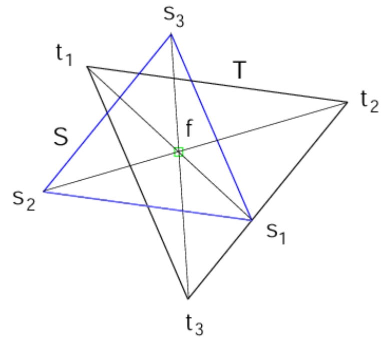

Figure 1: The focus f of a pair of simplices (e.g. two triangles S, T in the 2-D case) is the unique intersection of the lines joining their corresponding vertices.

Their respective 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 intervals 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. In perspective, this paves the way for TBM-like frameworks based on those same transformations.

1.2 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 2). In particular we focus on the credal sets of probabilities consistent with a belief function and a set of probability intervals, respectively, and introduce what we call the "lower" and "upper" simplices, i.e. the sets of probability measures which meet the lower and upper probability constraints on singletons.
As the pignistic function has a strong credal interpretation in its capacity of center of mass of the polytope of consistent probabilities, we can conjecture the existence of an analogous credal interpretation for other major Bayesian transformations of belief functions (Section 3).
Drawing inspiration from the ternary case, we prove in Section 4 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. Finally, in Section 5 we comment on those results, and outline alternative reasoning frameworks

based on the introduced credal interpretations of upper and lower probability constraints and the associated probability transformations.

2 Credal semantics of belief functions and probability intervals

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. They determine different *credal sets* or sets of probability distributions on Θ.

2.1 Credal interpretation of belief functions

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

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

where $m_b : 2^\Theta \to [0,1]$ is a set function called "basic probability assignment" (b.p.a.) or "basic belief assignment", 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". *Bayesian* BFs are belief functions which assign non-zero mass to singletons only: $m_b(A) = 0$ $\forall A : |A| > 1$.
In the following we denote by b_A the unique *categorical* belief function assigning 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 [6]

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

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 event $A \subseteq \Theta$ can float freely among its elements $x \in A$. A probability distribution "consistent" with b emerges by redistributing the mass of each focal element to its singletons.

Example. Let us consider a little example, namely a belief function 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, for instance, to equally share 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.

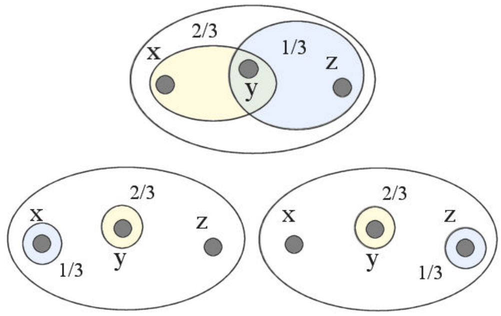

Figure 2: Top: A simple belief function in a frame of size 3. Bottom: two probabilities consistent with it on the same frame.

Belief function as lower bound. The credal set associated with a belief function b (i.e., the set of all the probability distributions consistent with b) is

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

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 [2], whose center of mass coincides with the pignistic transformation. Let us denote by Cl the convex closure operator: $Cl(b_1, ..., b_k) = \{b \in \mathcal{B} : b = \alpha_1 b_1 + \cdots + \alpha_k b_k, \sum_i \alpha_i = 1, \ \alpha_i \geq 0 \ \forall i\}$, where $\mathcal{B}$ is the space of all belief functions.

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 $P[b] = Cl(p^\rho[b] \ \forall \rho)$, 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 BF such that*

$$p^\rho[b](x_{\rho(i)}) = \sum_{A \ni x_\rho(i), A \not\ni x_\rho(j) \forall j < i} m_b(A). \qquad (6)$$

Each probability function (6) assigns to each singletons $x = x_{\rho(i)}$ the mass of all the focal elements of b which contain it, but do not contain the elements which precede x in the ordered list $(x_{\rho(1)}, ..., x_{\rho(n)})$ generated by the permutation ρ.

2.2 Credal interpretation of probability intervals

A *set of probability intervals* 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 such a set of intervals, 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, $\{b(x) \le p(x) \le pl_b(x), \forall x \in \Theta\}$. The plausibility function $pl_b(A) = 1 - b(A^c)$ expresses the evidence not against an event A.

Probability intervals possess themselves 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.

