Two k-additive generalizations of the pignistic transform for imprecise decision making

Thomas Burger

Universit Europenne de Bretagne, Universit de Bretagne-Sud, CNRS, Lab-STICC, Centre de Recherche Yves Coppens, BP 573, F-56017 Vannes CEDEX, FRANCE

Fabio Cuzzolin

Department of Computing of Communication Technologies, Oxford Brookes University, Oxford, U.K.

Abstract

The Transferable Belief approach to the Theory of Evidence is based on the pignistic transform which, mapping belief functions to probability distributions, allows to make "precise" decisions on a set of disjoint hypotheses via classical utility theory. In certain scenarios, however, such as medical diagnosis, the need for an "imprecise" approach to decision making arises, in which sets of possible outcomes are compared. We propose here a framework for imprecise decision derived from the TBM, in which belief functions are mapped to k-additive belief functions (i.e., belief functions whose focal elements have maximal cardinality equal to k) rather than Bayesian ones. We do so by introducing two alternative generalizations of the pignistic transform to the case of k-additive belief functions. The latter has several interesting properties: depending on which properties are deemed the most important, the two distinct generalizations arise. The proposed generalized transforms are empirically validated by applying them to imprecise decision in concrete pattern recognition problems.

Keywords: Theory of evidence, Transferable Belief Model, pignistic transform, k-additive belief functions, generalization, geometric approach

1. Introduction

Decision making is a common issue in applied science, as people or machines need to make inferences about the state of the external world, and take appropriate actions. Such state is typically assumed to be described by a probability distribution over a set of alternative hypotheses, which in turn needs to be inferred from the available data. Sometimes, however, as in the case of extremely rare events (e.g., a volcanic eruption), few statistics are available to drive such inference. Part of the data can be missing. In all practical cases the available evidence can only provide some sort of constraint on the unknown, "true" probability governing the process. Different kinds of constraints are associated with different uncertainty measures or *imprecise probabilities*.

The Theory of Evidence or Dempster-Shafer Theory (DST), [1, 2] is centered on a particular class of imprecise probabilities [3] called *belief functions*. DST allows to distinguish between *imprecision* due to the subjectivity of the point of view, and *randomness* inherent to the phenomenon of interest [4], and is efficient at handling conflicting, not entirely reliable or partial data.

Making decisions based on evidence in the form of a belief function, however, is not trivial, even though a number of approaches have been proposed in the past [5, 6, 7]. On the other hand, decision making is well established and intuitive in a probabilistic context: decisions can be evaluated by assessing their ability to provide a winning strategy on the long run in a game theory context, or maximize return in a utility theory framework. For these reasons probabilistic decision making is the basis of the most popular approach to DST, the *Transferable Belief Model* (TBM) [2]. In the latter, the various pieces of information are combined in the form of belief functions, while the result is eventually converted into a probability distribution to make a decision. Several methods for mapping a belief function to a probability or "probability transforms" [8] have been proposed, the most popular being the original *pignistic transform* [2] proposed in the TBM, which is motivated by the *principle of insufficient reason* [9] (even though its justification is based on more elaborated arguments).

1.1. Motivations

In some practical scenarios, however, *imprecise* decision making is of interest. We call imprecise decision a setting in which several groups of hypotheses of different cardinalities are compared, and one of them is selected. When the selected group is the set gathering all the possible hypotheses the decision is so imprecise that, in practice, no decision is made. Conversely, in the case in which a singleton set is selected, the decision is "precise". Imprecise decisions have important applications in, among others, gesture recognition [10] or handwriting recognition [11].

Email addresses: thomas.burger@univ-ubs.fr (Thomas Burger), fabio.cuzzolin@brookes.ac.uk (Fabio Cuzzolin)

URL: http://www-labsticc.univ-ubs.fr/~burger/ (Thomas Burger), http://cms.brookes.ac.uk/staff/FabioCuzzolin/ (Fabio Cuzzolin)

In this paper we propose an approach to imprecise decision making based on k-additive belief functions, i.e., belief functions whose non-zero mass subsets or "focal elements" have size at most k. In particular, we propose to generalize the pignistic transform to k-additive mass functions, where k is a parameter to tune. The interest of this generalization is manyfold.

In first place, k-additive mass functions are intermediate objects between probabilities and classical belief functions, whose study can shed light on the various links between probability theory and theory of evidence [12, 13]. Secondly, as they limit the cardinality of the focal elements, they can be useful to limit the computational complexity of DST which has been recognized as its most crucial drawback. Finally, and more relevantly to decision making, a k-additive pignistic transform provides an interesting solution to the problem of imprecise decisions, as k-additive belief functions can help tuning the imprecision of the decision thanks their parameter k. We propose two different k-additive generalizations of the pignistic transform, derived one from decision making considerations, the other for a geometric analysis of the set of k-additive belief functions dominating the belief function to map.

1.2. Paper outline

Section 2 recalls the basic notions of DST and its geometric interpretation. Section 3 recalls the basis of the Transferable Belief Model and survey the literature around the pignistic transform: the pignistic level in the TBM, the axiomatic justifications of the pignistic transform, and its geometrical properties. We also summarize the state-of-the-art on the probability transformation problem, including pre-existing generalizations and inverse transformations of the pignistic function. Section 4 motivates our interest in imprecise decision making with the help of some toy examples. As it is recalled in Section 3, the pignistic transform has several interesting properties. Depending on the properties which are deemed most important, two distinct generalizations are introduced. Section 5 presents a first generalization derived from decision making arguments. Section 6 presents instead a generalization derived from geometrical considerations, starting from the behavior of the pignistic transform as center of mass of the set of dominating 1-additive (Bayesian) belief functions. These two generalizations are then investigated from a mathematical and a geometrical point of view, in order to understand their properties and verify that they are admissible generalizations of the pignistic transform (Section 7). Finally, their behavior is compared in real-world pattern recognition problems for which imprecise decisions are relevant (Section 8). Finally, we discuss our results and provide concluding remarks in Section 9, while Appendix A collects some mathematical proofs.

2. Belief functions and their geometric representation

2.1. Dempster-Shafer theory

Let $\Omega = \{\omega_1, ..., \omega_{|\Omega|}\}$ be a finite set, called *frame* or *state-space*, which is made of exclusive and exhaustive hypotheses. A *mass function* m is defined on the power set of Ω, denoted by $\mathcal{P}(\Omega)$, with values in $[0, 1]$, such that $\sum_{A\subseteq\Omega} m(A) = 1$ and $m(\emptyset) = 0$. The value $m(A)$ measures the subjective belief an agent commits to the subset of hypotheses A. As m has the structure of a Choquet's capacity, it is possible to define other functions which are equivalent to m via Möbius inversions [14]. The *belief function* b is defined as:

$$b(A) = \sum_{B\subseteq A} m(B), \qquad \forall A \subseteq \Omega \tag{1}$$

Basically, $b(A)$ is the sum of the masses of all the pieces of evidence which imply A, and corresponds to the lower bound of all the subjective probabilities which are consistent with the given evidence. Dually, the *plausibility function*

$$pl(A) = \sum_{B\cap A\neq\emptyset} m(B), \forall A \subseteq \Omega \tag{2}$$

determines an upper bound for such probability values, and measures the evidence which does not contradict A. Finally, the *commonality function*

$$q(A) = \sum_{A\subseteq B} m(B), \forall A \subseteq \Omega \tag{3}$$

measures the amount of support that a set could potentially receive from its supersets, if the knowledge were more precise. Thus, the commonality function also measures the imprecision of the knowledge encoded by the mass function. Finally, m, b, pl and q are four equivalent representations of the same knowledge, and:

$$\begin{array}{rcl} m(A) & = & \sum_{A\subseteq B} (-1)^{|B|-|A|} q(B), \\ q(A) & = & \sum_{B\subseteq A, B\neq\emptyset} (-1)^{|B|+1} pl(B) \\ m(A) & = & \sum_{B\subseteq A} (-1)^{|A|-|B|} b(B). \end{array} \tag{4}$$

A subset $F \subseteq \Omega$ such that $m(F) > 0$ is called a *focal element* of m. The union of all the focal elements of m is called the *core* of m, and is denoted by $\mathcal{F}(m)$. If the c focal elements of m are nested ($F_1 \subset F_2 \subset \ldots \subset F_c$), m is said to be *consonant*. If all the focal elements are singletons, then, m is said to be *Bayesian*. A Bayesian mass function is in a trivial one-to-one correspondence with a probability distribution. We call *degrees of freedom* of a mass function m the elements of $\mathcal{P}(\Omega)$ that can potentially be focal elements of m. For instance, a Bayesian belief function has $|\Omega|$ degrees of freedom, while a generic belief function has $2^{|\Omega|}$ degrees of freedom, corrsponding to all the elements of $\mathcal{P}(\Omega)$.

Two mass functions $m^{[1]}$ and $m^{[2]}$, based on pieces of evidence provided by two independent and reliable sources can be combined into a new mass function $m^{[\cap]}$ by means of the *conjunctive combination* $\bigcirc\!\!\!\!\cap$, defined as:

$$m^{[\cap]}(A) = \left[m^{[1]} \bigcirc\!\!\!\!\cap\, m^{[2]}\right](A) = \mathcal{K}_{12} \sum_{B\cap C=A} m^{[1]}(B)\, m^{[2]}(C) \tag{5}$$

$\forall A \subseteq \Omega$, where

$$\mathcal{K}_{12} = \frac{1}{1 - \sum_{B\cap C=\emptyset} m^{[1]}(B) \cdot m^{[2]}(C)} \tag{6}$$

measures the conflict between $m^{[1]}$ and $m^{[2]}$. The operator $\textcircled{\cap}$ is symmetric and associative, and thus, it can be extend to a N-ary operator.

The *least commitment principle* [15] postulates that, given a set of mass functions compatible with a number of constraints, the most appropriate one is the *least informative*. As pointed out by Denoeux [16], the principle plays a role similar to that of maximum entropy in probability theory. However, there are many ways of measuring the information content of belief functions, which in turn implies a partial order in their space [17, 18, 19], from the less informative (or less committed) belief functions to the more informative (or more committed) ones. Several such partial orders exist. In this paper, we mainly use the partial ordering called *weak inclusion*, related to the notion of *b-dominance*: a belief function $b^{[2]}$ dominates another one $b^{[1]}$ if the belief values of b' are greater than or equal to those of $b^{[1]}$ for all events $A \subseteq \Omega$

$$b^{[1]} \ll b^{[2]} \;\equiv\; b^{[1]}(A) \leq b^{[2]}(A) \;\; \forall A \subseteq \Omega. \qquad (7)$$

2.2. Geometric approach

As shown in [20, 21, 22], it is possible to interpret belief functions as points of a Cartesian space, and study the interplay of Dempster-Shafer theory objects from a geometrical point of view.

Given a frame Ω, each belief function $b : 2^{\Omega} \to [0,1]$ is completely specified by its $N \doteq 2^{|\Omega|} - 2$ belief values $\{b(A)$, such that $\emptyset \subsetneq A \subsetneq \Omega\}$, (as $b(\emptyset) = 0$, $b(\Omega) = 1 \; \forall b$) and can therefore be represented as a vector of $\mathbb{R}^N$:

$$\vec{b} = [b(A), \emptyset \subsetneq A \subsetneq \Omega]' \qquad (8)$$

If we denote by b_A the *categorical* belief function [2] assigning all the mass to a single subset $A \subseteq \Omega$, $m_{b_A}(A) = 1$, $m_{b_A}(B) = 0$ $\forall B \subseteq \Omega$, $B \neq A$, we can prove that [20, 22] the set of points of $\mathbb{R}^N$ which correspond to a belief function, called the *belief space* and noted $\mathcal{B}$, coincides with the convex closure[1] Cl of all the vectors representing categorical belief functions: $\mathcal{B} = Cl(\vec{b}_A, \emptyset \subsetneq A \subseteq \Omega)$. The belief space $\mathcal{B}$ is a simplex [22], and each vector $\vec{b} \in \mathcal{B}$ representing a belief function b can be written as a convex sum as:

$$\vec{b} = \sum_{\emptyset \subsetneq A \subseteq \Omega} m_b(A) \vec{b}_A. \qquad (9)$$

The set $\mathcal{B}_1$ of all Bayesian belief functions on Ω is the simplex determined by all basis belief functions associated with singletons: $\mathcal{B}_1 = Cl(b_\omega, \omega \in \Omega)$.

