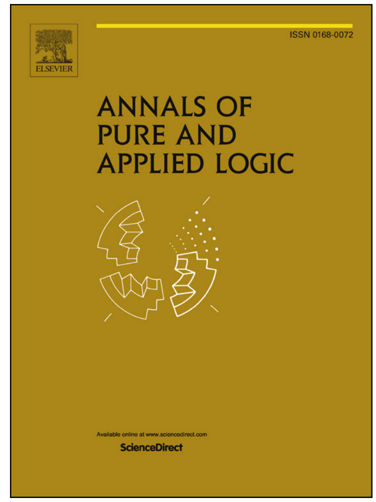
Journal Pre-proof



A logico-geometric comparison of coherence for non-additive uncertainty measures

Esther Anna Corsi, Tommaso Flaminio, Hykel Hosni

PII: S0168-0072(23)00099-4
DOI: https://doi.org/10.1016/j.apal.2023.103342
Reference: APAL 103342

To appear in: *Annals of Pure and Applied Logic*

Received date: 30 September 2021
Revised date: 5 April 2023
Accepted date: 10 July 2023

A logico-geometric comparison of coherence for non-additive uncertainty measures

Esther Anna Corsi*[1], Tommaso Flaminio†[2], and Hykel Hosni‡[1]

[1]Department of Philosophy, University of Milan, Italy
[2]Artificial Intelligence Research Institute, IIIA – Spanish National Research Council, CSIC

July 13, 2023

Abstract

We investigate the notion of coherence for (non-)additive uncertainty measures from a logico-geometric point of view. Our main result is to the effect that distinct criteria for coherence are not always matched by axiomatically distinct measures of uncertainty. In addition we introduce a metalogic within which this kind of result can be captured formally.

Keywords– Coherence, Probability logic, Belief Functions, Lower probability.

AMS Classification– 03B48, 60A05, 68T37.

1 Introduction and motivation

The investigation reported in this paper has its origins is [24, 21, 9] and identifies the common root of logic and probability in the notion of *coherence*. Our key research question is whether distinct criteria for coherence are always matched by axiomatically distinct measures of uncertainty. Moreover we ask whether this distinction, when available, can be expressed within a suitable metalogic.

In particular, we study the notion of coherence for several uncertainty measures by geometric means. By leveraging on this geometric representation we show that coherence is not sufficient to distinguish books on nontrivial sets of events that are extendible to lower probabilities and belief functions. The same geometric representation will be then used for a purely logical analysis of coherence that will be done by Riesz infinite-valued logic.

Since coherence features a (perhaps unique) combination of logico-mathematical, foundational and practical interest, we begin by helping the reader to appraise the wide landscape which the present work helps systematising. This will also serve the purposes of introducing some of the terminology and notation that will be used throughout the paper.

*esther.corsi@unimi.it (Corresponding Author)
†tommaso@iiia.csic.es
‡hykel.hosni@unimi.it

For the founders of mathematical logic, Boole and De Morgan, logic and probability belonged on equal terms to the mathematical analysis of sound inference. In the intervening two centuries though, the fields have grown largely independentently, as is apparent from textbook presentations. And yet there is much to be understood about the foundations and applications of reasoning under uncertainty by looking very closely at how logic and probability concur in forming a constellation of methods, models and axiomatisations [33, 37, 28].

Logic plays a twofold role in the foundations and applications of uncertain reasoning. Syntactically, it provides an unambiguous definition of the objects of an agent's reasoning, i.e. the algebra of events. Semantically, it provides the mechanism necessary for the agent to quantify the uncertainty which is not resolved by the information they possess. In the most familiar case of probability, two Boolean algebras are at work (see Subsection 2.1 below for basic terminology): $\mathbf{A}$, which formalises events, and $\mathbf{2}$, which formalises the values that indicator functions take. Homomorphisms v from $\mathbf{A}$ to $\mathbf{2}$ correspond to classical evaluations of formulas, i.e. in agreement with Tarskian semantics. Events which do not get a binary truth-value are the objects of the distribution of a probability mass function, which is done according to criteria exceeding the semantics of classical logic.

Against this background the classical problem of coherence can be put as follows:

Given

1. An idealised agent who is required to quantify their uncertainty concerning a (finite) subset of $\mathbf{A}$,
$$\Psi = \{a_1, a_2, \ldots, a_n\}.$$
2. Uncertainty is resolved by Tarskian semantics using evaluations as specified above.

Want the conditions under which the assignment

$$\beta : \Psi \ni a_i \mapsto \beta_i \in [0, 1]$$

can be said to be *coherent*.

In this context, the term "coherence" has been popularised by Bruno de Finetti [11, 12], who argued persuasively that the desired conditions should be represented by normalisation and finite additivity (see below), but as we shall recall, de Finetti's word was far from being the last one on the subject. Note that coherence bears the following desiderata. We require β to

1. represent the information (if any) available to the agent;
2. pin down the appropriate quantification of the uncertainty which is not resolved by information.

For present purposes, *information resolves uncertainty* to the extent that the (classical, propositional) sentences which represent it decide other sentences of interest. So for instance, the information represented by ϕ resolves (classically) any uncertainty about $\phi \vee \psi$. Note that the converse does not hold.

Thus the problem of coherence boils down to identifying the conditions under which β quantifies uncertainty "appropriately". This is an extra-logical requirement, which typically depends on the purpose for which we are considering a model of uncertainty quantification in the first place. Following a tradition which has played a crucial role in the development of the theory of probability, we restrict our attention to situations in which the agent's decision-making is spelt-out in terms of (highly abstract and idealised) betting behaviour. Depending on the nature of the betting problem, certain desiderata become particularly compelling, as discussed in Section 2, and summarised in Table 1.