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

$$\mathcal{P}[b] = \bigcap_{i=1}^{n-1} \mathcal{P}^i[b], \qquad (7)$$

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

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

Note that for $i = n$ the constraint is trivially met by all the probability distributions: $\mathcal{P}^n[b] = \mathcal{P}$.

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

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

where $\mathcal{P}'$ denotes the set of all *pseudo*-probabilities[1] 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: it 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*

$$T^1[b] \doteq \{p \in \mathcal{P}' : p(x) \ge b(x) \quad \forall x \in \Theta\}, \qquad (8)$$

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

$$\begin{array}{ll} \doteq & \{p \in \mathcal{P}' : p(A) \ge b(A) \quad \forall A : |A| = n-1\} \\ = & \{p \in \mathcal{P}' : p(\{x\}^c) \ge b(\{x\}^c) \ \ \forall x \in \Theta\} \\ = & \{p \in \mathcal{P}' : p(x) \le pl_b(x) \ \ \forall x \in \Theta\}, \end{array} \qquad (9)$$

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

Simplicial form. The generalization to pseudo-probabilities allows to give the credal sets (8) and (9) the form of *simplices*. A *simplex* is the convex closure of a collection of "affinely independent" points $v_1, ..., v_k$ of a vector space, i.e., points which cannot be expressed as an affine combination of the others:

$$\nexists \left\{\alpha_j, j \ne i : \sum_{j\ne i} \alpha_j = 1\right\} \ s.t. \ v_i = \sum_{j\ne i} \alpha_j v_j.$$

[1] Also called "normalized signed measures" in measure theory.

The notation introduced in Equation (4) is extensively used in the following [4].

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

$$T^1[b] = Cl(t^1_x[b], x \in \Theta) \qquad (10)$$

of the vertices

$$t^1_x[b] = \sum_{y\ne x} m_b(y) b_y + \Big(1 - \sum_{y\ne x} m_b(y)\Big) b_x. \qquad (11)$$

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

$$T^{n-1}[b] = Cl(t^{n-1}_x[b], x \in \Theta) \qquad (12)$$

of the vertices

$$t^{n-1}_x[b] = \sum_{y\ne x} pl_b(y) b_y + \Big(1 - \sum_{y\ne x} pl_b(y)\Big) b_x. \qquad (13)$$

To clarify the above results, let us denote by

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

the total mass and plausibility of singletons, respectively. By Equation (11) each vertex $t^1_x[b]$ of the lower simplex is the distribution that adds the mass $1 - k_b$ of non-singletons to the original mass of the element x, leaving all the others unchanged:

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

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

$$T^1[b] = \mathcal{P}^1[b] \qquad (14)$$

is *completely included* in the probability simplex $\mathcal{P}$.

On the other hand the vertices (13) of the upper simplex are not guaranteed to be valid probabilities, but only *pseudo*-probabilities in the sense that 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" $k_{pl_b} - 1$:

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

As $1 - k_{pl_b}$ can be a negative quantity, $m_{t^{n-1}_x[b]}(x)$ can be negative too and $t^{n-1}_x[b]$ is not guaranteed to be a "true" probability. We will see this in Section 4.

In conclusion, by Equations (2), (14) and (9) the credal set of probabilities consistent with a probability interval is the intersection[2] $\mathcal{P}[b, pl_b] = T^1[b] \cap T^{n-1}[b]$.

[2] This credal set is an outer approximation [10] of $\mathcal{P}[b]$.

3 Bayesian transformations

The relation between belief and probability measures or "Bayesian belief functions" is central in uncertainty theory [1, 10, 11, 13, 27], and in the theory of evidence [16] in particular.

3.1 Pignistic function as center of mass of consistent probabilities

In Smets' "Transferable Belief Model" [17, 18, 20, 21] beliefs are represented as convex sets of probabilities, while decisions are made by resorting to a Bayesian BF called *pignistic function*:

$$BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}. \tag{15}$$

The rationality principle behind the pignistic function can be explained in terms of the "floating mass" interpretation of focal elements exposed in Section 2.1. If the mass of each focal element is *uniformly* distributed among all its elements, the probability we obtain is (15). The pignistic function $BetP[b]$ has a strong credal interpretation, as it is known [2, 12] to be the center of mass of the set $\mathcal{P}[b]$ of probabilities consistent with 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 [24], relative belief of singletons [7], and intersection probability [5] possess such credal interpretations in terms of the probability intervals associated with a belief function.