2.3. k-additive belief functions

Definition 1. *A **k-additive belief function** b on Ω (with $k \in \mathbb{N}$ and $k \leq |\Omega|$) is a belief function such that its mass function m has at least one focal element of cardinality k and none of cardinality $> k$.*

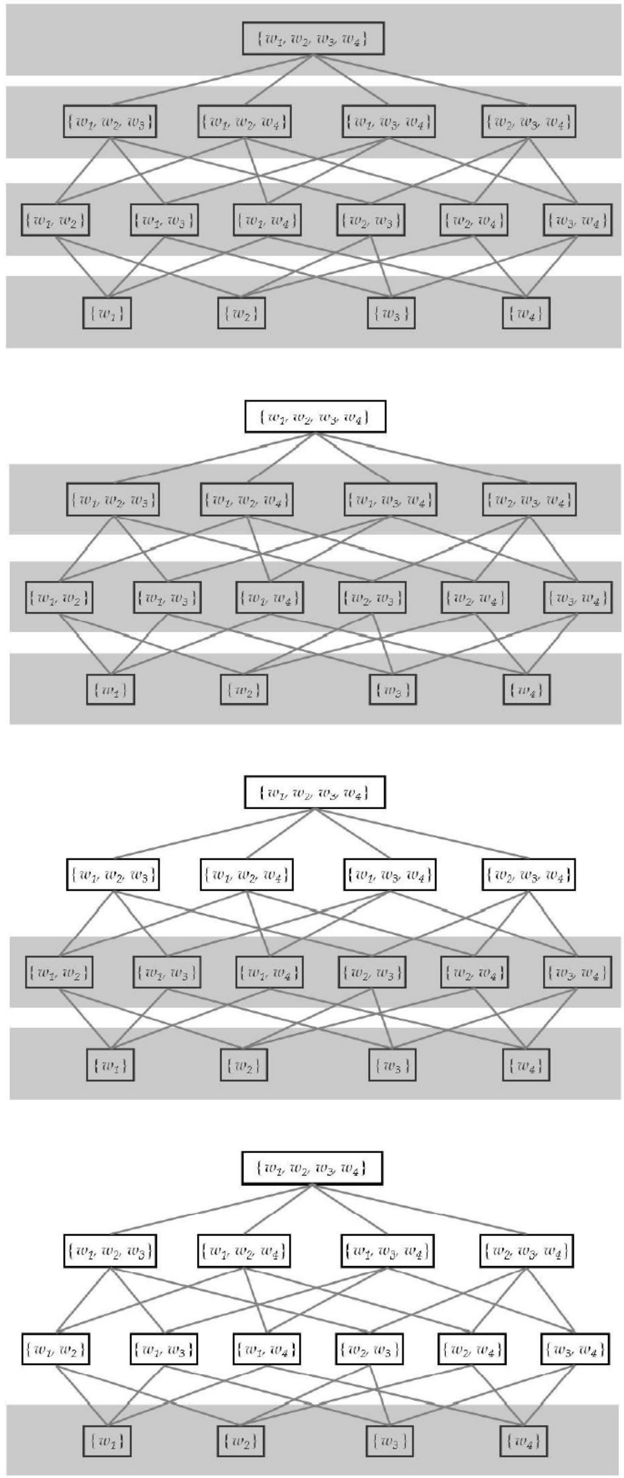

Figure 1: The power set of a frame of 4 elements, represented by a lattice, and the degrees of freedom of several belief functions (in grey). The first figure the case of a classical belief function. Then, the cases of a 3-additive belief function and a 2-additive belief function. Finally, a probability distribution (or Bayesian belief function, or 1-additive belief function.)

According to such a definition, a ℓ-additive belief function with $\ell < k$ is strictly speaking not k-additive. Hence, for sake of convenience, we define a wider class of belief functions as follows:

[1] $Cl(\vec{b}_1, ..., \vec{b}_k) = \{\vec{b} \in \mathcal{B} : \vec{b} = \alpha_1 \vec{b}_1 + \cdots + \alpha_k \vec{b}_k, \sum_i \alpha_i = 1, \; \alpha_i \geq 0 \; \forall i\}$.

Definition 2. *An **at most** k**- additive belief function** b on* Ω *(with* $k \in \mathbb{N}$ *and* $k \leq |\Omega|$*) is a belief function such that its mass function m has no focal element of cardinality* $> k$.

A ℓ-additive belief function with $\ell < k$ is not k-additive, but is at most k- additive. More generally, any belief function on Ω is at most $|\Omega|$-additive. This will simplify our notations in the sequel. Clearly:

1. a probability distribution over Ω, or equivalently, a Bayesian belief function, is a 1-additive belief function, (see Figure 1);
2. for any belief function b defined over Ω, there is a $k \in \{1, ..., |\Omega|\}$, such that b is k-additive.

We denote by $\mathcal{B}_k$ the region of the belief space $\mathcal{B}$ associated with k-additive belief functions. The set of degrees of freedom of a k-additive belief function is denoted by: $\mathcal{P}_k(\Omega) = \{A \subseteq \Omega, |A| \leq k\}$.

k-additive belief functions were originally introduced in [23]. Their interest is manyfold. Firstly, the number of degrees of freedom of a k-additive belief function is:

$$|\mathcal{P}_k(\Omega)| = \sum_{i=1}^{i=k} \binom{|\Omega|}{i}$$

whereas the number of degrees of freedom of a generic belief function. is:

$$|\mathcal{P}(\Omega)| = \underbrace{\sum_{i=1}^{i=\Omega} \binom{|\Omega|}{i}}_{2^{|\Omega|}} = |\mathcal{P}_k(\Omega)| + \underbrace{\sum_{i=k+1}^{i=|\Omega|} \binom{|\Omega|}{i}}_{>0 \text{ if } |\Omega|>k}.$$

Hence, for a fixed cardinality $|\Omega|$ of the frame, the number of degrees of freedom of a k-additive belief function is smaller, making the latter a more compact representation of knowledge. Thus, from a computational point of view, the use of k-additive representations is of real interest, as computation cost is one of the major drawbacks of the theory of belief functions most frequently criticized in the literature.

As mentioned in [23], another advantage of k-additive belief functions is that they are easier to handle from a perceptive point of view. Humans find it rather difficult to attach meaning to focal elements of larger cardinality. Hence, limiting the focal elements to subsets of Ω with bounded cardinality is a sensible way of ensuring that the corresponding mathematical representation of knowledge is intuitive and easy to handle.

Finally, as k-additive belief functions are somehow intermediate objects between belief functions and probabilities (in terms of their degrees of freedom), they constitute an interesting trade-off between the full expressive power of belief functions and the simplicity of interpretation of probability measures (as 1-additive belief functions).

2.4. Geometry of k-additive belief functions

Definition 3. *The **set of at most** k**-additive belief functions dominating a belief function** b* : $2^\Omega \to [0, 1]$ *is defined as* $\mathfrak{B}^k[b] = \{b' \in \mathcal{P}_k(\Omega) : b \ll b'\}$*, or equivalently:*

$$\mathfrak{B}^k[b] = \{b' \in \mathcal{P}_k(\Omega) : b(A) \leq b'(A)\ \forall A \subseteq \Omega\}. \quad (10)$$

Note that, strictly speaking, the space $\mathcal{B}_k$ of k-additive belief functions is not the same as the set of all at most k-additive belief functions. However, for each $\ell < k$ the set of ℓ-additive belief functions is a lower-dimensional face of the simplex associated with the set of at most k-additive belief functions. In other words, $\mathcal{B}_k$ and $\{\mathcal{B}_\ell\}_{\ell \leq k}$ have the same "volume".

As it has been proven in [24, 14], that $\mathfrak{B}^1[b]$, the set of dominating probabilities, is a polytope (or convex polyhedron) in the belief space, whose vertices are probabilities determined by permutations of the elements of Ω.

Proposition 1. *The set* $\mathfrak{B}^1[b]$ *of all the probability functions consistent with a belief function b (of mass m) is the polytope*

$$\mathfrak{B}^1[b] = Cl(p^\rho[b]\ \forall \rho),$$

where Cl(.) denotes the convex closure operator and where ρ *is any permutation* $\{\omega_{\rho(1)}, ..., \omega_{\rho(|\Omega|)}\}$ *of the singletons of* Ω*, and the vertex* $p^\rho[b]$ *is the Bayesian belief function such that*

$$p^\rho[b](\omega_{\rho(i)}) = \sum_{A \ni \omega_\rho(i);\ A \not\ni \omega_\rho(j)\ \forall j<i} m(A). \quad (11)$$

Each probability function (11) attributes to each singletons $\omega = \omega_{\rho(i)}$ the mass of all focal elements of b which contains it, but does not contain the elements which precede ω in the ordered list $\{\omega_{\rho(1)}, ..., \omega_{\rho(n)}\}$ generated by the permutation ρ.

In [25], the authors consider the dominance properties of k-additive belief functions for any type of capacities [3], define $\mathfrak{B}^k[b]$ as the polytope of k-additive belief functions dominating another belief function, and provide some results to characterize it. In this paper on imprecise decisions, we need to consider similar problems, but we focus on dominance with respect to belief functions rather than general capacities, and on the determination of the barycenter of $\mathfrak{B}^k[b]$ as a feasible generalization of the pignistic transform to k-additive belief functions, in an effort to extend the Transferable Belief Model to such computationally and semantically attractive objects, such as first proposed in [26]. But first, we will recall the bases of the TBM.

3. The TBM and the pignistic level

The *Transferable Belief Model* [2] is an interpretation of Shafer's original formulation of the Theory of Evidence [1] motivated by decision making rationales. It is characterized by the following main features.

First, it proposes an axiomatic definition of mass functions independently from any probabilistic setting, whereas Dempster's original definition [27] is based on a probabilistic view of the world, and whereas the alternative Shenoy-Shafer architecture [28, 29] description is centered on non-empty random sets (a set generalization of the notion of random variables in probability theory [30]). The bottom line of the TBM is that probabilities are suitable to model frequentist knowledge, i.e. knowledge derived from repeated statistical observations. Thus, probabilities correspond to long-run frequencies. Other forms of knowledge, instead, should not be described by subjective probabilities (such as those proposed by De Finetti [31], Cox and Jaynes [32], or part of the Bayesian community), as they do not provide a rich enough description. On the contrary, belief functions are adapted to quantized such subjective knowledge. However, for decision making, a probabilistic setting is appropriate as the decision maker is interested in betting on the outcome which is indeed the most frequent: A long-term winning strategy is sought.

This leads us to the second significant characteristic of the TBM, which is divided into two levels. In the first one, called the *credal level*, pieces of evidences are assessed, modified, combined, and so on, until our knowledge state is modeled by a single belief function. Then, decisions are made at the *pignistic level* using the knowledge encoded at the credal level: first the pignistic transform is applied in order to convert the resulting belief function into a probability distribution, called the *pignistic probability*. Then, a MAP (maximum a posteriori) decision is made based on such pignistic probability to derive the most appropriate element of Ω.

3.1. Pignistic transform

Definition 4. *Given a mass function $m : 2^{\Omega} \to [0,1]$, its **pignistic probability** $BetP : \Omega \to [0,1]$ is defined as:*

$$BetP(\omega_i) = \sum_{A \ni \omega_i} \frac{m(A)}{|A|} \qquad \forall \omega_i \in \Omega \tag{12}$$

where $|A|$ is the cardinality of the subset $A \subseteq \Omega$. The corresponding MAP decision ω_* is:

$$\omega_* = \arg\max_{\omega_i} [BetP(\omega_i)].$$

$BetP$ can be considered as a Bayesian mass function, denoted by $m_1^{[S]}$ (in order to stress the fact that it is a 1-additive belief function), whose belief values correspond to the Shapley values (see Equation 13) [33]:

$$m_1^{[S]}(\omega_i) = BetP(\omega_i), \quad m_1^{[S]}(A) = 0 \quad \text{if } |A| > 1.$$

Property 1. *The belief function $b_1^{[S]}$ which corresponds to the mass function $m_1^{[S]}$ is equal to the Shapley value $\mathcal{V}$ [34]:*

$$b_1^{[S]}(B) = \sum_{\{\omega_i\} \in B} m^{[S]}(\{\omega_i\}) = \sum_{A \subseteq \Omega} \frac{m(A) \cdot |A \cap B|}{|A|} = \mathcal{V}(B) \; \forall B \subseteq \Omega. \tag{13}$$

The Shapley value is a capacity used to determine how to share the profit (or loss) resulting on a bet booked by a coalition of gamblers [33]. We need to be careful, however, with the interpretation of $BetP$ as a belief function $m_1^{[S]}$, as the pignistic transform (12) does not commute with Dempster's rule (5):

Property 2. *The pignistic transform does not compute with Dempster's rule.*

3.2. Axiomatic justification of the pignistic transform

While the pignistic level's probabilistic setting for decision making was fully justified, in Smets' view, by the notion of long-run winning strategy, the nature of the mapping between belief functions in the credal level and probability distributions in the pignistic level was originally based on the **Principle of Insufficient Reason** (PIR) proposed by Bernoulli, Laplace, and Keynes [9].

Definition 5. *The **Principle of Insufficient Reason** states that "if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability".*

As understood in a probabilistic understanding of DST, the mass $m(A)$ associated with a non-singleton event $A \subseteq \Omega$ can be understood as a "floating probability mass" which can not be attached to any particular singleton event $\omega_i \in A, i \leq |\Omega|$ because of the lack of precision of the (multi-valued) operator that quantify our knowledge via the mass function. Then, according to the PIR, when considering the restriction of the mass function to the frame induced by the event A, it is wise to assume equiprobability amongst the singleton events $\omega_i \in A, \forall i \leq |\Omega|$. This yields the pignistic transform (12).

Later on, however, Smets [35, 36, 6, 34] advocated that the PIR could not justify by itself the uniqueness of the pignistic transform, and proposed a justification based on a number of axioms.

Definition 6. *The five rationality arguments which justify the existence and unicity of the pignsitic transform are:*

1. ***Linearity**: The pignistic transform must commute with the convex closure operator.*
2. ***Projectivity**: The pignistic tranform of a Bayesian belief function is (equivalent to) the belief function itself.*
3. ***Efficiency**: A bet on the whole frame is bound to win with a certain probability.*
4. ***Anonymity**: The result of the pignistic tranform does not depend on permutations of the elements of Ω.*
5. ***False event**: Any false event is bounded to have a zero pignistic probability.*

In [35] Smets refuted a Dutch Book which had been proposed against the pignistic transform [37], [38]. A Dutch Book is a scenario of bets and corresponding offered odds which guarantees a profit (or a loss) regardless the results on the experiments. Of course, the existence of such a Dutch Book in the TBM framework (whereas for the same scenario it does not exist in the probabilistic framework) would have been a proof of the lack of coherence of the pignistic transform. Even though this Dutch Book objection was indeed discarded, Smets admitted that a proof that his pignistic transform resists to all Diachronic Dutch Books in general had not been given [35].