Armed with this little terminology and notation, we can proceed to locate more precisely the present contribution relative to the wider area of probability logic. The problem of coherence as we recalled it, was initially motivated by foundational concerns which were absent from the coeval investigations on the mathematical problem of identifying the conditions for a Boolean algebra to carry a finitely

Key axiom	Betting	Uncertainty resolution
Additivity	2-sided	Classical
∞-monotonicity	2-sided	Partial
Super-additivity	1-sided	Classical

Table 1: A summary of the relation between the key axiomatic properties versus the characteristics of the decision problems and the corresponding uncertainty resolutions. The relevant definitions are provided in Section 2.

additive measure. This line of investigation is rooted in von Neumann's early work on $\sigma-$algebras dating back to the late 1930's, and goes through a conjecture of Horn's and Tarski's, its refutation due to Gaifman, and culminates with the 1959 representation of Kelley's – see [35] for a terse account.

In addition the measure theoretic questions, which does not constitute the focus of our work, the early 1980's witnessed an explosion of interest in probabilistic extensions of logical inference owing to the promises of expert systems (see [39] for a comprehensive historical overview). This brought the problem of combining logic and probability to the attention of the AI community see, e.g. [42, 20, 26, 28]. With a peculiar "applications-foundations feedback", a number of criticisms to the probabilistic representation of uncertainty became commonplace within AI. Among them to the inability of (finitely) additive measures of uncertainty to represent two important aspects of information processing systems, namely *vagueness* and *partial ignorance*.

The former concern arises because not all events of interest to an (idealised, artificially intelligent) agent have a binary realisation. To the contrary, many properties are best thought of as graded, from a person's age, to the business of a road junction. With this motivation, [58, 59] put forward an extension of classical logic, where truth-values in $[0, 1]$ are interpreted as "degrees of truth" of *fuzzy* sentences. Since fuzziness is a semantic (i.e. logical) property of the uncertainty resolution is distinct from a measure quantifying uncertainty. Hence it is mathematically meaningful and practically useful to understand how say, probability should be assigned to non-binary events. This lead to a renewed interest in the early 20th century approaches to many-valued logics, [29] and in particular the Łukasiewicz real-valued logic [41]. By the beginning of this century it became clear that the vast literature extending classical logics could be used in probabilistic uncertainty resolution. For our present purposes, a particularly important contribution in this respect is [44]. In it Paris shows that one of the key methods for justifying the quantification of uncertainty by means of probability – the Dutch Book method to be discussed at length below – does not necessitate Tarskian uncertainty resolution. The Riesz consequence relation which plays a central role in Section 4 below, sums up much of the unexpected interaction between many-valued logics and probability which has been motivated by Paris's note.

As to partial ignorance, the concern arises because probability may force agents to go unwarrantedly beyond the uncertainty resolved by Tarskian semantics. We illustrate with a classic problem popularised by Ellsberg [18].

Example 1.1. Consider an urn with red, blue and green balls. Suppose ψ_1 stands for "the ball is red", ψ_2 stands for "the ball is blue and ψ_3 stands for "the ball is green", Suppose further that the agent knows that the proportion of the red balls in the urn is 1/3. Representing this information probabilistically, leads to the straightforward quantification of the agent's uncertainty in the relevant event, i.e $P(\psi_1) = 1/3$. Observe now that the information available does not resolve "enough" uncertainty to allow the agent to come up with equally straightforward quantifications for $P(\psi_2)$ and $P(\psi_3)$. Any value in $[0, 2/3]$ will be consistent with the information available. In the absence of any further information, the probabilistic setting would lead to setting both $P(\psi_2)$ and $P(\psi_3)$ to 1/3. The problem with that is that 1/3 fails to represent the information to the effect that the information available does *not* support 1/3 any more than any value in $[0, 2/3]$.

A popular reaction to this kind of problem has been to weaken the *additivity* of probability functions.

Among the many proposals in this direction, two have proved to be particularly fruitful. Following earlier statistical work of Dempster, Shafer [49] sought to capture the uncertainty resolution provided by classical logic as determining the evidential support of a super-additive measure, which has become known as Dempster-Shafer belief function (to be defined below). The second non-additive approach to measuring the ignorance arising from partial uncertainty resolution, consists in taking (convex) sets of probability functions – the ones which are consistent with the information available. This results in so-called *credal sets* [38] which lead naturally to defining lower (and upper) probabilities for the events of interest. This approach, which is rooted in the investigation of Inner and Outer measures [30], has been championed by Walley [54]. Like belief functions, lower probabilities (to be defined below) are also super-additive.

In light of this rough and incomplete sketch of the landscape, it is not particularly surprising that a great variety of approaches to logic-based uncertain reasoning have been put forward over the past few decades, each following its own peculiar blend of mathematical, foundational and practical motivation. As a result it may not be always obvious which framework or measure of uncertainty is best suited to which kind of problem. The main aim of the present paper is to put forward logico-geometric tools that can contribute significantly to obtaining a unified picture. Far from being only a problem for applications, tying the properties of uncertainty measures with the kind of situations in which they are practically useful has a distinct foundational importance.

The remainder of paper is organised as follows. Section 2 reviews the measures of uncertainty of present interest (Subsection 2.2) and their associated notions of coherence (Subsections 2.3–2.5). Whilst the results of this Section are not novel, the way we present them can be of independent interest, in addition to anticipating, through numerical examples, some key geometric insights that will be used in our main results. Section 3 forms the core of our paper. In Subsection 3.1 we focus on the axiomatic comparison of lower probabilities and belief function. Using the geometric framework introduced in Subsection 3.2, Subsection 3.3 presents the main result of this paper, Theorem 3.10. In it we identify rather mild conditions under which belief functions and lower probability are coherence-wise indistinguishable despite being axiomatically distinct. Section 4 showcases a metatheoric role for many-valued logics in reasoning about coherent measures of uncertainty. In it we define the Riesz consequence relation and show in Theorem 4.10 how it can represent formally the geometric coherence-wise comparisons put forward in Subsection 3.3. As we shall observe in the concluding Section 5, this metalogical setting may prove useful in the abstract investigation of coherence-based uncertainty measures.