3.2 Relative plausibility and belief

The *relative plausibility of singletons* [24] $\tilde{pl}_b$ is the unique probability that, given a belief function b with plausibility pl_b, assigns to each singleton its normalized plausibility:

$$\tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_{y \in \Theta} pl_b(y)} = \frac{pl_b(x)}{k_{pl_b}}. \tag{16}$$

Voorbraak has proven that $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities through Dempster's orthogonal sum $\oplus$ [9], $\tilde{pl}_b \oplus p = b \oplus p \ \forall p \in \mathcal{P}$. Cobb and Shenoy [3] have later shown that (16) meets a number of other interesting properties with respect to $\oplus$.

Dually, a *relative belief of singletons* $\tilde{b}$ [7] can be defined. This probability function assigns to the elements of Θ their normalized belief values:

$$\tilde{b}(x) \doteq \frac{b(x)}{\sum_{y \in \Theta} b(y)}. \tag{17}$$

Even though the existence of (17) 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 number of belief and probability measures [7].

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 BF b (seen as a constraint) assures it belong to x.

It can be proven that relative belief meets a number of dual properties with respect to Dempster's sum which are the dual of those enjoyed by relative plausibility [7]. These two approximations form a strongly linked couple: we will see what this implies in terms of their geometry in the probability simplex.

3.3 Intersection probability

For any set of probability intervals (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 $x \in \Theta$ for some $\alpha \in [0, 1]$ such that:

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

(see Figure 3). This corresponds to the reasonable request that the desired probability, as a candidate to represent the set of intervals (1), should behave homogeneously for each element x of the domain. When

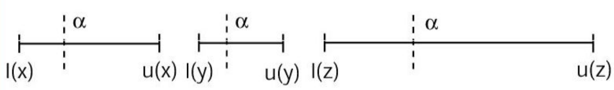

Figure 3: An illustration of the notion of intersection probability for an upper/lower probability system.

the set of intervals is that associated with a belief function, the upper bound to the probability of a singleton is obviously $u(x) = pl_b(x)$, its lower bound $l(x) = b(x) = m_b(x)$. The intersection probability can then be written as [5]

$$p[b](x) = \beta[b] pl_b(x) + (1 - \beta[b]) b(x) \tag{18}$$

as the quantity α of Figure 3 has value

$$\beta[b] = \frac{1 - k_b}{k_{pl_b} - k_b}. \tag{19}$$

Here k_{pl_b}, k_b denote again the total plausibility and belief of singletons, respectively.
The ratio $\beta[b]$ (19) measures the fraction of *each* 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

$$p[b](x) = b(x) + \Big(1 - \sum_x b(x)\Big) R[b](x) \qquad (20)$$

where

$$R[b](x) \doteq \frac{pl_b(x) - b(x)}{k_{pl_b} - k_b} = \frac{pl_b(x) - b(x)}{\sum_y (pl_b(y) - b(y))}. \qquad (21)$$

The quantity $pl_b(x) - b(x)$ 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 associated with the set of intervals (1). It is the natural to call it *relative uncertainty* on singletons.
According to (20), $p[b]$ re-distributes to each $x \in \Theta$ a fraction of the mass of non-singletons $(1 - \sum_x b(x))$ in proportion to the relative uncertainty $R[b](x)$ of each singleton in the set of intervals.

4 Credal interpretation of Bayesian approximations

4.1 The ternary case

Taking inspiration from the important case of the pignistic transformation, here we will be able to prove that other Bayesian transformations of belief functions possess a similar credal interpretation.
Let us first analyze the case of a frame of cardinality three: $\Theta = \{x, y, z\}$. Consider the BF

$$\begin{array}{lll} m_b(x) = 0.2, & m_b(y) = 0.1, & m_b(z) = 0.3, \\ m_b(\{x,y\}) = 0.1, & m_b(\{y,z\}) = 0.2, & m_b(\Theta) = 0.1. \end{array} \qquad (22)$$