The commutativity of the pignistic transform with respect to convex closure ("linearity") is of great interest. Conversely, Property 2, stating that the pignistic transform does not compute with Dempster's rule, is a major arguments against the pignistic transform: In [39], the commutation with the Dempster's rule is said to be far more important than the linearity argument. Thus, another set of axioms is proposed to defined the behavior of a transformation of a belief function onto a probability [39]. Of course, these two sets of axioms are not compatible and lead to different results. But, as pointed out by Shafer [32], the word "axiom" has two different meaning: It can either be an assertion, the truth of which cannot be established nor denied, as in Euclid's geometry, or it can be an assumption the consequences of which are later explored. This may explain the ongoing discussions around the axiomatic justification of probability transformations.

The original intuition driven by the PIR, however, retains its interest as confirmed by the following geometrical property of the pignistic transform.

Property 3. *Let b be a belief function, and let us consider the set $\mathfrak{B}^1[b]$ of probabilities dominating b. The pignistic transform $m_1^{[S]}$ of b is the center of mass of $\mathfrak{B}^1[b]$ in the belief space.*

A complete proof of this well known result [14, 40, 24] being hard to find, we presented one in Appendix A.1. This version of the demonstration is interesting, as it is based on techniques that will be required in the analysis of the barycenter of the set of dominating k-additive belief functions.

3.3. Other probability transformations and generalizations of the pignistic transform

Other probability transforms, not necessarily in relation to the TBM, have been proposed along the years. Here is a short review of them:

The only transform which is as popular as the pignistic probability is defined by normalizing the plausibility values of singletons in order for them to sum up to 1 over Ω:

$$m^{[RelPl]}(\omega_i) = \frac{pl(\omega_i)}{\sum_{\omega_j\in\Omega} pl(\omega_j)} = \frac{pl(\omega_i)}{\sum_{A\subseteq\Omega} m(A)\cdot|A|} \qquad \forall\omega_i\in\Omega. \tag{14}$$

This transform is really interesting as it is both a computationally efficient approximation of a belief function and useful for decision making. It was first introduced by Voorbraak in 1989 [41] as a normalization of the commonality values of the singletons,but its use for decision making came earlier as the problem of finding *the most plausible configuration* was previously addressed in [28, 42]. In spite of its lack of dominating properties, this transform has been studied by various authors, and possessed as a consequence a number of different names in the literature, such as: *Bayesian approximation* [41], *proportional plausibility probability* [43], *plausibility transform* [39], *cautious probabilistic transform* [8], *relative plausibility of singleton* [44, 45, 46].

Similarly, it is possible to derive another transform by normalizing the belief of singletons. As long as $\exists\omega_i\in\Omega$ such that $m(\omega_i)\neq 0$, it is defined by:

$$m^{[RelBel]}(\omega_i) = \frac{bel(\omega_i)}{\sum_{\omega_j\in\Omega} bel(\omega_j)} \qquad \forall\omega_i\in\Omega.$$

Using this transformation amounts to dropping the focal elements with cardinality greater than or equal to 2. Its result has several interesting properties in terms of both decision making and computational complexity. In [43] (2001), Sudano briefly introduced it as the *proportional belief probability*. Then, it was introduced and extensively studied [8] (2006) by Daniel as the *disjunctive probabilistic transform*. In the latter, Daniel also briefly discussed the interactions of several transforms with belief and plausibility functions from a geometric point of view. Cuzzolin extensively studied semantics and properties of this transform, that he called the *relative belief of singletons* [47, 48], proving in particular that (in analogy to the plausibility transform (14)) it commutes with Dempster's sum (of plausibility functions), and it perfectly represents plausibility functions when combined with Bayesian belief functions.

In a similar way, Cuzzolin derived from geometric considerations the *orthogonal projection* and the *intersection probability* [44]. The orthogonal projection

$$\pi[b](\omega_i) = \sum_{A\supseteq\{\omega_i\}} m(A)\Big(\frac{1+|A^c|2^{1-|A|}}{|\Omega|}\Big) + \sum_{A\not\supseteq\{\omega_i\}} m(A)\Big(\frac{1-|A|2^{1-|A|}}{|\Omega|}\Big) \tag{15}$$

is particularly interesting as it meets Smets' linearity axiom (Axiom 1), showing that the pignistic transform is only one of a family of transformations commuting with convex or affine combination, which Cuzzolin called *affine family* [46].

The *probability deficiency proportional plausibilities*, was very briefly introduced by Sudano [43] and shares important similarities with Cuzzolin's intersection probability. Both have the following structure:

$$bel(\omega_i) + \Big[1 - \sum_{\omega_j\in\Omega} bel(\omega_j)\Big]\cdot\frac{\Delta(\omega_i)}{\sum_{\omega_j\in\Omega}\Delta(\omega_j)}, \quad \forall\omega_i\in\Omega_X$$

with $\Delta = pl$ for Sudano's transform, and $\Delta = pl - m$ for Cuzzolin's.

Daniel [8] and Sudano [43] proposed several other transforms [49]. However, to our knowledge (apart from a few preliminary works of ours [50, 26]), there are only three generalizations of the pignistic transform, none of them generalizing to the k-additive case so interesting for imprecise decision.

The first one, presented by Daniel [8], is the *weighted pignistic transform*. The idea is to use a mass function $m^{[W]}$ to weight the pignistic transform according to some prior knowledge (encoded in $m^{[W]}$). $\forall\omega_i\in\Omega$:

$$m_X^{[WS]}(\omega_i) = \sum_{\substack{\omega_i\in A\\ A\subseteq\Omega}} \frac{m^{[W]}(\omega_i)}{\sum_{B\in A} m^{[W]}(B)}\frac{m(A)}{|A|}.$$

Baroni's [51] considers the problem of defining a pignistic transform which is not only valid for belief functions, but also for other types of imprecise probabilities, such as 2-monotone capacities. Finally, Dezert's *generalized pignistic transformation* [52] refers in fact to a complete different framework (the Dezert-Smarandache Theory of Plausible and Paradoxical Reasoning [53]), which has the same power of expression as the Dempster-Shafer theory [54].

4. Imprecise decisions via *k*-additive generalizations

As we have seen, a significant amount of work has been conducted around the pignistic transform. Nevertheless, to our knowledge, no k-additive generalization has so far been proposed, notwithstanding the interest such generalization would bear in the context of imprecise decision.

Here, as the prime goal of the pignistic transform is to provide an efficient framework for decision making, we first explain how k-additive belief functions can actually be used to implement imprecise decisions, and we stress the interest of imprecise decisions on different examples. Later we focus on alternative ways of deriving such a generalization.

4.1. The interest of imprecise decisions

Classically, any decision setting is made of a finite number of choices (or hypotheses), fitted with several estimators such as posterior probabilities, profits or losses. A decision maker promotes a single choice and discard the others, according to a specific strategy.

This classical framework is not always able to describe the decision making behavior of a human expert, as a human being is able, when necessary, to promote an entire set of disjoint hypotheses. Hence, as illustrated in [55], during a medical diagnosis, physicians need to discriminate all the potential diseases of a patient but the single one which really is the source of the illness. However, if symptoms do not concur, or if the disease is at an early stage, the physician is only capable of discarding a huge proportion of all the possible diseases, while still considering a few of them. In such a context, a decision making process can be seen as a *narrowing* of the set of hypotheses, rather than the selection of a specific one. If in this narrowing process all the hypotheses are discarded but one, we say that the decision is *precise* (or classical, or hard). If several hypotheses remain, the decision is *imprecise*. We call *the cardinality of the decision* the number of hypotheses that are not discarded.

Naturally, the cardinality of the imprecise decision should strongly depend on the situation and on the amount of knowledge, and should not depend on a predefined strategy (such as, for instance, keeping the N best hypotheses). This is well illustrated by the diagnosis example: The physician ought to be as as precise as possible, while preferring imprecision over error. The trade-off between precision and error means that, after having considered all the symptoms, the physician will have to decide between options such as "flu", "flu or bronchitis", "bronchitis or cancer or pneumothorax", etc. Finaly, the decision maker is led to compare all groups of hypotheses regardless to their cardinality, and choose one of these groups, the cardinality of which will only depend on the imprecision of the knowledge of the decision maker, and not on a predefined strategy.

When implementing such behavior two major difficulties occur. On one side we need to guarantee that the decision will not be too imprecise (what a patient would think about an physician who hesitates among two hundred different diagnoses?). To do so, we need to discard sets of hypotheses whose cardinality is above a maximum acceptable value k. Hence, k-additive belief functions seem to be a suitable tool for such a modeling. On the other hand, we need to provide a methodology to compare the respective significance of subsets of different cardinalities, so that the agent will not always promote the most precise or, on the contrary, the most imprecise decision.

To face this issue, ad-hoc methods are typically used. For instance, it is possible to consider compound hypotheses and practice hypothesis testing as in classical statistical theory. In such approaches the size of the selected compound hypothesis is related to the p-value which is expected (the probabaility of the null hypothesis). In a more subjective setting, it is possible to associate a cost with each decision and minimize such as cost function on the power set of Ω. Finally, it is possible to simply sort the individual decision outcomes in descending order, and select the first N hypotheses so that their total probability value is above a certain pre-defined threshold, such as it is classically done in handwritting recognition [56].

The goal of this paper is to provide an imprecise decision approach based on the philosophy of Dempster-Shafer theory (in its TBM interpretation) which fulfils the requirements for such kind of decision making scenarios, i.e., that allow for the comparison of the relative interest of sets of hypotheses of different cardinalities. Practically, the example below illustrates how imprecise decision can be cast in the TBM framework.

4.2. Example: Imprecise trajectories

Let us consider the position of a robot in a state-space so that its behavior is Markovian in time: the location of the robot at time $t + 1$ is independent from its past trajectory *conditionally to the present location* at time t [57].

The robot's future location only depends on (1) the current location, (2) an unknown random process that eventually drives the global motion of the robot, and (3) some bias due to noise or hidden variables. All the possible trajectories along a discrete time scale can be modeled by a lattice representing the cross product of (discrete) state-space and time. A particular trajectory is just a path in this lattice (Figure 2). Suppose that, at each time instant t, sensors provide distinct pieces of information to the robot which are processed in the TBM framework: they are merged (at the credal level) and the current state of the robot is inferred by a decision process that occurs at the pignistic level. Several stances are possible:

1. The classical pignistic transform is used. Unfortunately, as the sensors are error-prone, the inferred state is not always the right one, and the estimated trajectory is composed of correct and incorrect states with respect to the ground-truth (Fig. 2). Of course, the TBM provides tools to filter such trajectories [58, 59, 60] and, in spite of an additional computational cost, they are really efficient.

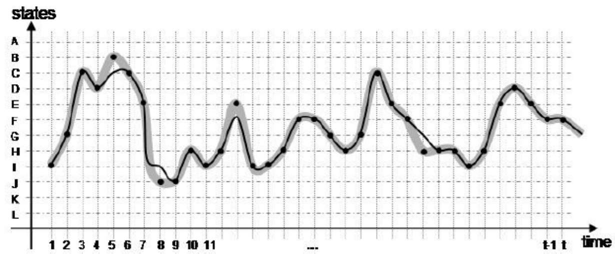

Figure 2: The space-time lattice: the horizontal axis represents the time iterations, and the vertical axis, the states. The real trajectory (ground truth) is represented by the black line, and the inferred states are presented by black dots linked by the grey line. The real and inferred trajectories differ, as few mistakes are made in the decision process.

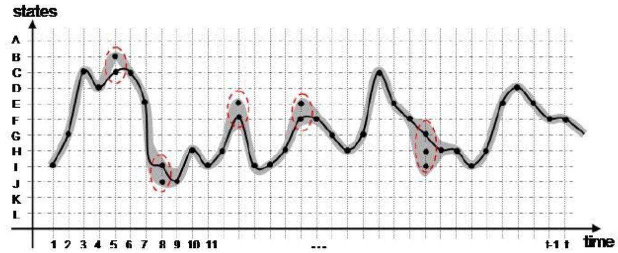

Figure 4: As in Figure 2, the real trajectory (ground truth) is compared to the inferred one. A trade-off between risky bets (a singleton state is assessed) and imprecise decisions (circled by a dot line) allows limiting the number of mistake while remaining quite precise.

2. Betting on compound hypotheses (knowing that, the more numerous they are, the smaller the chance of making a mistake) is safer, but the risk is that no real decision is eventually made and the inferred trajectory is too imprecise (Fig. 3).
3. A tradeoff between these two extreme stances is to automatically tune the cardinality of the decision: When the decision is difficult to make a compound hypothesis is selected to avoid mistakes, while otherwise a singleton hypothesis is awarded to ensure accuracy (Fig. 4).

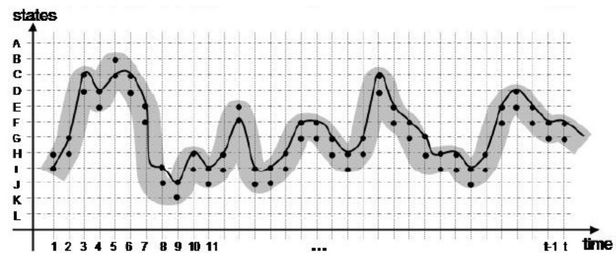

Figure 3: In a similar manner to figure 2, the real trajectory (ground truth) is compared to the inferred one. As a matter of fact, no mistake is made on the inferred trajectory, but, as a drawback, it is really imprecise.

The first stance corresponds to classical decision making. The second stance allows for imprecise decisions: such a decision process can indeed be useful, and as we have mentioned several manners of implementing it exist, in both belief and probability formalisms. The final option corresponds to situations where it is possible to bet on compound hypotheses, but in a different manner, as the cardinality of the decision *is not fixed*. Figures 2, 3 and 4 illustrate the kind of path obtained in the three scenarios. Clearly the last approach is rather interesting, as it proposes a trade-off between minimizing the risk of making an error and supporting a decision as precise as possible. This latter precisely corresponds to the type of decision one expect to define by generalizing the pignistic transform.