2 Preliminaries

This Section begins by recalling the basic definitions and results on Boolean algebras (Subsection 2.1) and of three key measures of uncertainty and their associated dual measures (Subsection 2.2). In it, we shall highlight the properties which will be of particular interest for our results. Then we review the original setup due to de Finetti (Section 2.3) and its extension to partly resolving uncertainty by Jaffray in Section 2.4 and to imprecise probabilities in Section 2.5. Note that this Section recalls material selected with the goal of making the paper essentially self-contained and does not attempt to do justice to the incredibly rich relevant literature, for which excellent surveys are available, including [43, 33, 37, 28, 52]. Again in the interest of brevity, some basic logico-algebraic notions are taken for granted. We refer, with apologies, readers who need to fill out the gaps to [7, 27, 47].

2.1 Finite Boolean algebras and their atoms

Finite Boolean algebras are the algebraic framework of this paper; hence its logical setting is that of classical propositional logic, CPL. Here, we will briefly recap on some needed notions and basic results about Boolean algebras and CPL, for a more exhaustive introduction about this subject we invite the reader to consult [47], and [6, 32, 31].

Given a countable (finite or infinite) set V of *propositional variables*, the CPL language $\mathsf{L}(V)$ (or simply L when V will be clear by the context) is the smallest set containing V and closed under the usual connectives $\wedge, \vee, \neg, \bot$, and $\top$ of type $(2, 2, 1, 0, 0)$. Along this paper we will use the notation φ, ψ, etc (with possible subscript) for formulas. Further, we shall adopt the following abbreviations:

$$\varphi \to \psi = \neg\varphi \vee \psi,\ \varphi \leftrightarrow \psi = (\varphi \to \psi) \wedge (\psi \to \varphi).$$

We shall denote by $\vdash_{CPL}$ the provability relation of CPL, in particular we will write $\vdash_{CPL} \varphi$ to denote that φ is a theorem.

A *logical valuation* (or simply a *valuation*) of L is a map v from V to the domain $\{0, 1\}$ of the Boolean algebra **2**, which uniquely extends to a function, that we denote by the same symbol v, from L to $\{0, 1\}$ in accordance with the usual Boolean truth functions, i.e. $v(\varphi \wedge \psi) = \min\{v(\varphi), v(\psi)\}$, $v(\bot) = 0$, $v(\neg\varphi) = 1 - v(\varphi)$, etc. We shall denote by Ω the set of all valuations of L. For a given formula φ and a given valuation $v \in \Omega$, we will write $v \models \varphi$ whenever $v(\varphi) = 1$.

We will broadly adopt, analogously to the above recalled logical frame, the signature $(\wedge, \vee, \neg, \bot, \top)$ of type $(2, 2, 1, 0, 0)$ for the algebraic language upon which Boolean algebras are defined. Thus, the same conventions and abbreviations of L can be adopted also in the algebraic setting. Further, in every Boolean algebra $\mathbf{A} = (A, \wedge, \vee, \neg, \bot, \top)$ we shall write $a \leqslant b$, whenever $a \to b = \top$. The relation $\leqslant$ is indeed the lattice-order in $\mathbf{A}$. Thus, $a \leqslant b$ iff $a \wedge b = a$ iff $a \vee b = b$.

Along this paper, in order to distinguish an algebra from its universe, we will denote the former by $\mathbf{A}$, $\mathbf{B}$ etc, and the latter by A, B etc, respectively.

Recall that a map $h : \mathbf{A} \to \mathbf{B}$ between Boolean algebras is a *homomorphism* if h commutes with the operations of their language, that is, $h(\top_{\mathbf{A}}) = \top_{\mathbf{B}}$, $h(\neg_{\mathbf{A}} a) = \neg_{\mathbf{B}} h(a)$, $h(a \wedge_{\mathbf{A}} b) = h(a) \wedge_{\mathbf{B}} h(b)$ etc (notice that we adopt subscripts to distinguish the operations of $\mathbf{A}$ from those of $\mathbf{B}$). Bijective (or 1-1) homomorphisms are called *isomorphisms* and if there is a isomorphism between $\mathbf{A}$ and $\mathbf{B}$, they are said to be *isomorphic* (and we write $\mathbf{A} \cong \mathbf{B}$).

Definition 2.1. An element α of a Boolean algebra $\mathbf{A}$ is said to be an *atom* of $\mathbf{A}$ if $\alpha > \bot$ and for any other element $b \in A$ such that $\alpha \geqslant b \geqslant \bot$, either $\alpha = b$ or $b = \bot$.

Every finite algebra has atoms that will be denoted by α, β, γ etc.

2.2 Uncertainty measures

As we already stated before, in this paper we only consider finite, and hence atomic, Boolean algebras. Those are the domains of the uncertainty measures we deal with.

Definition 2.2 (Probability functions)**.** A *probability function on an algebra* $\mathbf{A}$ is a $[0, 1]$-valued map P satisfying:

(P1) $P(\top) = 1$, $P(\bot) = 0$; (Normalisation)

(P2) $P(a \vee b) = P(a) + P(b)$, if $a \wedge b = \bot$. (Finite additivity)

The set of all probability functions on $\mathbf{A}$ is denoted by $\mathbb{P}^{\mathbf{A}}$. We omit the superscript when the algebra is clear from the context.

Remark 2.3. It is customary in measure-theoretic approaches to cast Definition 2.2 within a probability space, i.e. a countable set of elementary outcomes Ω with an algebra $\mathbf{A}$ on it, typically a field of sets. Then one defines a measure μ on $\mathbf{A}$ imposing normalisation on Ω and countable additivity, which generalises (P2) above to countable unions of (countably many incompatible) events. For our purposes it is sufficient to limit ourselves to finite additivity, noting that standard results, which start with de Finetti's own, show that there is no loss of generality in doing this, see [3] for a comprehensive overview. In addition, finite additivity provides the natural setting for probability logic, see Chapter 3 of [45].

Definition 2.4 (Belief functions [49])**.** A *belief function on an algebra* $\mathbf{A}$ is a $[0,1]$-valued map Bel satisfying:

(B1) $Bel(\top) = 1$, $Bel(\bot) = 0$;

(B2) $Bel\left(\bigvee_{i=1}^{n} a_i\right) \geqslant \sum_{i=1}^{n} \sum_{\{J\subseteq\{1,\ldots,n\}:|J|=i\}} (-1)^{i+1} Bel\left(\bigwedge_{j\in J} a_j\right)$, for $n \in \mathbb{N}$.