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

$$\begin{array}{l} \rho^1 = (x,y,z), \\ \quad \rho^1[b](x) = .4, \ \ \rho^1[b](y) = .3, \ \ \rho^1[b](z) = .3; \\ \rho^2 = (x,z,y), \\ \quad \rho^2[b](x) = .4, \ \ \rho^2[b](y) = .1, \ \ \rho^2[b](z) = .5; \\ \rho^3 = (y,x,z), \\ \quad \rho^3[b](x) = .2, \ \ \rho^3[b](y) = .5, \ \ \rho^3[b](z) = .3; \\ \rho^4 = (z,x,y), \\ \quad \rho^4[b](x) = .3, \ \ \rho^4[b](y) = .1, \ \ \rho^4[b](z) = .6; \\ \rho^5 = (z,y,x), \\ \quad \rho^5[b](x) = .2, \ \ \rho^5[b](y) = .2, \ \ \rho^5[b](z) = .6; \end{array} \qquad (23)$$

(as the permutations (y, x, z) and (y, z, x) yield the same probability distribution). We can notice that:

1. $\mathcal{P}[b]$ (the polygon delimited by little squares) is the intersection of two triangles (2-dimensional simplices) $T^1[b]$ and $T^2[b]$;

2. the relative belief of singletons

$$\tilde{b}(x) = \frac{1}{3}, \ \ \tilde{b}(y) = \frac{1}{6}, \ \ \tilde{b}(z) = \frac{1}{2}$$

   is the *intersection of the lines joining the corresponding vertices of the probability simplex $\mathcal{P}$ and the lower simplex $T^1[b]$*;

3. the relative plausibility of singletons

$$\tilde{pl}_b(x) = \frac{4}{15}, \ \ \tilde{pl}_b(y) = \frac{1}{3}, \ \ \tilde{pl}_b(z) = \frac{2}{5}$$

   is the *intersection of the lines joining the corresponding vertices of $\mathcal{P}$ and upper simplex $T^2[b]$*;

4. finally, the intersection probability

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

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

Point 1. is easily explained by noticing that in the ternary case, by Equation (7), $\mathcal{P}[b] = T^1[b] \cap T^2[b]$. Figure 4 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 not guaranteed to be true in the general case.

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

A counterexample similar to that of the proof of Theorem 1 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 set of probability intervals 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 in the general case too.

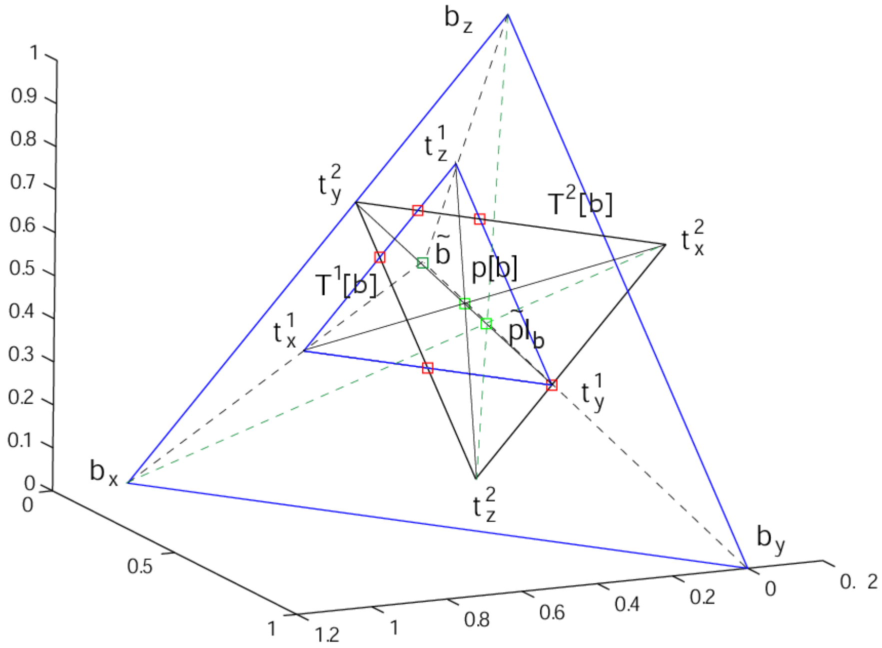

Figure 4: The polytope of all the probabilities consistent with the belief function (22) is shown here in the simplex $\mathcal{P} = Cl(b_x, b_y, b_z)$ of all probability distributions on $\Theta = \{x, y, z\}$. Its vertices (red squares) are given by Equation (23). Intersection probability $p[b]$, relative belief $\tilde{b}$ and plausibility $\tilde{pl}_b$ 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 the lower and upper simplices $T^1[b]$ and $T^2[b]$ are nothing but triangles. Their focus is geometrically the intersection of the lines joining corresponding vertices (dashed lines for $\{\mathcal{P}, T^1[b]\}$ and $\{\mathcal{P}, T^2[b]\}$, solid ones for $\{T^1[b], T^2[b]\}$).