4.3. Two possible generalizations

From the discussion of Section 3.2 it emerges that there are at least two ways of generalizing the pignistic transform to k-additive belief functions. Of course additional generalizations could in principle exist, but a strong case can be made in favor of the following ones.

The first option is to generalize the manner in which the pignistic transform redistributes the mass of each focal element of the frame to atoms of Ω. Once understood why this leads to a probability distribution suitable for decision making, we can extend it to situations in which the mass is redistributed from subsets of cardinality $> k$ to subsets of cardinality $\leq k$. In the rest of the paper, we will call this generalization *asymmetric*, as it does not consider all the vertice of $\mathfrak{B}^k[b]$ with the same weight.

Another option is to consider the geometrical feature of the pignistic transform as the barycenter of the polytope of dominating probabilities and to define the barycenter of the polytope of dominating k-additive belief functions, whatever the value of k. In the sequel we will call this the *geometric* generalization of the pignistic transform.

These two different generalizations lead to different outcomes, and have advantages and drawbacks.

5. The asymmetric generalization

As we have seen in Section 3.2, besides the proposed axiomatic justification, the Principle of Insufficient Reason plays an important role in the justification of the pignistic transform. The latter is a redistribution process of mass assignments in which the mass $m(A)$ associated with each $A \subseteq \Omega$ is equally shared by its singleton elements $\omega_i \in \Omega \cap A$. Thus, as stated by the PIR, $m(A)$ is divided into $|A|$ parts of which each $\omega_i \in \Omega \cap A$ receives one:

$$m(\omega_i) \leftarrow \frac{m(A)}{|A|} \quad \forall \omega_i \in \Omega.$$

Any generalization to k-additive belief functions will have to redistribute the mass of subsets of cardinality $> k$ to subsets of cardinality $\leq k$. As a case study, let us consider A such that $|A| > k$, and B and C two subsets of A of cardinality $\leq k$, such that:

$$|B| = n \times |C|$$

with $n \in \mathbb{N}$ and $n < k$ (for instance, C is a singleton, whereas $|B| = n$). In absence of any additional knowledge, the PIR states that there is no reason to privilege any of the singletons in B

or C with respect to the others. Then, it is natural to expect that, in a generalized redistribution process, B receives a mass from A which is n times the mass that C receives from A, as all the singletons they are composed of are equally believed in. In other words, the mass $m(A)$ has to be divided into $\mathcal{N}$ parts, where

$$\mathcal{N} = \sum_{j=1}^{k} \binom{|A|}{j} \cdot j = \sum_{j=1}^{k} \frac{|A|!}{(j-1)!(|A|-j)!}$$

is the total cardinality of the subsets of A of size not greater than k.

Finally, each subset D of A with a cardinality $\leq k$ will receive exactly $|D|$ parts from A. For instance, if $A = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, and the mass of A is redistributed in a 2-additive pignistic transform, $m(A)$ is divided into $\mathcal{N} = 4 + 2 \times 6 = 16$ parts, and any of the 4 singletons will inherits 1 parts, whereas the 6 subsets of 2 elements will receive 2 parts each (i.e. a share proportional to their cardinality).

Formally, let us consider a *hesitation threshold* of value $k \leq |\Omega|$ defined according to the problem under consideration, and a mass function m.

Definition 7. *Given a belief function b with mass $m : 2^\Omega \to [0,1]$, the result of the* ***asymmetric*** *transform, denoted by $m_k^{[\mathcal{A}]}$, is defined as the k-additive belief function induced by the mass assignment*

$$m_k^{[\mathcal{A}]}(B) = m(B) + \sum_{A \supset B,\, A \subseteq \Omega,\, |A|>k} \frac{m(A) \cdot |B|}{\mathcal{N}(|A|, k)} \tag{16}$$

$\forall B \subseteq \Omega$ such that $|B| \leq k$, $m_k^{[\mathcal{A}]}(B) = 0$ $\forall B \subseteq \Omega$ such that $|B| > k$, where

$$\mathcal{N}(|A|, k) = \sum_{\ell=1}^{k} \binom{|A|}{\ell} \cdot \ell = \sum_{\ell=1}^{k} \frac{|A|!}{(\ell-1)!(|A|-\ell)!}$$

is the average cardinality of the subsets of A of size at most k.

This transformation was first empirically introduced, without any theoretical justification, as a means to perform classification into imprecise clusters for an American Sign Language recognition task in videos [10, 61]. The good classification rates demonstrated there support the interest of imprecise decision making in pattern recognition, as well as the validity of the asymmetric k-additive pignistic transform.

6. The geometric generalization

A different generalization of the pignistic transform to k-additive belief functions emerges when we focus on its geometric property of being the barycenter of the set of dominating probabilities. Property 3 states that the pignistic transform of b corresponds to $\overline{\mathfrak{B}^1[b]}$, the barycenter of $\mathfrak{B}^1[b]$. Quite obviously, the asymmetric generalization does not meet with this property. Indeed, the redistribution process which defines it in not "fair", as subsets with greater cardinality receives a larger share of mass than subsets with lower cardinality. On the other hand, the center of mass of a polytope in the belief space is not defined according to any weighting, including the cardinality of A. No additional support is given to subsets of greater cardinality with respect to lower cardinality ones. Thus, a k-additive belief function meeting the trivial generalization of Property 3 would constitute a different generalization of the pignistic transform. Then, let us therefore characterize the barycenter $\overline{\mathfrak{B}^k[b]}$ of the polytope $\mathfrak{B}^k[b]$.

6.1. The polytope of dominating belief functions

Proposition 1 states that the polytope of dominating probabilities (1-additive belief functions) $\mathfrak{B}^1[b]$ has vertices associated with permutations of the list of element of Ω. This suggests that the set of dominating k-additive belief functions could have a similar form, with each vertex associated with a permutation of the list of focal elements of size *smaller than or equal to k*.

Conjecture 1. *Given a belief function $b : \mathcal{P}(\Omega) \to [0,1]$, with mass function m, the region $\mathfrak{B}^k[b]$ of all the k-additive belief functions on Ω which dominate b according to order relation (7) is the polytope:*

$$\mathfrak{B}^k[b] = Cl(b^\rho[b]\ \forall \rho),$$

where ρ is any permutation $\{A_{\rho(1)}, ..., A_{\rho(|\mathcal{P}^k(\Omega)|)}\}$ of the focal elements of Ω of size at most k ($\mathcal{P}^k(\Omega)$), and the vertex $b^\rho[b]$ is the k-additive belief function with the following mass function:

$$m^\rho[b](A_{\rho(i)}) = \sum_{B \supseteq A_{\rho(i)};\ B \not\supset A_{\rho(j)}\ \forall j<i} m(A). \tag{17}$$

Moreover, each actual vertex $b^\rho[b]$ of $\mathfrak{B}^k[b]$ is associated with the same number of permutations of $\mathcal{P}^k(\Omega)$.

This allows us to deal with the computation of the center of mass of $\mathfrak{B}^k[b]$ in a straightforward manner.

Theorem 1. *If Conjecture 1 holds, given a belief function $b : \mathcal{P}(\Omega) \to [0,1]$ of mass function m, the center of mass $\overline{\mathfrak{B}^k[b]}$ of the simplex $\mathfrak{B}^k[b]$ of k-additive belief functions dominating b is given by the mass assignment:*

$$m_k^{[\mathcal{B}]}(A) = \begin{cases} \displaystyle\sum_{B \supseteq A} \frac{m(B)}{|\mathcal{P}_k(B)|}, & \forall A \in \mathcal{P}_k(\Omega) \\ 0 & A \notin \mathcal{P}_k(\Omega), \end{cases} \tag{18}$$

where $|\mathcal{P}_k(B)|$ is the number of subsets of B of size not greater than k.

Note that $|\mathcal{P}_k(B)| = \sum_{\ell=1}^{k} \binom{|B|}{\ell} \neq \mathcal{N}(|B|, k) = \sum_{\ell=1}^{k} \binom{|B|}{\ell} \cdot \ell$. The proof can be found in Appendix A.3. It follows the classical one given for the barycenter $\overline{\mathfrak{B}^1[b]} = BetP[b]$ of the set of dominating Bayesian belief functions (Appendix A.1).

6.2. A generalization of the pignistic transform induced by the barycenter of dominating k-additive belief functions

As expected, for $k = 1$, the expression (18) reduces to the pignistic probability distribution (12), since $|\mathcal{P}_1(B)| = |B|$.

Nonetheless, we cannot rigorously consider it a generalization of the pignistic transform according to the axiomatic Definition 6, as it does not fulfill the axiom of Projectivity[2]. One would expect the k-additive pignistic transform to reduce to the identity transformation when applied to an already k-additive belief function b_k. Unfortunately, the barycenter of the at most k-additive belief functions dominating such a b_k cannot be b_k. This can be pictured in the binary case of Figure A.5, but here is a more illustrative example on the ternary case: Let us consider $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and m the mass function on Ω:

$$\begin{aligned} m(\{\omega_1, \omega_2\}) &= 0.2 \\ m(\omega_1) &= 0.3 \\ m(\omega_2) &= 0.5 \end{aligned}$$

and $m(.) = 0$ otherwise. The barycenter of the set of at most 2-additive belief functions dominating m is given by:

$$\begin{aligned} m_2^{[\mathcal{B}]}(\{\omega_1, \omega_2\}) &= 0.6667 \\ m_2^{[\mathcal{B}]}(\omega_1) &= 0.3667 \\ m_2^{[\mathcal{B}]}(\omega_2) &= 0.5667 \end{aligned}$$

and $m_2^{[\mathcal{B}]}(.) = 0$ otherwise, whereas we would expect $m_2^{[\mathcal{B}]} = m$.

Hence, (18) is not a projective transform. Nonetheless, the close relationship between the geometric barycenter of Theorem 1 and the pignistic transform is enticing, and just a minor modification of Equation (18) is required to obtain an interesting k-additive generalization which fulfills Smets' axioms: It is sufficient to restrict the distribution process to mass assignments to focal elements of cardinality strictly greater than k.

Definition 8. *Given an arbitrary belief function $b : \mathcal{P}(\Omega) \rightarrow [0, 1]$, its* ***geometric k-additive pignistic transform*** *is defined as the k-additive belief function with mass assignment:*

$$m_k^{[\mathcal{G}]}(A) = \begin{cases} m(A) + \sum\limits_{\substack{B \supset A,\\ B \subseteq \Omega,\\ |B| > k}} \dfrac{m(B)}{|\mathcal{P}_k(B)|} & \forall A \subseteq \Omega : |A| \leq k \\ 0 & \forall A \subseteq \Omega : |A| > k. \end{cases} \quad (19)$$

The pignistic transform corresponds to a redistribution process in which the mass of each focal element is re-assigned *on an equal basis* among its elements (size 1 subsets). Equation (19) represents an analogous redistribution process in which the mass of each focal elements is re-assigned *to each subset of size $\ell \leq k$ on an equal basis.*

6.3. Geometric interpretation

A simple geometric interpretation can be provided for the geometric transform of Definition 8 as well, in terms of the barycenter (18) of the polytope of k-additive dominating belief functions, and the k-additive part of the original b.f.

Consider a belief function b whose order of additivity is greater than k, denote by $\vec{b}$ the corresponding vector in the belief space. By Equation (9) it is possible to rewrite $\vec{b}$ as a sum of two components:

$$\vec{b} = \sum_{|A| \leq k} m_b(A) \vec{b}_A + \sum_{|A| > k} m_b(A) \vec{b}_A = \vec{b}_{\leq k} + \vec{b}_{\leq k}.$$

Note that the vector $\vec{b}_{\leq k} \notin \mathcal{B}_k$ does not represent a valid belief function, as it is not normalized: $||\vec{b}_{\leq k}||_1 < 1$. Nevertheless, $\frac{\vec{b}_{\leq k}}{||\vec{b}_{\leq k}||_1} \in \mathcal{B}_k$. Similarly, $\frac{\vec{b}_{>k}}{||\vec{b}_{>k}||_1} \in \mathcal{B} \setminus \mathcal{B}_k$.

Let us then consider $\tilde{m}_k^{[\mathcal{B}]} = \overline{\mathfrak{B}^k \left[\frac{\vec{b}_{>k}}{||\vec{b}_{>k}||_1} \right]}$, the barycenter of the the set of k-additive belief functions dominating $\vec{b}_{>k}$, normalized to be a valid belief function. We have, by Theorem 1:

$$\tilde{m}_k^{[\mathcal{B}]}(A) = \begin{cases} \sum\limits_{B \supseteq A} \dfrac{\tilde{m}(B)}{|\mathcal{P}_k(B)|}, & \forall A \in \mathcal{P}_k(\Omega) \\ 0 & A \notin \mathcal{P}_k(\Omega), \end{cases} \quad (20)$$

$$= \begin{cases} \dfrac{1}{||\vec{b}_{>k}||_1} \cdot \sum\limits_{\substack{B \supset A,\\ B \subseteq \Omega,\\ |B| > k}} \dfrac{m(B)}{|\mathcal{P}_k(B)|}, & \forall A \subseteq \Omega : |A| \leq k \\ 0 & \forall A \subseteq \Omega : |A| > k. \end{cases} \quad (21)$$

where, $\tilde{m}$ and m stand, respectively, for the mass functions of $\frac{\vec{b}_{>k}}{||\vec{b}_{>k}||_1}$, and of $\vec{b}$.