Remark 2.5. Note that (B2) does not take the form of a straightforward generalisation of (P2):

$$Bel(\theta \vee \phi) \geqslant Bel(\theta) + Bel(\phi), \quad \text{if} \quad \theta \wedge \phi = \bot. \qquad \text{(Finite super-additivity)}$$

Belief functions are indeed those super-additive set functions which are *completely* additive, or *monotone*, as it is sometimes said.

In the finite setting, belief functions on Boolean algebras can be characterized in terms of the associated *mass functions* as follows. Let $\mathbf{A}$ be any finite Boolean algebra with atoms $\alpha_1, \ldots, \alpha_t$. A mass function is a map m that assigns to each subset X of atoms (called *focal elements*) a real number such that $m(\varnothing) = 0$ and $\sum_X m(X) = 1$. Given a mass function m, the map

$$Bel(a) = \sum_{X\subseteq\{\alpha_i|\alpha_i\leqslant a\}} m(X) \tag{1}$$

is a belief function and every belief function on $\mathbf{A}$ arises from (1).

Example 2.6. Consider the finite Boolean algebra $\mathbf{A}$ with two atoms α_1 and α_2 and the following mass assignment:

$$\begin{aligned} m(\{\alpha_1\}) &= 0.2 \\ m(\{\alpha_2\}) &= 0.4 \\ m(\{\alpha_1, \alpha_2\}) &= 0.4 \\ m(\varnothing) &= 0. \end{aligned}$$

Since $m(\varnothing) = 0$ and $\sum_{X\subseteq A} m(X) = m(\{\alpha_1, \alpha_2\}) + m(\{\alpha_1\}) + m(\{\alpha_2\}) + m(\varnothing) = 1$, it is the case that m is a mass function. Moreover, since $\sum_{y\subseteq\{\alpha_1\}} m(y) = 0.2$ and $\sum_{y\subseteq\{\alpha_2\}} m(y) = 0.4$, the map $\beta : \alpha_1 \mapsto 0.2, \alpha_2 \mapsto 0.4$ is a belief function.

An element φ of a Boolean algebra $\mathbf{A}$ is said to be *covered m times by a multiset* $\{\{a_1, \ldots, a_n\}\}$ *of elements of A* if every homomorphism of $\mathbf{A}$ to $\{0,1\}$ that maps φ to 1, also maps to 1 at least m propositions from $a_1, \ldots, a_n$ as well. An (m,k)-cover of $(\varphi, \top)$ is a multiset $\{\{a_1, \ldots, a_n\}\}$ that covers $\top$ k times and covers φ $n + k$ times.

Definition 2.7 (Lower probability functions [51])**.** A *lower probability on an algebra* $\mathbf{A}$ is a monotone $[0,1]$-valued map $\underline{P}$ satisfying:

(L1) $\underline{P}(\top) = 1$, $\underline{P}(\bot) = 0$;

(L2) For all natural numbers n, m, k and all $a_1, \ldots, a_n$, if $\{\{a_1, \ldots, a_n\}\}$ is an (m,k)-cover of $(\varphi, \top)$, then $k + m\underline{P}(\varphi) \geqslant \sum_{i=1}^{n} \underline{P}(a_i)$.

Although this definition does not make the name *lower probabilities* particularly obvious, [1, Theorem 1] puts forward the following enlightening characterisation, anticipated by [51]. Let $\underline{P} : \mathbf{A} \to [0,1]$ be a lower probability and denote with $\mathcal{M}(\underline{P})$ the set of probability functions which bound $\underline{P}$ from above, i.e. $\mathcal{M}(\underline{P}) = \{P \in \mathbb{P} \mid \underline{P}(a) \leqslant P(a), \forall a \in A\}$. Then, for all $a \in A$,

$$\underline{P}(a) = \min_{P\in\mathcal{M}(\underline{P})} P(a). \tag{2}$$

Example 2.8. Let us consider the finite Boolean algebra $\mathbf{A}$ with two atoms α_1 and α_2 and the assignment $\beta\colon \bot \mapsto 0, \alpha_1 \mapsto 0.2, \alpha_2 \mapsto 0.4, \top \mapsto 1$. β is a lower probability. To see this note that $\mathcal{M}(\beta) = \{P \in \mathbb{P} \mid 0.2 \leqslant P(\alpha_1),\, 0.4 \leqslant P(\alpha_1)\}$. In particular, the probability functions P_1 and P_2 that belong to $\mathcal{M}(\beta)$ and that *generate* β are $P_1\colon \alpha_1 \mapsto 0.2, \alpha_2 \mapsto 0.8$ and $P_2\colon \alpha_1 \mapsto 0.6, \alpha_2 \mapsto 0.4$, respectively.

Remark 2.9. Belief functions and lower probabilities constitute the best-known axiomatic generalisations of probability functions. As it is apparent from the definitions above, probability functions are the *additive* special case of both belief functions and lower probabilities. Moreover, belief functions and lower probabilities can be related through probability functions. However, the question of interpreting this relation in terms of *coherence*, turns out to be remarkably difficult, as the main results of our paper illustrate in Section 3.

We end this subsection by recalling the definition of two ordinal measures of uncertainty which stand in duality relation to one another, and which will play an important role in what follows.

Definition 2.10 (Normalized necessity measures)**.** A *normalized necessity measure on an algebra* $\mathbf{A}$ is a $[0,1]$-valued map N satisfying:

(N1) $N(\top) = 1$, $N(\bot) = 0$;

(N2) $N(a_1 \wedge a_2) = \min\{N(a_1), N(a_2)\}$.

To each necessity measure on a Boolean algebra $\mathbf{A}$ is associated a dual *possibility measure*, usually denoted by Π and defined on $\mathbf{A}$ by letting

$$\Pi(a) = 1 - N(\neg a), \tag{3}$$

for all $a \in A$ – see [17, 33] for details.