4.2 Focus of a pair of simplices

In the ternary case 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, or the probability simplex. This remark can be formalized through the notion of *focus* of a pair of simplices, laying the foundations for a credal interpretation of these three Bayesian transformations.

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* $S \cap T$ *which has the same affine coordinates* $\{\alpha_1, ...\alpha_n\}$ *in both simplices:*

$$f(S, T) = \sum_{i=1}^{n} \alpha_i s_i = \sum_{i=1}^{n} \alpha_i t_j, \qquad \sum_{i=1}^{n} \alpha_i = 1. \tag{24}$$

Such point always exists, even though it does not always fall into the intersection of the two simplices. In the latter case, though, the focus coincides with the unique intersection of the lines joining corresponding vertices of S and T (see Figure 1 again).
Suppose indeed that a point p is such that

$$p = \alpha s_i + (1 - \alpha) t_i, \qquad \forall\, i = 1, ..., n \tag{25}$$

(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

$$\begin{aligned} p &= \sum_{i=1}^{n} \alpha_i t_i = \frac{1}{1-\alpha} \sum_i \alpha_i \left[p - \alpha s_i\right] \\ &= \tfrac{1}{1-\alpha} \left[p \textstyle\sum_i \alpha_i - \alpha \sum_i \alpha_i s_i\right] \\ &= \frac{1}{1-\alpha} \left[p - \alpha \sum_i \alpha_i s_i\right] \end{aligned}$$

which implies $p = \sum_i \alpha_i s_i$, i.e. p is the focus of (S, T). Notice that the center of mass 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 center of mass of the opposite ($d-1$ dimensional) face (see Figure 5). But those centers of mass for all $d-1$ dimensional faces form themselves the vertices of a simplex T. Therefore, the pignistic function itself can be thought of as the focus of two simplices.

4.3 Bayesian transformations as foci

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

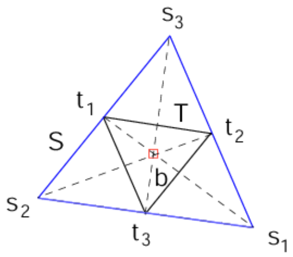

Figure 5: The center of mass b of a simplex S is the focus of the simplex itself and the simplex T formed by the centers of mass of all its $n-1$-dimensional faces. Here a 2-dimensional example is shown.

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 of simplices* $\{\mathcal{P}, T^{n-1}[b]\}$.

The notion of focus of upper and lower simplices provides indeed the desired credal semantics for the family of Bayesian transformations linked to Dempster's rule of combination, in terms of the credal set associated with the related set of probability intervals.
The coordinate of the focus on the intersecting lines also has a meaning in terms of degrees of belief.

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

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.*

An analogous result has recently been proven [4] for the intersection probability (18).

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

As we could have expected, 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 set of probability intervals).

Theorem 6. *The coordinate of the intersection probability as focus of* $\{T^1[b], T^{n-1}[b]\}$ *on the corresponding intersecting lines is the ratio* $\beta[b]$ *(19).*

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

5 Comments and conclusions

The notion of focus of a pair of simplices provides a unifying geometric framework for a number of different Bayesian transformations of belief functions. In fact, as we pointed out here, it is more correct to think of relative belief, plausibility, and intersection probability as transformations/approximations/representatives of lower, upper, and interval probability systems respectively. While $\tilde{b}$, $\tilde{pl}_b$ and $p[b]$ are potentially inconsistent with the original BF, they are perfectly consistent with the associated lower/upper probability systems (as they fall into the corresponding credal set). Therefore we can argue that simply replacing the pignistic transform with a different transformation when operating on BFs in the TBM would not be semantically correct.

The geometric notion of focus has 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*. Notice that the second constraint can be empty, like in the case of upper or lower probability systems. If we assume homogeneous behavior in the two sets of constraints as a rationality principle for a probability transformation, then the above Bayesian functions follow as the necessary unique solutions to the corresponding transformation problems. The notion can be easily extended to more than two constraints.