By comparing Equations (19) and (20) it follows that, after denoting by m_k the mass function of $\frac{\vec{b}_{\leq k}}{||\vec{b}_{\leq k}||_1}$,

$$m_k^{[\mathcal{G}]} = ||\vec{b}_{\leq k}||_1 \cdot m_k + ||\vec{b}_{>k}||_1 \cdot \tilde{m}_k^{[\mathcal{B}]}.$$

As $||\vec{b}_{\leq k}||_1 + ||\vec{b}_{>k}||_1 = 1$, the latter amounts to a linear combination of the k-additive part of b (i.e. $b_{\leq k}$) and of the barycenter of k-additive belief functions dominating the remaining part of b (i.e. $b - b_{\leq k}$), whose coefficient is the proportion of mass which is assigned to a focal element of cardinality lower than or equal to k.

In the next section, we provide some mathematical properties of the proposed asymmetric and geometric k-additive pignistic transforms. Moreover, we establish their coherence with Smets' axioms, making them eligible as k-additive generalizations of the pignistic transform.

7. Properties of the sets of pignistic k-additive transforms

7.1. Dominance properties

Clearly, for any K-additive belief function b (note that any b is at least $|\Omega|$-additive) it is possible to define $K - 1$ pairs of k-additive belief functions $\{b_k^{[\mathcal{A}]}, b_k^{[\mathcal{G}]}\}$ with $1 \leq k \leq K - 1$ by applying the asymmetric and geometric k-additive pignistic transforms for all the values of k up to K.

Definition 9. *Let b be a K-additive belief function. The two* ***sets of asymmetric and geometric at most K-additive pignistic***

[2]Another option would be to generalize the Projectivity axiom itself in a sensible way. We will not pursuit that line of reasoning in the present paper.

belief functions *of b are defined as:*

$$\mathcal{A}[b] = \left\{b^{[S]}, b_2^{[\mathcal{A}]}, \cdots, b_{K-1}^{[\mathcal{A}]}, b\right\}, \quad (22)$$

$$\mathcal{G}[b] = \left\{b^{[S]}, b_2^{[\mathcal{G}]}, \cdots, b_{K-1}^{[\mathcal{G}]}, b\right\}. \quad (23)$$

The next proposotion shows that applying a sequence of pignistic k-additive transforms with different values of hesitation threshold k is equivalent to applying directly the transform with the smallest k.

Proposition 2. *Let b be a k-additive belief function and $k_1, k_2 < k$. We have that*

$$\left(b_{k_1}^{[\mathcal{A}]}\right)_{k_2}^{[\mathcal{A}]} = b_{\min(k_1,k_2)}^{[\mathcal{A}]}, \quad \left(b_{k_1}^{[\mathcal{G}]}\right)_{k_2}^{[\mathcal{G}]} = b_{\min(k_1,k_2)}^{[\mathcal{G}]}.$$

The proof is in Appendix A.4. As a consequence, it is possible to compute in a recursive manner all the elements of $\mathcal{A}[b]$ (resp. $\mathcal{G}[b]$) using decreasing values of the hesitation threshold. Now, let us study the dominating properties of $\mathcal{A}[b]$ and of $\mathcal{G}[b]$.

Proposition 3. *Let b be a K-additive belief function, and let $k < K$. Then*

$$b \ll b_k^{[\mathcal{A}]}, \quad b \ll b_k^{[\mathcal{G}]},$$

or, in other words, both the asymmetric and geometric k-additive pignistic transforms of b dominate b.

The proof can be found in Appendix A.5. All these propositions are useful to prove several interesting properties, which are summarized below:

Property 4. *Let b be a K-additive belief function. The sets $\mathcal{A}[b]$ and $\mathcal{G}[b]$ of pignistic k-additive belief functions, $1 \le k \le K$, have the following properties:*

1. *if $k = K$ both transforms are idle: $b = b_k^{[\mathcal{A}]} = b_k^{[\mathcal{G}]}$;*
2. *if $k = 1$, the result corresponds to the Shapley value: $b_1^{[\mathcal{A}]} = b_1^{[\mathcal{G}]} = b^{[S]}$;*
3. *$\forall k \le K$, $b_k^{[\mathcal{A}]}$ and $b_k^{[\mathcal{G}]}$ are k-additive belief functions which are uniquely defined, and which dominate b;*
4. *$\forall k \le K$, $\mathcal{A}[b_k^{[\mathcal{A}]}] \subseteq \mathcal{A}[b]$, and $\mathcal{G}[b_k^{[\mathcal{G}]}] \subseteq \mathcal{G}[b]$;*
5. *$\forall k_2 < k_1 \le K$, $b_{k_1}^{[\mathcal{A}]} \ll b_{k_2}^{[\mathcal{A}]}$ and $b_{k_1}^{[\mathcal{G}]} \ll b_{k_2}^{[\mathcal{G}]}$*
6. *We have:*

$$b = b_K^{[\mathcal{A}]} \ll b_{K-1}^{[\mathcal{A}]} \ll \cdots \ll b_2^{[\mathcal{A}]} \ll b_1^{[\mathcal{A}]} = b^{[S]}$$

and

$$b = b_K^{[\mathcal{G}]} \ll b_{K-1}^{[\mathcal{G}]} \ll \cdots \ll b_2^{[\mathcal{G}]} \ll b_1^{[\mathcal{G}]} = b^{[S]};$$

The proofs are given in Appendix A.6.

7.2. Coherence with Smets' axioms

Most importantly, we now consider the coherence of the two proposed generalizations of the pignistic transform with respects to the 5 rationale of Smets [35, 36, 6, 34].

Theorem 2. *Both asymmetric and geometric k-additive pignistic transforms fulfil Smets' rationality arguments. In other words:*

1. ***Linearity:*** *Asymmetric and geometric k-additive pignistic transforms commute with the convex closure operator.*
2. ***Projectivity:*** *Asymmetric and geometric k-additive pignistic transforms are idle for all at least k-additive belief functions.*
3. ***Efficiency:*** *A bet on the entire decision space is bound to win with a certain probability.*
4. ***Anonimity:*** *The outcome of both asymmetric and geometric k-additive pignistic transforms is not sensitive to permutations of the elements of Ω.*
5. ***False Event:*** *Asymmetric and geometric k-additive pignistic transforms assign nil mass to every false event.*

A proof of this theorem is given in Appendix A.7. This theorem definitely shows that the asymmetric and the geometric k-additive pignistic transforms are valid generalizations of Smets' pignistic transform.

8. Evaluation on a decision making context

As interesting as their formal properties are, the rationale of the proposed k-additive pignistic transforms is their potential application to imprecise decision making. In this last part of the paper we therefore compare them in a real world scenario. First, we present a framework to evaluate the interest of any imprecise decision method with respect to classical decision making. Second, we provide some basic background on handwriting recognition and we explain how imprecise decision are crucial in this field. Finally, we apply the asymmetric and geometric k-additive pignistic transforms on real handwriting datasets, and we compare them together and with the classical pignsitic transform.

8.1. How to evaluate imprecise decisions ?

First of all, we need to work out a sensible way of evaluate imprecise decisions. In [62], a method is indeed proposed to evaluate and compare the efficiency of precise and imprecise decision algorithms, based on the following idea.

A dataset is considered for a classification task, and for each item of the dataset, a decision is made on the class to which this item should belong, according to the different decision processes to compare. By performing several classification tasks in exactly the same setting while changing the decision process, the variations of the accuracy rate is only due to the efficacy of the decision process itself. The highest the accuracy rates, the more effective the decision rule. The problem reduces then to finding a measure of accuracy which allow comparisons of decision rules whose outcomes do not necessarily have the same cardinality.

In a classical setting, accuracy can be defined as follows. For each item in the dataset the potential classes $\{1, ..., C\}$ are ranked according to some confidence measure. The first N such classes are selected, so that for each item in the dataset a decision of constant cardinality N is made, and the accuracy $Acc(N)$ of the decision is the fraction of items for which the correct class is one of the first N. If $N = 1$, we have a precise decision. When decision outcomes of different cardinalities are allowed, as in imprecise decision making, the notion of *mean cardinality* becomes relevant.

Let us assume that the dataset contains 100 items, and that a decision of cardinality 1 is made for 60 of them, whilst a decision of cardinality 2 is made for the remaining 40. The mean cardinality of the decision process is $\frac{60+2\times40}{100} = 1.4$, a value which belongs to the set of rational numbers $\mathbb{Q}$ (contrarily to the classical case). Formally, if T is the size of the dataset and $\alpha_j,\ j > 0$ represents the number of items for which a decision of cardinality j is made, then the *mean cardinality* of the decision is defined as:

$$Q = \frac{\sum_j j \cdot \alpha_j}{T}.$$

The classical definition of accuracy $Acc(N)$ can be reformulated using the Kronecker symbol: $\delta_i^N = 1$ if and only if the true class of the i-th item of the dataset is ranked in the first N classes proposed by the classifier, whilst $\delta_i^N = 0$ otherwise:

$$Acc(N) = \frac{\sum_{i=1}^T \delta_i^N}{T} = \frac{\sum_{i=1}^T \delta_i^N \times N}{T \times N} = \frac{\sum_{i=1}^T \delta_i^N \times N}{\sum_{i=1}^T \delta_i^C \times N}$$

as $T = \sum_{i=1}^T \delta_i^C$ since $\delta_i^C = 1\ \forall i$.

Let L_i, be the list of proposed classes for item i, $|L_i|$ being the length of this list. A more general definition of accuracy is therefore

$$Acc(Q) = \frac{\sum_{i=1}^T \delta_i^{|L_i|} \times |L_i|}{\sum_{i=1}^T \delta_i^T \times |L_i|} = \frac{\sum_j j \cdot \beta_j}{\sum_j j \cdot \alpha_j}$$

where $\overline{|L_i|} = Q$ is the mean cardinality of the decision, α_j is the number of items for which a decision of cardinality j is made (correct or not), while β_j is the number of items for which a decision of cardinality j is *correctly* made. In our toy example, if among the 60 precise decisions, 50 are correct, and among the 40 imprecise decisions, 30 are correct, then the accuracy is:

$$Acc(1.4) = \frac{50 \times 1 + 30 \times 2}{60 \times 1 + 40 \times 2} = \frac{110}{140} = 78.6\%$$

To compare the accuracy $Acc(Q)$ of an imprecise classifier with the accuracy $Acc(N)$ of a precise one, we can simply consider the accuracy of the latter for the lower $\lfloor Q \rfloor$ and upper $\lceil Q \rceil$ integer approximations of Q, from which the linear interpolation

$$iAcc(Q) = (\lceil Q \rceil - Q) \cdot Acc(\lfloor Q \rfloor) + (Q - \lfloor Q \rfloor) \cdot Acc(\lceil Q \rceil)$$

provides a good indicator. For instance, if the reference precise algorithm has $Acc(1) = 80\%$ and $Acc(2) = 90\%$, then, $Acc(1.7)$ can be compared to $iAcc(1.7) = 87\%$.

It is also useful to consider the set of *partial accuracy* rates $pAcc(j) = \beta_j/\alpha_j$ for all ranks j, i.e., the accuracy rates computed on each set of items for which a decision of cardinality j is made. If an imprecise decision process always provides better $pAcc(j)$'s than a second one for decisions of small cardinality j, it is arguably robust: In such a case, it is likely that such precise, narrow decisions are less frequent in the first process than in the second, as the first decision strategy focuses only when the decision is robust enough.

8.2. Basis of handwritting recognition

Classically, handwritting recognition algorithms are based on the following structure :

1. An image processing module segments the pen mark from the background and a connectivity analysis provides the locations of the spaces between the letters, in order to extract each word separately.
2. Each word is processed to extract meaningfull descriptors of the shape of the penmark, and these features are used as variable inputs for a classification task (where the classes corresponds to the words of a lexicon). The classifier output is classically made of a list of several words, ordered by decreasing interest (most of the time, the interest of each word is quantized by a probability). Eventually, to improve the word-level recognition, a DST-based combination of several classifiers can be used instead of a single classifier [11].
3. Finally, a linguistic layer is used, so that grammar or syntaxic rules can be used to eraze some mistake locally produced by the isolated word recognition module.

Obviously, the longer the lists from the classifier are, the larger the chances of having the right classes within the lists. Hence, to make sure that the minimum number of errors occur, classicaly, rather long lists are used. Hence, for instance, when recognizing a short sentence of 8 words, with lists of 10 propositions for each word, the linguistic layer has to compare 8^{10} potential sentences. Among those words, it is likely to assume that at least, one or two of them are rather simple to recognize (a short list of 2 or 3 items is safe enough), which drastically reduce the combinatory of the sentences. This illustates well the interest of imprecise decisions at step 2. As the cardinality of the decision is adaptative, it is possible to have more propositions on words which are difficult to recognize, whereas less are given for words which are more easily processed.

8.3. Application to handwriting recognition

Equipped with these measures of accuracy, let us compare precise decisions (as provided by the classical pignistic transform) and two imprecise decision strategies associated with the two proposed generalizations to k-additive belief functions in a classification task which concerns handwritten word recognition on three publicly available databases: the IFN/ENIT benchmark dataset of Arabic words and the RIMES and IRONOFF databases of Latin words. We adopt here the protocol published in [62], with the crucial assistance of its authors. We provide therefore only a brief description of the experimental setting, while we invite readers interested in more details to refer to [62].

The IFN/ENIT [63] contains 32,492 handwritten words (Arabic symbols) of 946 Tunisian town/villages names written by 411 different writers. Four different sets (a, b, c, d) are used for training and 3000 word images from set (e) for testing. The RIMES database [64] is composed of isolated handwritten word snippets extracted from handwritten letters (latin symbols). In our experiments, 36000 snippets of words are used to train different HMM classifiers and 3000 words are used in the test.

IRONOFF [65] is both an on-line and off-line dataset. The sub-dataset IRONOFF-Chèque only contains a small lexicon of roughly 30 words used on French checks (numbers, currencies, etc.). 7956 words are used for training and 3987 are used for testing. As the absolute accuracy of each classifier is not an issue here, a rather simple protocol is applied. A HMM classifier based on the upper contour description of the image of the word is used to derive posterior probabilities for the word to recognize to belong to each class [56]. As it clearly appears in Table 1, the three datasets present heterogeneous levels of difficulty with respect to the adopted classifier: RIMES is rather challenging, IFN/ENIT is of intermediate difficulty, while IRONOFF is the easiest to cope with.