Each possibility measure Π on a finite Boolean algebra $\mathbf{A}$ gives a *normalized possibility distribution* π once restricted to the atoms $\alpha_1, \ldots, \alpha_t$ of $\mathbf{A}$. In this context, normalization means that $\pi(\alpha_i) = 1$, for at least one α_i. Conversely, each normalized possibility distribution π on the α_i's uniquely determines a possibility Π and a necessity measure N by the following stipulations:

$$\Pi(a) = \bigvee_{j=1}^{t} \pi(\alpha_j) \wedge a(\alpha_j) \text{ and } N(a) = \bigwedge_{j=1}^{t} (1 - \pi(\alpha_t)) \vee a(\alpha_j), \tag{4}$$

where $a(\alpha_j)$ stands for 1 if $\alpha_j \leqslant a$ and 0 otherwise.

Remark 2.11 (Dual uncertainty measures)**.** In the same way as possibility measures can be defined by duality in the sense of equation (3), *dual companions* (also known as *conjugate measures*) can be defined also for the other uncertainty measures we considered so far. More precisely, while probability functions and self-dual in the sense that every probability P on a Boolean algebra $\mathbf{A}$ satisfies $P(a) = 1 - P(\neg a)$ for all $a \in A$, the following cases arise for non-additive measures:

- For every belief function Bel on $\mathbf{A}$, the map $Pl : \mathbf{A} \to [0,1]$ such that for all $a \in A$, $Pl(a) = 1 - Bel(\neg a)$ is called *plausibility function.*
- The dual companion of a lower probability $\underline{P} : \mathbf{A} \to [0,1]$, defined as $\overline{P}(a) = 1 - \underline{P}(\neg a)$ is known in the literature as an *upper probability function.*

Needless to say that, in general, Bel and Pl, and $\underline{P}$ and $\overline{P}$ differ on a Boolean algebra. Indeed, if on a Boolean algebra $\mathbf{A}$, one has that for all $a \in A$, $Bel(a) = Pl(a)$ (or $\underline{P}(a) = \overline{P}(a)$), then Bel ($\underline{P}$ respectively) is a probability function. More details can be found in [33].

The diagram depicted in Figure 1 sums up the mutual relations between the uncertainty measures and their duals described in this Subsection.

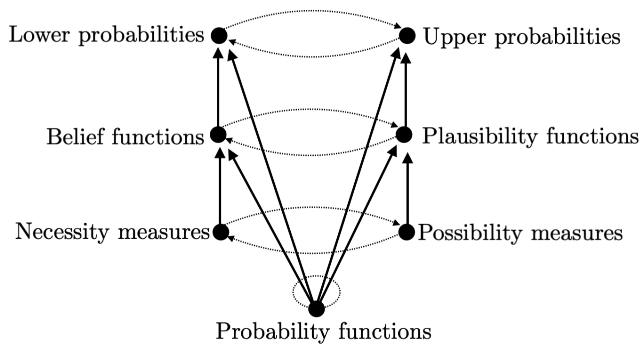


Figure 1: The uncertainty measures considered so far arranged by generality (solid arrows) and their dual companions (dashed arrows).

2.3 Two-sided betting with fully resolvable uncertainty

As anticipated in Section 1, the main focus of this paper is on coherence. As this notion has been the trademark of the foundational point of view championed by Bruno de Finetti, we begin by recalling the framework he laid down in his seminal [11].

Suppose that $a_1, \ldots, a_n$ are elements of a finite Boolean algebra $\mathbf{A}$, which are interpreted as the events of interest to a bookmaker $\mathbf{B}$. Suppose that this interest materialises with the publication of a *book* $\beta\colon a_1 \mapsto \beta_1, \ldots, a_n \mapsto \beta_n$ where $\beta_i \in [0,1]$ for $i = 1, \ldots, n$. This assignment is made by $\mathbf{B}$ under a number of constraints. The most important of which forces $\mathbf{B}$ to let a gambler $\mathbf{G}$ choose real-valued stakes $\sigma_1, \ldots, \sigma_n$ for each a_i in the book. Hence, for $i = 1, \ldots, n$, $\mathbf{G}$ pays $\sigma_i \beta_i$ to $\mathbf{B}$ in return for $\sigma_i v(a_i)$. Thus, $\mathbf{G}$'s payoff is $\sum_{i=1}^{n} \sigma_i (v(a_i) - \beta_i)$ whereas $\mathbf{B}$'s payoff is $\sum_{i=1}^{n} \sigma_i (\beta_i - v(a_i))$.

Remark 2.12. Throughout v is a (Boolean algebra) homomorphism from $\mathbf{A}$ to $\{0,1\}$ which agrees, that is, with the classical propositional semantics. Hence uncertainty is resolved classically – and we use the notation as a reminder that Boolean algebra homomorphisms play the role of classical propositional valuations (recall Subsection 2.1). We referred to this situation as the "classical problem of coherence" in Section 1.

This set-up is sufficient for de Finetti to put forward a definition which has had a profound impact on the foundations of probability since the second half of the nineteenth century.

Definition 2.13 (Coherence)**.** β is *coherent* if there is no $a_1, \ldots, a_n \in A$ and $\sigma_1, \ldots, \sigma_n$ in $\mathbb{R}$ such that for every v,

$$\sum_{i=1}^{n} \sigma_i (\beta_i - v(a_i)) < 0. \quad (5)$$

Equivalently, β is coherent if for every $a_1, \ldots, a_n \in A$ and $\sigma_1, \ldots, \sigma_n \in \mathbb{R}$, there is a v such that,

$$\sum_{i=1}^{n} \sigma_i (\beta_i - v(a_i)) \geqslant 0. \quad (6)$$

In other words, the book published by $\mathbf{B}$ is *coherent* if there is no choice of events and of (possibly negative) stakes which $\mathbf{G}$ can make, exposing $\mathbf{B}$ to a sure loss.