Finally, the credal interpretation of upper, lower, and interval probability constraints on singletons lays in perspective the foundations of 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 center of mass, the pignistic function $BetP[b]$: $\{\mathcal{P}[b], BetP[b]\}$. We can therefore imagine similar frameworks

$$\left\{\{\mathcal{P}, T^1[b]\}, \tilde{b}\right\}, \quad \left\{\{\mathcal{P}, T^{n-1}[b]\}, \tilde{pl}_b\right\}, \quad \left\{\{T^1[b], T^{n-1}[b]\}, \tilde{pl}_b\right\} \tag{26}$$

in which lower, upper, and interval constraints on a probability distribution on $\mathcal{P}$ are represented by the associated credal sets. This would involve replacing the TBM's disjunctive/conjunctive combination rules [19] by specific evidence elicitation/revision operators for lower, upper, and interval probability systems. Decisions would then be made based on the appropriate focus probability: relative belief, plausibility,

or interval probability respectively.
Notice that, even though in the case of belief functions such systems are simply less informative than the original BF, and their credal sets outer approximations of the credal set of consistent probabilities $\mathcal{P}[b]$, they can be defined independently in their own right, according to the available evidence at hand. In such a case, the use of the appropriate transformation according to the above rationality principle would ensure the consistency of the result. We plan to elaborate on this line of research in the near future.

Appendix: proofs

Proof of Theorem 1. Consider a belief function $b : 2^\Theta \rightarrow [0,1]$, $\Theta = \{x_1, x_2, ..., x_n\}$ such that $m_b(x_i) = k_b/n$, $m_b(\{x_1,x_2\}) = 1 - k_b$. Then

$$b(\{x_1,x_2\}) = 2 \cdot \frac{k_b}{n} + 1 - k_b = 1 - k_b\left(\frac{n-2}{n}\right),$$
$$\tilde{b}(x_1) = \tilde{b}(x_2) = \frac{1}{n} \Rightarrow \tilde{b}(\{x_1,x_2\}) = \frac{2}{n}.$$

For $\tilde{b}$ to be consistent with b it is necessary that $\tilde{b}(\{x_1,x_2\}) \geq b(\{x_1,x_2\})$, i.e.

$$\frac{2}{n} \geq 1 - k_b\frac{n-2}{n} \Rightarrow k_b \geq 1$$

i.e. $k_b = 1$. Therefore if $k_b < 1$ (b is not a probability) its relative belief of singletons is not consistent.

Proof of Theorem 2. We need to prove that $\tilde{b}$ has the same simplicial coordinates in $\mathcal{P}$ and $T^1[b]$. By definition (17) $\tilde{b}$ 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

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

in terms of the vertices $t^1_x[b]$ of $T^1[b]$. Replacing (11) in the above equation yields $\sum_{x\in\Theta} \frac{m_b(x)}{k_b} t^1_x[b] =$

$$\begin{aligned} &= \sum_{x\in\Theta} \frac{m_b(x)}{k_b}\left[\sum_{y\neq x} m_b(y) b_y + \left(1 - \sum_{y\neq x} m_b(y)\right) b_x\right] = \\ &= \sum_{x\in\Theta} b_x\left(\frac{m_b(x)}{k_b}\sum_{y\neq x} m_b(y)\right) + \sum_{x\in\Theta}\frac{m_b(x)}{k_b} b_x + \\ &- \sum_{x\in\Theta} b_x\left(\frac{m_b(x)}{k_b}\sum_{y\neq x} m_b(y)\right) = \sum_{x\in\Theta}\frac{m_b(x)}{k_b} b_x = \tilde{b}. \end{aligned}$$

Proof of Theorem 3. We just need to replace belief with plausibility values in the proof of Theorem 2.

Proof of Theorem 4. 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 (25). The latter assumes the following form (being $s_i = b_x$, $t_i = t^1_x[b]$): $\sum_{x\in\Theta} \frac{m_b(x)}{k_b} b_x =$

$$\begin{aligned} &= t^1_x[b] + \alpha(b_x - t^1_x[b]) = (1-\alpha)t^1_x[b] + \alpha b_x \\ &= (1-\alpha)\left[\sum_{y\neq x} m_b(y) b_y + \big(1 - k_b + m_b(x)\big) b_x\right] + \alpha b_x \\ &= b_x\Big[(1-\alpha)\big(1 - k_b + m_b(x)\big) + \alpha\Big] + \\ &\quad + \sum_{y\neq x} m_b(y)(1-\alpha) b_y, \end{aligned}$$