Datasets	$Acc(1)$	$Acc(2)$	$Acc(3)$	$Acc(4)$
RIMES	54.10	66.40	72.13	75.87
IFN/ENIT	73.60	79.77	82.83	84.60
IRONOFF	85.65	91.51	93.84	95.55

Table 1: Accuracy rates for the RIMES, IFN/ENIT and IRONOFF datasets.

After the classification step, the posterior probability distribution of the HMM classifier is either directly used to make a precise decision, or converted into a consonant mass function by inverse pignistic transform [66] to make an imprecise decision. Then, the two proposed generalizations of the pignistic transform are applied, with different values of $k \in \{2, 3, 4\}$. Performances are measured as explained in the previous section, and are presented in Table 2. They are based on the values T, α_j, β_j summarized in Table 3.

In Table 2 we consider the accuracy differential $\Delta = Acc(Q) - iAcc(Q)$ for both asymmetric and geometric transform. Positive values (≥ 0.4) show therefore an improvement with respect to classical precise decision, whereas values close to zero ($|\Delta| < 0.1$) indicate that the imprecise and precise decisions are quite equivalent. There are no negative values (≤ -0.4) to indicate a lower accuracy of imprecise decisions. In particular, the improvement is always significant for the RIMES and IRONOFF datasets. On the other hand, no real variations appears in the case of the IFN/ENIT dataset.

Now, let us compare the performances of asymmetric and geometric k-additive pignistic transform. The improvements associated with the geometric transform are clearly greater on the RIMES and IRONOFF datasets, whereas they are neck to neck with those produced by the asymmetric k-additive pignistic transform on the IFN/ENIT dataset. From this comparison, the geometric transform seems to perform better than the asymmetric one, in spite of a less clear semantic intepretation.

Nonetheless, a more refined analysis shows that the asymmetric transform has in fact some advantages over the geometric one. Whatever the cardinality of k it appears that $pAcc(i)$, $\forall i \leq k-1$ is greater for the asymmetric case than for the geometric case (with only one exception with the RIMES dataset where $k = 4$ and $i = 2$). In other words, "asymmetric" decisions seem to be always more trustworthy when the cardinality of the decision is small. This comes from the fact that the asymmetric redistribution pattern supports the largest focal elements, so that it is less likely to encourage decisions focused on "small" focal sets. When this happens, the decision is rather robust.

9. Conclusion

In this article, we have introduced the notion of imprecise decision, as a narrowing of the decision space, and we have illustrated how such a formalism is interesting in the framework of Dempster-Shafer Theory, and more precisely, in the Transferable Belief Model. In this context, we have justified why k-additive belief functions seem adapted to model imprecise decision, which motivates for the search of a generalization of the pignistic transform, the point being to define a pignistic k-additive belief function, to which the classical pignistic probability would be a particular case (probabilities corresponding to the case of 1-additive belief functions).

The pignsitic probability, in addition to be a posteriori justified by different rationality arguments, was first intuited by the principle of insufficient reason, and is known to correspond to the Shapley value. Thus, we have proposed two distinct generalizations, both of them fulfilling the rationality arguments: The first one (called *asymmetric*) is based on an intuitive redistribution process of the mass assignments, according to the original motivation of the principle of insufficient reason, while the second one (called *geometric*) is based on generalizing a well-known geometrical proprety of the Shapley value, that is, it corresponds to the barycenter of the polytope of dominating probabilities.

From a formal point of view, the results may appear as disappointing, as the barycenter of dominating k-additive belief functions obviously does not meet one of the rationality arguments. This lead us to modify the expression of this barycenter to meet an acceptable *geometric* generalization of the pignistic transform. As a consequence, our results on the barycenter of dominating k-additive belief functions are left incompleted and based on a conjecture, as these geometrical problems becomes partially unrealated to our original concern. Nonetheless, the complexity of the demonstration of this conjecture and its self-interest, as an open mathematical problem, is appealing and future works of us will focus on it, besides any consideration for imprecise decisions.

From an applicative point of view, we have applied three decisions strategies (the classical pignistic transform as well as our *asymmetric* and *geometric* generalizations) to a real world problem, i.e. multi-script handwriting recognition. In addition to a prime study of an adapted protocol to evaluate imprecise decisions, we have conducted experiments on three publicly available datasets, on which concuring results appears. The first conclusion stresses the intrest of imprecision decision in general, as both *asymmetric* and *geometric* generalizations provide interesting results. More precisely, it appears that, even if the *asymmetric* generalization provides more accurate focused decision, the overall accuracy is slightly better with the geometric k-additive pignistic transform, which experimentally justifies

Datasets	RIMES		IFN/ENIT		IRONOFF	
Pignistic Transform	**Asymmetric**	**Geometric**	**Asymmetric**	**Geometric**	**Asymmetric**	**Geometric**
k **= 2**						
Q	1.778	1.520	1.705	1.511	1.675	1.408
$Acc(Q)$ (%)	64.96	63.91	78.00	77.913	90.29	89.69
$iAcc(Q)$ (%)	63.67	60.50	77.95	76.75	89.60	88.04
Δ	+1.29	**+3.41**	+0.06	**+1.16**	+0.69	**+1.64**
$pAcc(1)$ (%)	**70.12**	60.11	**85.67**	78.81	**96.05**	92.27
$pAcc(2)$ (%)	64.22	65.66	76.40	77.48	88.91	87.82
k **= 3**						
Q	2.235	2.081	2.075	1.956	2.068	1.897
$Acc(Q)$ (%)	69.22	68.91	79.95	79.66	92.11	92.12
$iAcc(Q)$ (%)	67.75	66.86	80.00	79.50	91.67	90.90
Δ	+1.47	**+2.04**	-0.04	**+0.16**	+0.45	**+1.21**
$pAcc(1)$ (%)	**70.95**	66.25	**85.91**	83.44	**96.00**	94.36
$pAcc(2)$ (%)	**69.48**	69.32	**80.62**	79.90	**91.07**	90.59
$pAcc(3)$ (%)	68.81	69.36	77.97	78.03	91.53	92.06
k **= 4**						
Q	2.72	2.632	2.467	2.401	2.536	2.417
$Acc(Q)$ (%)	71.76	71.37	81.19	80.89	94.06	93.91
$iAcc(Q)$ (%)	70.53	70.02	81.20	81.00	92.76	92.48
Δ	+1.23	**+1.35**	**-0.01**	-0.11	+1.31	**+1.43**
$pAcc(1)$ (%)	**74.21**	72.37	**87.10**	86.67	**96.76**	95.93
$pAcc(2)$ (%)	72.45	**72.64**	**83.21**	82.75	**93.19**	93.11
$pAcc(3)$ (%)	**71.76**	70.72	**80.09**	79.63	**93.61**	93.52
$pAcc(4)$ (%)	71.14	71.14	79.35	78.96	94.02	93.88

Table 2: Comparison between our algorithm and the classical approach.

Datasets	T	$k = 2$	$k = 3$	$k = 4$
Asymmetric Geometric		$\alpha_1, \beta_1, \alpha_2, \beta_2$ $\alpha_1, \beta_1, \alpha_2, \beta_2$	$\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3$ $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3$	$\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3, \alpha_4, \beta_4$ $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3, \alpha_4, \beta_4$
RIMES	3000	666, 467, 2334, 1499 1439, 865, 1561, 1025	661, 469, 973, 676, 1366, 940 883, 585, 991, 687, 1126, 781	539, 400, 686, 497, 850, 610, 925, 658 619, 448, 709, 515, 830, 587, 842, 599
IFN/ENIT	3000	886, 759, 2114, 1615 1468, 1157, 1532, 1187	887, 762, 1001, 807, 1112, 867 1045, 872, 1040, 831, 915, 714	775, 675, 786, 654, 703, 563, 736, 584 833, 722, 800, 662, 697, 555, 670, 529
IRONOFF	3979	1292, 1241, 2687, 2389 2354, 2172, 1625,1427	1299, 1247, 1109, 1010, 1571, 1438 1631, 1539, 1126, 1020, 1222, 1125	1080, 1045, 808, 753, 971, 909, 1120, 1053 1229, 1179, 842, 784, 927, 867, 981, 921

Table 3: Nessary values for the computations of the results of Table 2.

the choice we made on its definition, as a variant of the strictly speaking barycenter of dominating k-additive belief functions.

These results being particularly encouraging, future works will focus on (1) other applications of imprecise decision, (2) the definition of yet other transforms of belief functions to k-additive belief functions, (3) the study of the barycenter of dominating k-additive belief functions, and (4) re-questioning or generalizing the rationality arguments to link in a more elegant manner the geometric k-additive pignistic transform and the barycenter of dominating k-additive belief functions.

Acknowledgements

The authors would like to thank very much D. Yousri Kessentini and Pr. Thierry Paquet for providing the experimental setting on the handwriting script recognition experiments.

References

[1] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[2] P. Smets, R. Kennes, The transferable belief model, Artificial intelligence 66 (2) (1994) 191–234.

[3] P. Walley, Towards a unified theory of imprecise probability, International Journal of Approximate Reasoning 24 (2000) 125–148.

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

[5] T. M. Strat, Decision analysis using belief functions, International Journal of Approximate Reasoning 4 (1990) 391–417.

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

[7] M. Troffaes, Decision making under uncertainty using imprecise probabilities, International Journal of Approximate Reasoning 45 (1) (2007) 17–29.
[8] M. Daniel, On transformations of belief functions to probabilities, International Journal of Intelligent Systems, special issue on Uncertainty Processing.
[9] J. M. Keynes, Fundamental ideas, A Treatise on Probability, Ch. 4.
[10] O. Aran, T. Burger, A. Caplier, L. Akarun, A belief-based sequential fusion approach for fusing manual and non-manual signs, Pattern Recognition 42 (5) (2009) 812–822.
[11] Y. Kessentini, T. Burger, T. Paquet, Evidential ensemble hmm classifier for handwriting recognition, in: Proceedings of IPMU, 2010.
[12] T. Burger, O. Aran, A. Caplier, Modeling hesitation and conflict: A belief-based approach for multi-class problems, Machine Learning and Applications, Fourth International Conference on (2006) 95–100.
[13] T. Burger, Defining new approximations of belief function by means of dempster's combination, in: Proceedings of the Workshop on the theory of belief functions, 2010.
[14] A. Chateauneuf, J. Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, Mathematical Social Sciences 17 (1989) 263–283.
[15] P. Smets, Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem, International Journal of Approximate Reasoning 9 (1993) 1–35.
[16] T. Denoeux, A new justification of the unnormalized dempster's rule of combination from the Least Commitment Principle, in: Proceedings of FLAIRS'08, Special Track on Uncertaint Reasoning, 2008.
[17] R. R. Yager, The entailment principle Dempster-Shafer granules, International Journal of Intelligent Systems 1 (1986) 247–262.
[18] D. Dubois, H. Prade, A set theoretical view of belief functions, International Journal of Intelligent Systems 12 (1986) 193–226.
[19] T. Denoeux, Conjunctive and disjunctive combination of belief functions induced by non distinct bodies of evidence, Artificial Intelligence.
[20] F. Cuzzolin, R. Frezza, Geometric analysis of belief space and conditional subspaces, in: Proceedings of the 2^{nd} International Symposium on Imprecise Probabilities and their Applications (ISIPTA2001), Cornell University, Ithaca, NY, 26-29 June 2001.
[21] F. Cuzzolin, Visions of a generalized probability theory, PhD dissertation, Università di Padova, Dipartimento di Elettronica e Informatica (19 February 2001).
[22] F. Cuzzolin, A geometric approach to the theory of evidence, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 38 (4) (2008) 522–534.
[23] M. Grabisch, K-order additive discrete fuzzy measures and their representation, Fuzzy sets and systems 92 (1997) 167–189.
[24] F. Cuzzolin, On the credal structure of consistent probabilities, in: Logics in Artificial Intelligence, Vol. 5293/2008, Springer Berlin / Heidelberg, 2008, pp. 126–139.
[25] P. Miranda, M. Grabisch, P. Gil, Dominance of capacities by k-additive belief functions, European Journal of Operational Research 175 (2006) 912–930.
[26] T. Burger, F. Cuzzolin, The barycenters of the k-additive dominating belief functions & the pignistic k-additive belief functions.
[27] A. P. Dempster, A generalization of bayesian inference, in: Classic Works of the Dempster-Shafer Theory of Belief Functions, 2008, pp. 73–104.
[28] G. Shafer, P. P. Shenoy, Local computation on hypertrees, Working paper No. 201, School of Business, University of Kansas.
[29] P. P. Shenoy, G. Shafer, An axiomatic framework for Bayesian and belief function propagation, in: Proceedings of the AAAI Workshop of Uncertainty in Artificial Intelligence, 1988, pp. 307–314.
[30] H. Nguyen, On random sets and belief functions, in: Classic Works of the Dempster-Shafer Theory of Belief Functions, 2008, pp. 105–116.
[31] J. K
"uhr, D. Mundici, De Finetti theorem and Borel states in [0, 1]-valued algebraic logic, International journal of approximate reasoning 46 (3) (2007) 605–616.
[32] G. Shafer, Comments on "constructing a logic of plausible inference: a guide to cox's theorem", by kevin s. van horn, Int. J. Approx. Reasoning 35 (1) (2004) 97–105.
[33] L. Shapley, A value for n-person games, in: Contributions to the Theory of Games,.
[34] P. Smets, Decision making in a context where uncertainty is represented by belief functions, Belief functions in business decisions (2002) 17–61.
[35] P. Smets, No Dutch book can be built against the TBM even though update is not obtained by Bayes rule of conditioning, in: Workshop on probabilistic expert systems, Societa Italiana di Statistica, Roma, 1993, pp. 181–204.
[36] P. Smets, The axiomatic justification of the transferable belief model, Tech. rep., Universite' Libre de Bruxelles, Technical Report TR/IRIDIA/1995-8.1 (1995).
[37] R. Jeffrey, Conditioning, kinematics, and exchangeability, Causation, chance, and credence 1 (1988) 221–255.
[38] P. Teller, Conditionalization and observation, Synthese 26 (2) (1973) 218–258.
[39] B. R. Cobb, 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.
[40] F. Cuzzolin, Geometry of upper probabilities, in: Proceedings of the 3^{rd} Internation Symposium on Imprecise Probabilities and Their Applications (ISIPTA'03), July 2003.
[41] F. Voorbraak, A computationally efficient approximation of Dempster-Shafer theory, International Journal on Man-Machine Studies 30 (1989) 525–536.
[42] R. Giles, Foundations for a theory of possibility, Fuzzy Information and Decision Processes (1982) 183–195.
[43] J. J. Sudano, Pignistic probability transforms for mixes of low- and high-probability events, in: Proceedings of the International Conference on Information Fusion, 2001.
[44] F. Cuzzolin, Two new Bayesian approximations of belief functions based on convex geometry, IEEE Transactions on Systems, Man, and Cybernetics, Part B 37 (4) (2007) 993–1008.
[45] F. Cuzzolin, Credal semantics of Bayesian transformations in terms of probability intervals, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 40 (2) (2010) 421–432.
[46] F. Cuzzolin, Geometry of relative plausibility and relative belief of singletons, Annals of Mathematics and Artificial Intelligence 59 (2010) 47–79.
[47] F. Cuzzolin, Semantics of the relative belief of singletons, Interval/Probabilistic Uncertainty and Non-Classical Logics (2008) 201–213.
[48] F. Cuzzolin, Dual properties of the relative belief of singletons, PRICAI 2008: Trends in Artificial Intelligence (2008) 78–90.
[49] M. Daniel, Consistency of probabilistic transformations of belief functions, in: IPMU, 2004, pp. 1135–1142.
[50] T. Burger, A. Caplier, A Generalization of the Pignistic Transform for Partial Bet, in: Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), Verona, Italy, July 1-3, Springer-Verlag New York Inc, 2009, pp. 252–263.
[51] P. Baroni, P. Vicig, Transformations from imprecise to precise probabilities, in: ECSQARU, 2003, pp. 37–49.
[52] J. Dezert, F. Smarandache, M. Daniel, The generalized pignistic transformation, Arxiv preprint cs/0409007.
[53] F. Smarandache, J. Dezert, An introduction to the DSm theory for the combination of paradoxical, uncertain and imprecise sources of information, in: Proceedings of the 13th International Congress of Cybernetics and Systems, 2005, pp. 6–10.
[54] L. Cholvy, Using logic to understand relations between DSmT and Dempster-Shafer Theory, Symbolic and Quantitative Approaches to Reasoning with Uncertainty (2009) 264–274.
[55] A. Appriou, Knowledge propagation in information fusion processes, Keynote talk, wtbf'10, Brest, France, mars 2010.
[56] Y. Kessentini, T. Paquet, A. B. Hamadou, Off-line handwritten word recognition using multi-stream hidden markov models, Pattern Recognition Letters 30 (1) (2010) 60–70.
[57] L. R. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, in: Proceedings of the IEEE, 1989, pp. 257–286.
[58] E. Ramasso, M. Rombaut, D. Pellerin, Forward-Backward-Viterbi procedures in the Transferable Belief Model for state sequence analysis using belief functions, Symbolic and Quantitative Approaches to Reasoning with Uncertainty (2007) 405–417.
[59] H. Xu, P. Smets, Reasoning in evidential networks with conditional belief functions, International Journal of Approximate Reasoning 14 (2-3)