As shown in [11, 12], coherence is necessary and sufficient for the existence of a finitely additive measure P that extends the book β over $\mathbf{A}$, i.e. a probability function P such that for $i = 1, \ldots, n$, $P(a_i) = \beta_i$. De Finetti sought to use this result to justify the representation of "rational degrees of belief" by means of probabilities, a line of reasoning which has become known as the Dutch Book argument. There are many more details to this argument, which are immaterial for our present purposes (we refer interested readers to [23] for a logical presentation consonant to the present setting). One key point though, is that de Finetti concocted his argument in such a way that **B** avoids sure loss exactly by publishing a *fair price* book, i.e. one carrying null expectation of either gain or loss.

Fair prices are among the very first devices which have been used to give meaning to probability. They reflect an idealised and abstract situation in which every aspect of the book is perfectly known to **B**, except of course, which events will eventually obtain. As recalled in Section 1 several authors, starting with the early contributions by Frank Knight and John M. Keynes, have taken issue with the unattainable demands of such idealisations. Over the past century, a rich landscape of generalisations of probability measures has arisen in response to such foundational as well as practical concerns, with a particularly interesting line of research taking place within AI. As recalled above, the motivations, as well as the mathematical approaches pursued, diverge significantly and yet the relaxation of additivity is one important feature that wide array of formalisms have in common. Since our main focus in Section 3 will be on the relation between lower probabilities and belief functions – two prominent non-additive measures – we now review how those arise naturally by suitably relaxing some conditions in de Finetti's argument.

2.4 Two-sided betting with partly resolvable uncertainty

The first generalisation we consider is due to J-Y Jaffray who in [34], characterised coherent degrees of belief under *partly resolved uncertainty.* This led to a representation for belief functions which he linked to the then-emerging field of decision theory under ambiguity (see [53] for a thorough contextualisation).

In Jaffray's setting, coherence is defined essentially as in Definition 2.13 above, except that the underlying uncertainty resolution mechanism is no longer provided by classical logic. Instead of homomorphisms v, uncertainty is resolved by functions $C_a(a_i)$ defined as follows:

$$C_a(a_i) = \begin{cases} 1 & \text{if } \models_{cl} a \to a_i \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

So, for events $\Psi = \{a_1, \ldots, a_n\}$ the book $\beta\colon a_1 \mapsto \beta_1, \ldots, a_n \mapsto \beta_n$ is published as above by bookmaker **B**. For the gambler **G** to place real-valued stakes $\sigma_1, \ldots, \sigma_n$ on $a_1, \ldots, a_n$ at the betting odds written in β, means that **G** pays **B** for each a_i the amount $\sigma_i\beta_i$ and **B** gains from **G** the amount $\sigma_i C_a(a_i)$. Hence, the total balance for **B** is now

$$\sum_{i=1}^{n} \sigma_i(\beta_i - C_a(a_i)).$$

Definition 2.14 (Bf-coherence)**.** The book β is *bf-coherent* if it is not the case that, for every fixed non-contradictory event a, $\sum_{i=1}^{n} \sigma_i(\beta_i - C_a(a_i)) < 0$, i.e. there is a fixed non-contradictory event a s.t. $\sum_{i=1}^{n} \sigma_i(\beta_i - C_a(a_i)) \geqslant 0$.

Note that, in full analogy with Definition 2.13 this leads to a verbalisation of coherent degrees of belief in terms of avoiding sure loss. Hence the "intended semantics" of coherence is unchanged. What does change is the uncertainty resolution which underlies this formulation of the problem of coherence. Generalising classical valuations to (7), avoiding sure loss characterises Dempster-Shafer belief functions.

Theorem 2.15 ([34])**.** *An assignment β on events $\phi_1, \ldots, \phi_k \in A$ is* bf-coherent *if and only if it can be extended to a belief function on the Boolean algebra A.*

The following example, in addition to illustrating numerically coherent belief functions, provides a useful anticipation for our own framework.

a	$C_a(a_1)$	$C_a(a_2)$	$C_a(a_3)$
α_1	1	0	1
α_2	1	1	0
α_3	0	1	1
$\alpha_1 \vee \alpha_2$	1	0	0
$\alpha_2 \vee \alpha_3$	0	1	0
$\alpha_1 \vee \alpha_3$	0	0	1
$\alpha_1 \vee \alpha_2 \vee \alpha_3$	0	0	0

Table 2: Values of $C_a(a_i)$ over $\mathbf{A}$ and the set of events Ψ

Example 2.16. Let us consider an algebra $\mathbf{A}$ with atoms $\alpha_1, \ldots, \alpha_t$ ($t \geqslant 3$) and probability distributions $p_1(\alpha_1) = q$, $p_1(\alpha_2) = 1-q$, $p_1(\alpha_i) = 0$ for all $i \neq 1, 2$; $p_2(\alpha_2) = q$, $p_2(\alpha_3) = 1-q$, $p_2(\alpha_i) = 0$ for all $i \neq 2, 3$; $p_3(\alpha_1) = 1-q$, $p_3(\alpha_3) = q$, $p_3(\alpha_i) = 0$ for all $i \neq 1, 3$ where q is any value $1/3 < q \leqslant 1/2$. Suppose that $\mathbf{A}$ has three atoms $\alpha_1, \alpha_2, \alpha_3$. Let us consider the set of events $\Psi = \{a_1, a_2, a_3\}$ defined as in the proof of Theorem 3.10: $a_1 = \alpha_1 \vee \alpha_2, a_2 = \alpha_2 \vee \alpha_3$, and $a_3 = \alpha_1 \vee \alpha_3$. To show that the book $\beta : a_i \mapsto q$ for every $i = 1, 2, 3$ is bf-coherent, we need to show that for every stakes $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}$ chosen by $\mathbf{G}$ there is at least one non-contradictory event a s.t. $\sum_{i=1}^{n} \sigma_i(q - C_a(a_i)) \geqslant 0$. The values the function $C_a(a_i)$ are summarised in Table 2.16.