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

Proof of Theorem 5. Again we can compute the line coordinate α of $\tilde{pl}_b = f(\mathcal{P}, T^{n-1}[b])$ by imposing condition (25). The latter assumes the form (being $s_i = b_x$, $t_i = t^{n-1}_x[b]$): $\sum_{x\in\Theta} \frac{pl_b(x)}{k_{pl_b}} b_x =$

$$\begin{aligned} &= t^{n-1}_x[b] + \alpha(b_x - t^{n-1}_x[b]) = (1-\alpha)t^{n-1}_x[b] + \alpha b_x \\ &= (1-\alpha)\left[\sum_{y\neq x} pl_b(y) b_y + \big(1 - k_{pl_b} + pl_b(x)\big) b_x\right] + \alpha b_x \\ &= b_x\Big[(1-\alpha)\big(1 - k_{pl_b} + pl_b(x)\big) + \alpha\Big] + \\ &\quad + \sum_{y\neq x} pl_b(y)(1-\alpha) b_y, \end{aligned}$$

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

Proof of Theorem 6. Again, we need to impose condition (25) 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 (11), (13) of $t^1_x[b]$ and $t^{n-1}_x[b]$)

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

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)

$$\frac{1-k_b}{k_{pl_b}-k_b}\big[pl_b(x) + (1-k_{pl_b}) - m_b(x) - (1-k_b)\big] + \\ +(1-k_b) + m_b(x) = \beta[b]\big[pl_b(x) - m_b(x)\big] + (1-k_b) \\ +m_b(x) - (1-k_b) = p[b](x)$$

while obviously $m_b(y) + \alpha(pl_b(y) - m_b(y)) = m_b(y) + \beta[b](pl_b(y) - m_b(y)) = p[b](y)$ for all $y \neq x$, no matter the choice of x.

References

[1] M. Bauer, *Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study*, IJAR **17** (1997), 217–237.

[2] A. Chateauneuf and J.Y. Jaffray, *Some characterization of lower probabilities and other monotone capacities through the use of Moebius inversion*, Math. Soc. Sci. **17** (1989), 263–283.

[3] B. Cobb and P.P. Shenoy, *On the plausibility transformation method for translating belief function models to probability models*, IJAR **41** (2006), no. 3, 314–330.

[4] F. Cuzzolin, *Rationale and properties of the intersection probability*, submitted to AI (2007).

[5] F. Cuzzolin, *Two new Bayesian approximations of belief functions based on convex geometry*, IEEE Trans. SMC-B **37** (2007), no. 4, 993–1008.

[6] F. Cuzzolin, *A geometric approach to the theory of evidence*, IEEE Trans. SMC-C **38** (2008), no. 4, 522–534.

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

[8] L. de Campos, J. Huete, and S. Moral, *Probability intervals: a tool for uncertain reasoning*, IJUFKS **1** (1994), 167–196.

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

[10] T. Denoeux, *Inner and outer approximation of belief structures using a hierarchical clustering approach*, IJUFKS **9** (2001), no. 4, 437–460.

[11] T. Denoeux and A. Ben Yaghlane, *Approximating the combination of belief functions using the Fast Moebius Transform in a coarsened frame*, IJAR **31** (2002), no. 1-2, 77–101.

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

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

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

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

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

[17] Ph. Smets, *Belief functions versus probability functions*, Uncertainty and Intelligent Systems, Springer Verlag, Berlin, 1988, pp. 17–24.

[18] Ph. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in Artificial Intelligence, 5, Elsevier Science Publishers, 1990, pp. 29–39.

[19] Ph. Smets, *Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem*, IJAR **9** (1993), 1–35.

[20] Ph. Smets, *Decision making in the TBM: the necessity of the pignistic transformation*, IJAR **38** (2005), no. 2, 133–147.

[21] Ph. Smets and R. Kennes, *The transferable belief model*, AI **66** (1994), 191–234.

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

[23] B. Tessem, *Approximations for efficient computation in the theory of evidence*, AI **61** (1993), no. 2, 315–329.

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

[25] K. Weichselberger and S. Pohlmann, *A methodology for uncertainty in knowledge-based systems*, Lecture Notes in Artificial Intelligence, vol. 419, Springer, Berlin, 1990.

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

[27] A. Ben Yaghlane, T. Denoeux, and K. Mellouli, *Coarsening approximations of belief functions*, ECSQARU'2001, 2001, pp. 362–373.