(1996) 155–185.
[60] P. Smets, B. Ristic, Kalman filter and joint tracking and classification in the TBM framework, in: Proceedings of the Seventh International Conference on Information Fusion, Vol. 1, Citeseer, 2004, pp. 46–53.
[61] O. Aran, T. Burger, A. Caplier, L. Akarun, Sequential Belief-Based Fusion of Manual and Non-manual Information for Recognizing Isolated Signs, Gesture-Based Human-Computer Interaction and Simulation (2009) 134–144.
[62] T. Burger, Y. Kessentini, T. Paquet, Dealing with precise and imprecise decisions with a Dempster-Shafer theory based algorithm in the context of handwritten word recognition, in: 2010 12th International Conference on Frontiers in Handwriting Recognition, IEEE, 2010, pp. 369–374.
[63] M. Pechwitz, S. Maddouri, V. Maergner, N. Ellouze, H. Amiri, Ifn/enit - database of handwritten arabic words, Colloque International Francophone sur l'Ecrit et le Doucement (2002) 129–136.
[64] E. Grosicki, M. Carre, J. Brodin, E. Geoffrois, Results of the rimes evaluation campaign for handwritten mail processing, International Conference on Document Analysis and Recognition 0 (2009) 941–945.
[65] C. Viard-Gaudin, P. M. Lallican, P. Binter, S. Knerr, The ireste on/off (ironoff) dual handwriting database, International Conference on Document Analysis and Recognition 0 (1999) 455–458.
[66] D. Dubois, H. Prade, P. Smets, New semantics for quantitative possibility theory, in: S. Benferhat, P. Besnard (Eds.), Proc. of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU 2001), Springer-Verlag, Toulouse, France, 2001, pp. 410–421.

Appendix A. Proofs

In this section, we present the demonstrations of the theorems and propositions given in the article.

Appendix A.1. Property 3

If we use the shorthand notation $\#\rho$ for the cardinality of the set of the permutations ρ of Ω, then, the center of mass $\overline{\mathfrak{B}^1[b]}$ of $\mathfrak{B}^1[b]$ is given by

$$\sum_{\rho} \frac{p^{\rho}[b]}{\#\rho}$$

which, by Equation (11), corresponds to a Bayesian mass function which assigns to any focal element $\{\omega_i\}$ the value

$$\sum_{B \supseteq \{\omega_i\}} m(B) \frac{\#\rho : \forall \omega_j <_{\rho} \omega_i : B \not\supset \{\omega_j\}}{\#\rho}.$$

where $\omega_j <_{\rho} \omega_i$ indicates that ω_j comes before ω_i in the list of elements associated with the permutation ρ. To simplify this expression, we need to compute for each singleton focal element $B \supseteq \{\omega_i\}$ the number of permutations ρ of Ω such that B does not include any singleton ω_j which comes before ω_i ($\omega_j <_{\rho} \omega_i$) in the associated list $\{\omega_{\rho(1)}, ..., \omega_{\rho(|\Omega|)}\}$.

For all possible positions of ω_j in the list, the permutation must be such that all elements before ω_j are extracted from B^c, the complement of B. In any admissible permutation, ω_j has to appear in one of the first $|\Omega|-|B|+1$ locations (as otherwise some other elements of B would come before ω_j in the list). For each position i of ω_j, the number of admissible permutations is given by the possible dispositions $\frac{(|\Omega|-|B|)!}{[(|\Omega|-|B|)-(i-1)]!}$ of $(|\Omega| - |B|)$ points (the elements of B^c) in $i - 1$ locations (the elements of the list before ω_j), multiplied by the number $(|\Omega| - i)!$ of permutations of the remaining $n - i$ singletons, which can appear after ω_j in any order.

Then, $\overline{\mathfrak{B}^1[b]}$ is given by a mass function which assigns to $\{\omega_i\}$ the value:

$$\sum_{B \supseteq \{\omega_i\}} m(B) \sum_{i=1}^{|\Omega|-|B|+1} \frac{(|\Omega| - |B|)!}{[(|\Omega| - |B|) - (i-1)]!} \frac{(|\Omega| - i)!}{|\Omega|!}.$$

We can further simplify the multiplicative coefficient of $m(B)$ in the above expression, as follows:

$$\begin{aligned}
& \sum_{i=1}^{|\Omega|-|B|+1} \frac{(|\Omega| - |B|)!}{[(|\Omega| - |B|) - (i-1)]!} \frac{(|\Omega| - i)!}{|\Omega|!} \\
= & \sum_{i=1}^{|\Omega|-|B|+1} \frac{(|\Omega| - |B|)!}{[(|\Omega| - i) - (|B| - 1)]!} \frac{(|\Omega| - i)!}{|\Omega|!} \\
= & \sum_{i=1}^{|\Omega|-|B|+1} \frac{(|\Omega| - |B|)!}{[(|\Omega| - i) - (|B| - 1)]!} \frac{(|B|-1)!}{(|B|-1)!} \frac{(|\Omega| - i)!}{|\Omega|!} \\
= & \frac{(|\Omega| - |B|)!(|B| - 1)!}{|\Omega|!} \\
& \times \sum_{i=1}^{|\Omega|-|B|+1} \frac{(|\Omega| - i)!}{[(|\Omega| - i) - (|B| - 1)]!(|B| - 1)!} \\
= & \frac{(|\Omega| - |B|)!(|B| - 1)!}{|\Omega|!} \sum_{i=1}^{|\Omega|-|B|+1} \binom{|\Omega| - i}{|B| - 1},
\end{aligned}$$

which, after recalling that $\sum_{i=1}^{|\Omega|-|B|+1} \binom{|\Omega|-i}{|B|-1} = \binom{|\Omega|}{|B|}$ becomes

$$= \frac{(|\Omega| - |B|)!(|B| - 1)!}{|\Omega|!} \binom{|\Omega|}{|B|} = \frac{1}{|B|}.$$

As a consequence,

$$\overline{\mathfrak{B}^1[b]} = \sum_{B \supseteq \{\omega_i\}} \frac{m(B)}{|B|} = m^{[S]}(x), \qquad \text{(A.1)}$$

i.e., $\overline{\mathfrak{B}^1[b]}$ corresponds to the pignistic probability $m^{[S]}$ [2].

Appendix A.2. Conjecture 1

In the case of a binary frame $\Omega = \{x, y\}$ the list of focal elements of size at most $k = 2$ obviously reads as $\mathcal{P}^2(\Omega) = \{\{x\}, \{y\}, \{x, y\}\}$, so that its possible permutations are six:

$$\begin{aligned}
\rho_1 &= (\{x\}, \{y\}, \Omega) & \rho_2 &= (\{x\}, \Omega, \{y\}) \\
\rho_3 &= (\{y\}, \{x\}, \Omega) & \rho_4 &= (\{x\}, \Omega, \{y\}) \\
\rho_5 &= (\Omega, \{x\}, \{y\}) & \rho_6 &= (\Omega, \{y\}, \{x\}).
\end{aligned}$$

According to Conjecture 1, both permutations in each row generate the same 2-additive belief function.

Namely, having denoted by $\vec{m} = [m(x), m(y), m(\Omega)]'$ an arbitrary mass vector, the above pairs of permutations generate the following vertices:

$$\begin{aligned}
\rho_1, \rho_2 &\rightarrow [m(x) + m(\Omega), m(y), 0]' \\
\rho_3, \rho_4 &\rightarrow [m(x), m(y) + m(\Omega), 0]' \\
\rho_5, \rho_6 &\rightarrow [m(x), m(y), m(\Omega)]'.
\end{aligned} \qquad \text{(A.2)}$$

Figure A.5 depicts the belief space and the polytope $\mathfrak{B}^2[b]$ of 2-additive belief functions dominating a given belief function b

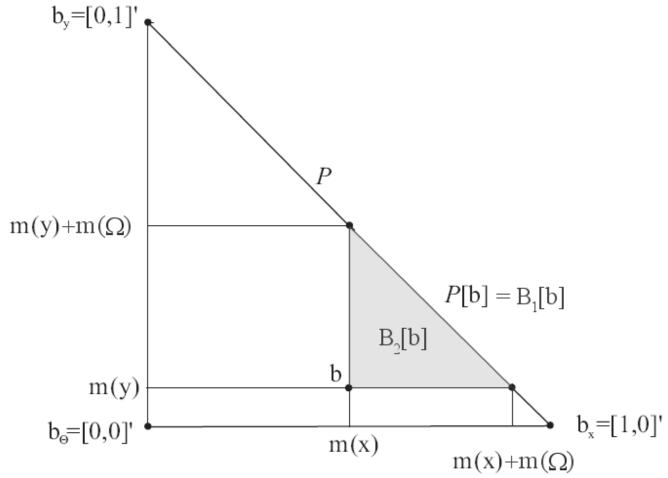

Figure A.5: The polytope $\mathfrak{B}^2[b]$ of the 2-additive belief functions dominating a given belief function b defined on a frame of size 2. The vertices of such polytope meet the conjectured form (17), and are given by the basic probability assignments of Equation (A.2).

for a frame Ω of cardinality 2. Here each belief function is a vector $b = [m(x), m(y)]'$ and $\mathfrak{B} = Cl(b_x, b_y, b_\Omega)$. As it can be appreciated, the last vertex in (A.2) of $\mathfrak{B}^2[b]$ corresponds to the original belief function b, while the first two are nothing but the vertices of the set $\mathfrak{B}^1[b]$ of dominating probabilities.

We can notice two facts: On one side, the Equation 17 seems confirmed by the analysis of the binary case with $k = 2$ (the cas where $k = 1$ corresponding to the pignistic transform). On the other side, unlike the case of dominating probabilities, there is no 1-1 correspondence between vertices of the polytope and the permutations of focal elements, as each vertex is produced by two different permutations. However, all vertices are associated with the same number of permutations. Finally, the conjecture holds in the binary case.