For every non-contradictory event $a \in \mathbf{A}$ we ask under which constraints on σ_1, σ_2 and σ_3, do we have $\sum_{i=1}^{n} \sigma_i(q - C_a(a_i)) \geqslant 0$. We denote with S_a the points in $\mathbb{R}^3$ (representing the values of the σ_is) that makes the payoff for $\mathbf{B}$ positive. We then show that the union of all these S_a with $a \in \mathbf{A}$ is $\mathbb{R}^3$.

If $a = \alpha_1$, then the payoff of $\mathbf{B}$ is $\sigma_1(q-1) + \sigma_2(q) + \sigma_3(q-1)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_1} = \{(x, y, z) \in \mathbb{R}^3 |\ z \leqslant \frac{-qx-qy+x}{q-1}\}$.

If $a = \alpha_2$, then the payoff of $\mathbf{B}$ is $\sigma_1(q-1) + \sigma_2(q-1) + \sigma_3(q)$ and it is greater than 0 if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_2} = \{(x, y, z) \in \mathbb{R}^3 |\ z \geqslant \frac{-qx-qy+x+y}{q}\}$.

If $a = \alpha_3$, then the payoff of $\mathbf{B}$ is $\sigma_1(q) + \sigma_2(q-1) + \sigma_3(q-1)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_3} = \{(x, y, z) \in \mathbb{R}^3 |\ z \leqslant \frac{-qx-qy+y}{q-1}\}$.

If $a = \alpha_1 \vee \alpha_2$, then the payoff of $\mathbf{B}$ is $\sigma_1(q-1) + \sigma_2(q) + \sigma_3(q)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_1 \vee \alpha_2} = \{(x, y, z) \in \mathbb{R}^3 |\ z \geqslant \frac{-qx-qy+x}{q}\}$.

If $a = \alpha_2 \vee \alpha_3$, then the payoff of $\mathbf{B}$ is $\sigma_1(q) + \sigma_2(q-1) + \sigma_3(q)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_2 \vee \alpha_3} = \{(x, y, z) \in \mathbb{R}^3 |\ z \geqslant \frac{-qx-qy+y}{q}\}$.

If $a = \alpha_1 \vee \alpha_3$, then the payoff of $\mathbf{B}$ is $\sigma_1(q) + \sigma_2(q) + \sigma_3(q-1)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_1 \vee \alpha_3} = \{(x, y, z) \in \mathbb{R}^3 |\ z \leqslant \frac{-qx-qy}{q-1}\}$.

If $a = \alpha_1 \vee \alpha_2 \vee \alpha_3$, then the payoff of $\mathbf{B}$ is $\sigma_1(q) + \sigma_2(q) + \sigma_3(q)$ and it is positive if $(\sigma_1, \sigma_2, \sigma_3) \in S_{\alpha_1 \vee \alpha_2 \vee \alpha_3} = \{(x, y, z) \in \mathbb{R}^3 |\ z \geqslant -x-y\}$.

To show that $\bigcup_{a \in \mathbf{A}} S_a = \mathbb{R}^3$ we only need to verify that $\bigcap_{a \in \mathbf{A}} S_a^c = \emptyset$, i.e. that the system of inequalities, as in this case,

$$\begin{cases} z > \frac{-qx-qy+x}{q-1} \\ z < \frac{-qx-qy+x+y}{q} \\ z > \frac{-qx-qy+y}{q-1} \\ z < \frac{-qx-qy+x}{q} \\ z < \frac{-qx-qy+y}{q} \\ z > \frac{-qx-qy}{q-1} \\ z < -x-y, \end{cases}$$

has no real solution.

2.5 One-sided betting with fully resolvable uncertainty

In his seminal [55], Walley termed "Bayesian dogma of precision" the conceptual underpinning of the conditions which force **B** to map each element of a book to a real number. Coherence-wise, we have recalled in Section 2.3 that this amounts to forcing **B** to publish fair prices for the book β. For if the book had a non-zero expectation, **G** will use this to lead **B** to sure loss. Thus fairness amounts two-sided betting: coherent degrees of belief correspond to those which make the bookmaker indifferent between either selling or buying any bets at the chosen prices. Showing that a mathematically sophisticated theory of uncertainty could dispense with this assumption, is one of the key achievements of [55].

Since then, considerable attention has been devoted to showing that the ensuing generalisations of (6) can be used to provide a foundation to a much broader class of imprecise probability measures (see e.g.[56, 40, 52]). Although the *Imprecise probability* community does not regard Walley's work as a generalisation of de Finetti's, as pointed out e.g. in Section 5 of [40], there is substantial foundational continuity between the two, as we are about to recall.

Note first that is customary in this field to frame coherence in terms of individual, rather than the interactive kind of decision problems recalled in Subsections 2.3 and 2.4 above. Hence coherence is spelt-out in terms of *desirable gambles*, which are thought of as the individual's expectation of a non-negative real-valued quantity X_i. In this context, a *lower prevision* $\underline{P}$ on a linear subspace $\mathcal{L}$ is said to be *coherent* if, for every non-negative integers m, n and for every $X_0, X_1, \ldots, X_n \in \mathcal{L}$ do we have

$$\sup \left\{ \sum_{j=1}^{n} (X_j - \underline{P}(X_j)) - m(X_0 - \underline{P}(X_0)) \right\} \geqslant 0. \tag{8}$$

Walley proves that a lower prevision is *coherent* if and only if for all $X, Y \in \mathcal{L}$ and $\lambda > 0$, the following hold.

$$\begin{aligned} \underline{P}(X) &\geqslant \inf X \\ \underline{P}(\lambda X) &= \lambda \underline{P}(X) \\ \underline{P}(X+Y) &\geqslant \underline{P}(X) + \underline{P}(Y). \end{aligned}$$

This characterisation, of which several equivalent formulations exist makes the theory of lower previsions one of the most general approaches to imprecise probabilities, as testified by [52]. However, we shall not need to exploit this generality here, starting from the fact that instead of gambles we shall consider their indicators a. Hence we shall consider lower *probabilities*, as anticipated with Definition 2.7. As an immediate consequence of this, and in contrast with the setup of Subsection 2.4, we are now back to a case in which uncertainty is resolved classically.