Appendix A.3. Theorem 1

Under the assumption that Conjecture 1 is true, the barycenter of $\mathfrak{B}^k[b]$ is

$$\sum_\rho \frac{b^\rho[b]}{\#\rho}.$$

By Equation (11), this corresponds to a mass function which assigns to each focal set A: $|A| \leq k$ the value:

$$\sum_{B \supseteq A} m(B) \frac{\#\rho : \forall A' <_\rho A : B \not\supset A'}{\#\rho}. \tag{A.3}$$

As in the proof of Property 3, the coefficient of $m(B)$ in the above equation is proportional to the number of permutations ρ of $\mathcal{P}^k(\Omega)$ such that B does not contain any element of $\mathcal{P}^k(\Omega)$ that comes before A in the permutation.

Obviously, there are $|\mathcal{P}^k(\Omega)|$ elements in $\mathcal{P}^k(\Omega)$. Of these, $|\mathcal{P}^k(\Omega)| - |\mathcal{P}^k(B)|$ are not included in B. Let $l = |B|$. Let us introduce for sake of simplicity the notation $\mathcal{M}(l,k) = \sum_{i=1}^{i=k} \binom{l}{i} = |\mathcal{P}^k(B)|$. As in Appendix A.1, for each position i of A, the number of admissible permutations is given by the dispositions

$$\frac{(\mathcal{M}(n,k) - \mathcal{M}(l,k))!}{[(\mathcal{M}(n,k) - \mathcal{M}(l,k)) - (i-1)]!}$$

of the $\mathcal{M}(n,k) - \mathcal{M}(l,k)$ subsets of size $\leq k$ which are not included in B over $i-1$ locations (the elements of the list before A), multiplied by the number $(\mathcal{M}(n,k) - i)!$ of permutations of the remaining $\mathcal{M}(n,k) - i$ elements of $\mathcal{P}^k(\Omega)$, which can appear after A in any order.

The same derivations of Appendix A.1 hold then for the case of dominating k-additive belief functions as well, when we replace $|\Omega|$ with $\mathcal{M}(|\Omega|, k)$ and $|B|$ with $\mathcal{M}(|B|, k)$. Therefore, the multiplicative coefficient of $m(B)$ in Equation (A.3) turn out to be $\frac{1}{\mathcal{M}(l,k)} = \frac{1}{|\mathcal{P}^k(B)|}$.

Appendix A.4. Proposition 2

Let us first consider the asymmetric transform. To establish this proposition, the simplest way is to consider the redistribution process of the asymmetric k-additive pignistic transform in the following two scenarios: First, when two consecutive transformations with thresholds k_1 and k_2 (with $k_1 > k_2$) are applied, and second, when a single transformation with the threshold $\min(k_1, k_2) = k_2$ is applied. Then, it is sufficient to check that the redistribution process of the mass attributed to a set of cardinality $> k_1$ leads to the same results in these two scenarios.

Let us consider A, a subset of Ω with $|A| > k_1$. In both scenarios $m(A)$ is redistributed to subsets of cardinality $\leq k_1$. Let us call B any subset of Ω such that $k_2 < |B| \leq k_1$, and C any subset with $|C| \leq k_2$.

In the first scenario, a single transform ($k = k_2$) is used. Each $C \subseteq \Omega$ with $|C| \leq k_2$ receives *directly* a number of parts of $m(A)$ which is, by definition, proportional to $|C|$: $H^A_{k_2}(C) \propto |C|$. In the second scenario, two transforms (first $k = k_1$, and then, $k = k_2$) are used. After the first transform, the sets C and B receive some part of $m(A)$. Then, after the second transform, the mass of the sets B is redistributed to the sets C. As the B have received some part of $m(A)$ after the first transform, these parts of $m(A)$ are redistributed to C after the second transform. Thus, C-type sets receive *directly* some mass from A (first transform) but also receive *indirectly* some mass from A that has transited via the sets B. If we note $H^{A\to B}_{k_1,k_2}(C)$ the mass that has transited from A, via B to C, we have that:

$$H^{A\to B}_{k_1,k_2}(C) \propto |C|.$$

This can be verified as, first we have $H^A_{k_1}(B) \propto |B|$, and then, for each B, $H^A_{k_1}(B)$ is shared and redistributed in a manner $\propto |C|$, which explains the previous equation. Hence, C's receive from A the mass:

$$\left(\underbrace{H^{A\to B}_{k_1,k_2}(C)}_{\propto|C|} + \underbrace{H^A_{k_2}(C)}_{\propto|C|} \right) \propto |C|.$$

Finally, it is easy to check that, whatever the scenario, C-type sets receive all the mass initially associated with A, so that it is shared among such C's in a manner proportional to their cardinality. As $m(A)$ and the sum of all the cardinality of the sets C is determined once and for all, both scenarios lead to the same mass redistribution. A similar demonstration holds for the geometric generalization.

Appendix A.5. Proposition 3

The proof is the same for the asymmetric and geometric transforms. Thus, we only consider the case of the geometric one. We need to show that, $\forall A \subseteq \Omega$, $b(A) \leq b_k^{[\mathcal{G}]}(A)$. By definition, $b(A) = \sum_{B\subseteq A} m(B)$ and $b_k^{[\mathcal{G}]}(A) = \sum_{B\subseteq A} m_k^{[\mathcal{G}]}(B)$.

Let us denote by $H_k^A(B)$ the mass inherited by B from A, and by $H_k(B)$ the total mass inherited by B from focal elements of cardinality k. Of course, we have

$$H_k(B) = m_k^{[GP]}(B) - m(B) = \sum_{A\supset B,\ |A|>k} H_k^A(B). \tag{A.4}$$

Moreover, by Equation (16), one has that:

$$H_k(B) = m_k^{[\mathcal{G}]}(B) - m(B) > 0 \text{ if } |B| \leq k,$$

as the terms $H_k(B)$ correspond to some mass inherited from focal elements of cardinality $> k$, redistributed to focal elements of cardinality $\leq k$. Now:
∘ If $|A| \leq k$, then, $b_k^{[\mathcal{G}]}(A) - b(A) = \sum_{B\subseteq A} H_k(B) > 0$.
∘ If $|A| > k$, then,

$$b_k^{[\mathcal{G}]}(A) = \sum_{\substack{B\subseteq A,\\ |B|\leq k}} m_k^{[\mathcal{G}]}(B) + \underbrace{\sum_{\substack{B\subseteq A,\\ |B|>k}} m_k^{[\mathcal{G}]}(B)}_{=0} = \sum_{\substack{B\subseteq A,\\ |B|\leq k}} m(B) + H_k(B) \tag{A.5}$$

According to the previous notation (A.4), it is possible to decompose $H_k(B)$ with respect to the origin of the mass received by B from all $C \subseteq \Omega$ s.t. $|C| > k$. Some of them are included in A, some others are not:

$$H_k(B) = \sum_{\substack{C\subseteq A\\ |C|>k}} H_k^C(B) + \sum_{\substack{C\not\subseteq A\\ |C|>k}} H_k^C(B)$$

so that

$$b_k^{[\mathcal{G}]}(A) = \sum_{\substack{B\subseteq A\\ |B|\leq k}} m(B) + \sum_{\substack{B\subseteq A\\ |B|\leq k}} \sum_{\substack{C\subseteq A\\ |C|>k}} H_k^C(B) + \sum_{\substack{B\subseteq A\\ |B|\leq k}} \sum_{\substack{C\not\subseteq A\\ |C|>k}} H_k^C(B). \tag{A.6}$$

Now we can notice that:

$$\sum_{\substack{B\subseteq A\\ |B|\leq k}} \sum_{\substack{C\subseteq A\\ |C|>k}} H_k^C(B) = \sum_{\substack{B\subseteq A\\ |B|>k}} m(B),$$

as the mass associated to subsets of A with cardinality $> k$ is redistributed to the subsets of A with cardinality $\leq k$. Thus,

$$b_k^{[\mathcal{G}]}(A) = \underbrace{\sum_{\substack{B\subseteq A\\ |B|\leq k}} m(B) + \sum_{\substack{B\subseteq A\\ |B|>k}} m(B)}_{b(A)} + \underbrace{\sum_{\substack{B\subseteq A\\ |B|\leq k}} \sum_{\substack{C\not\subseteq A\\ |C|>k}} H_k^C(B)}_{\geq 0}$$

i.e. $b_k^{[\mathcal{G}]}(A) \geq b(A)$, and $b \ll b_k^{[\mathcal{G}]}$.

Appendix A.6. Proposition 4

1) and 2) see [50].
3) • Existance and unicity: By construction.
• k-additivity: see [50].
• Dominance: Proposition 3.
4) By definition, we have:

$$\mathcal{A}[b_k^{[\mathcal{A}]}] = \left\{ \left(b_k^{[\mathcal{A}]}\right)^{[S]}, \left(b_k^{[\mathcal{A}]}\right)_2^{[\mathcal{A}]}, \cdots, \left(b_k^{[\mathcal{A}]}\right)_{k-1}^{[\mathcal{A}]}, \left(b_k^{[\mathcal{A}]}\right) \right\}$$

which by Proposition 2 reads:

$$\mathcal{A}[b_k^{[\mathcal{A}]}] = \left\{ b^{[S]}, b_2^{[\mathcal{A}]}, \cdots, b_{k-1}^{[\mathcal{A}]}, b_k^{[\mathcal{A}]} \right\}$$

On the other hand, we have:

$$\begin{aligned} \mathcal{A}[b] &= \left\{ b^{[S]}, b_2^{[\mathcal{A}]}, \cdots, b_{K-1}^{[\mathcal{A}]}, b \right\} \\ &= \left\{ \underbrace{b^{[S]}, b_2^{[\mathcal{A}]}, \cdots, b_{k-1}^{[\mathcal{A}]}, b_k^{[\mathcal{A}]}}_{\mathcal{A}[b_k^{[\mathcal{A}]}]}, b_{k+1}^{[\mathcal{A}]}, \cdots, b_{K-1}^{[\mathcal{A}]}, b \right\} \end{aligned}$$

Finally, $\mathcal{A}[b_k^{[\mathcal{A}]}] \subseteq \mathcal{A}[b]$. A similar proof holds for $\mathcal{G}[b_k^{[\mathcal{G}]}] \subseteq \mathcal{G}[b]$.
5) By Proposition 2, we have $b_{k_2}^{[\mathcal{A}]} = \left(b_{k_1}^{[\mathcal{A}]}\right)_{k_2}^{[\mathcal{A}]}$, and, by Proposition 3, we have $b_{k_1}^{[\mathcal{A}]} \ll \left(b_{k_1}^{[\mathcal{A}]}\right)_{k_2}^{[\mathcal{A}]}$, which leads to $b_{k_1}^{[\mathcal{A}]} \ll b_{k_2}^{[\mathcal{A}]}$. Again, a similar proof holds for $b_{k_1}^{[\mathcal{G}]} \ll b_{k_2}^{[\mathcal{G}]}$.
6) By direct application of 5).

Appendix A.7. Efficiency and Linearity arguments

Projectivity is immediate.
Anonimity is also immediate.
False Event is a direct consequence of Property 4.6: As the pignistic transform dominates both asymmetric and geometric k-additive pignistic transforms, and as $b^{[S]}$ assigns null mass to every false event, this is necessarily the case for our two generalizations as well.
Efficiency: The result of the two k-additive pignistic transforms is a belief function whose decision space is $\mathcal{P}_k(\Omega)$, and there are no focal elements outside $\mathcal{P}_k(\Omega)$. Hence, the mass assignments associated with all the element of the decision space $\mathcal{P}_k(\Omega)$ sum up to 1. Therefore, a bet on the entire such decision space is bound to be a winning bet.
Linearity: We only consider the asymmetric transform, as the proof is similar for the geometric one. Let us call $m^{[1]}$ and $m^{[2]}$ two mass functions, and $m_k^{[\mathcal{A}1]}$ and $m_k^{[\mathcal{A}2]}$ their corresponding asymmetric k-additive transforms. Let $\alpha, \beta \in [0, 1]$ be scalars such that $\alpha + \beta = 1$. We define their affine combination as $m^{[3]} = \alpha m^{[1]} + \beta m^{[2]}$, denote by $m_k^{[\mathcal{A}3]}$ the asymmetric k-additive transform of such affine combination $m^{[3]}$. The Linearity argument amounts to the following equality:

$$\alpha \cdot m_k^{[\mathcal{A}1]} + \beta \cdot m_k^{[\mathcal{A}2]} = m_k^{[\mathcal{A}3]}.$$

The latter can be easily verified, as $\forall B \in \mathcal{P}_k(\Omega)$, we have:

$$\begin{aligned}
&\alpha m_k^{[\mathcal{A}1]}(B) + \beta m_k^{[\mathcal{A}2]}(B) \\
&= \alpha\Big[m^{[1]}(B) + \textstyle\sum_{\substack{A\supset B,\\ A\subseteq\Omega,\\ |A|>k}} \frac{m^{[1]}(A)|B|}{\mathcal{N}(|A|,k)}\Big] + \beta\Big[m^{[2]}(B) + \textstyle\sum_{\substack{A\supset B,\\ A\subseteq\Omega,\\ |A|>k}} \frac{m^{[2]}(A)|B|}{\mathcal{N}(|A|,k)}\Big] \\
&= \alpha m^{[1]}(B) + \beta m^{[2]}(B) + \sum_{\substack{A\supset B, A\subseteq\Omega,\\ |A|>k}} \frac{[\alpha m^{[1]}(A) + \beta m^{[2]}(A)]|B|}{\mathcal{N}(|A|,k)} \\
&= m^{[3]}(B) + \sum_{A\supset B, A\subseteq\Omega, |A|>k} \frac{m^{[3]}(A)\cdot|B|}{\mathcal{N}(|A|,k)} = m_k^{[\mathcal{A}3]}(B).
\end{aligned} \tag{A.7}$$