So what distinguishes the notion of coherence to be defined for lower probabilities compared to their additive counterpart, is a key feature of the decision problem which gives coherence its operational meaning. With an insight which goes back to Smith [51], this is achieved by interpreting $\underline{P}(a_i)$ as an agent's *supremum buying price* for a_i, and $\overline{P}(a_i)$ as (the possibly distinct) agent's *infimum selling price* for it. If one takes the two agents to coincide, then the interval $[\underline{P}(a_i), \overline{P}(a_i)]$ becomes a natural candidate for the quantification of the agent's *partial ignorance* about a_i. If those two bounds arise from distinct agents, then the above difference can be used to measure their disagreement about their individual assessment on a_i. In both cases, this interval allows us to quantify what the "Bayesian dogma of precision" neglects: some form of uncertainty about a_i exceeding its variability. From a foundational point of view, this answers some of the key methodological issues raised against probability which have been outlined in Section 1 above. In particular, it offers a very natural solution to the problem raised in Example 1.1 by putting, say $\underline{P}(a_2) = 0$ and $\overline{P}(a_2) = 2/3$.

The formulation of coherence with arises by distinguishing buying from selling prices, makes it possible to adapt de Finetti's Dutch Book argument to imprecise probabilities. All we need to do is relaxing the assumption to the effect that the interaction between **B** and **G** amounts to a zero-sum game. This can be achieved by restricting **G**'s choice of stakes only to *positive* σ_i.

Under such asymmetric conditions, the interactive decision problem which leads to the intended interpretation of coherence of lower probabilities can be put schematically as follows.

B publishes a book $\beta\colon a_1 \mapsto [\underline{\beta}_1, \overline{\beta}_1], \ldots, a_n \mapsto [\underline{\beta}_n, \overline{\beta}_n]$ over the events $a_1, \ldots, a_n$ with $0 \leqslant \underline{\beta}_i \leqslant \overline{\beta}_i \leqslant 1$ for $i = 1, \ldots, n$.

Then **G** works out the greatest price at which they will be willing to buy a_i, i.e. $\underline{P}(a_i)$, and decides whether to bet *on* or *against* each event:

- If **G** bets on the event a_i, then she will pay $\sigma_i \overline{\beta}_i$ with $\sigma_i \geqslant 0$ and receive from **B** $\sigma_i v(a_i)$.
- If **G** decides to bet against a_i, then she will receive $\sigma_i \underline{\beta}_i$ with $\sigma_i \geqslant 0$ from **B** and pay her back $\sigma_i v(a_i)$.

Let I_O denote the set of indexes relative to the events that **G** bets on, and denote by I_A the set of indexes relative to the events **G** bets against. Then, the balance for **G** is given by

$$\sum_{i \in I_O} \sigma_i(v(a_i) - \overline{\beta}_i) + \sum_{i \in I_A} \sigma_i(\underline{\beta}_i - v(a_i)). \tag{9}$$

Breaking the symmetry between buying and selling prices is not sufficient to extend the setup based on Definition 2.13 to the imprecise case. In this setting, agents may exceed in caution in the sense that the interval $[0, 1]$ clearly prevents sure loss under all circumstances. However, representing one's uncertainty on an event of interest with the whole real unit interval amounts to expressing one's *total* rather than partial ignorance about it. This suggests that avoiding sure loss is a necessary, yet no longer sufficient condition for the imprecise probability extension of Dutch Book. Indeed, in Section 1.6 of [55] Walley distinguishes between *avoiding sure loss* and *coherence*. Expressed in terms of preferences among gambles (rather than events, see above) – the key difference between the two notions is that the former amounts to acyclicity and the latter to transitivity. Owing to the duality between lower and upper probabilities, it will be sufficient to our purposes to focus on the former. Moreover, to avoid terminological confusion, in what follows we will refer the corresponding notion of coherence as to *l-coherence*.

Definition 2.17 (l-coherence)**.** Let $\beta\colon a_1 \mapsto \beta_1, \ldots, a_n \mapsto \beta_n$ be a book defined over the set of events $\Psi = \{a_1, \ldots, a_n\}$. Then β is *l-incoherent* for **G** if there exists an event $a_{i*} \in \Psi$, $\sigma_1, \ldots, \sigma_n \geqslant 0$ and $m \geqslant 0$ such that $\sum_{i=1}^{n} \sigma_i(\beta_i - v(a_i)) - m(\beta_{i*} - v(a_{i*})) < 0$ for every valuation v. The book β is *l-coherent* if it is not l-incoherent.

From the above definition, it follows that whenever $m = 0$, Definition 2.17 is equivalent to Definition 2.13. Hence, l-coherence is a stronger criterion compared to avoiding sure loss, for it additionally requires that the prices cannot be raised by any positive linear combination of alternative bets on the book, see e.g. Section 2.2.1 of [2].

Theorem 2.18 ([55])**.** *A book β on events $\phi_1, \ldots, \phi_n \in A$ is* l-coherent *if and only if it can be extended to a lower probability on the Boolean algebra A.*

It is of particular interest for our present purposes that, in the special case of lower probabilities, l-coherence is to Definition 2.13 as additivity is to super-additivity. Section 3.1 below develops this in full detail. In preparation for it, it is useful to illustrate some key properties of l-coherence, with which we close the preparatory material to our main results.

Example 2.19. Take again the setting of Example 2.16. For every $j = 1, \ldots, 3$, denote by P_j the probability given by the distribution p_j and let $\underline{P}$ the lower probability defined by letting

$$\underline{P}(a) = \min\{P_j(a) \mid j = 1, 2, 3\}, \tag{10}$$

for all $a \in A$.

To show that β can be extended to a lower probability we have to show that for every $a_i \in \Psi$, for every possible subset of $\Psi \backslash \{a_i\}$ and for every positive stake on these events, there is a homomorphism v that makes the payoff for $\mathbf{G}$ relative to the subset of $\Psi \backslash \{a_i\}$ greater or equal to the payoff relative to a_i. To do this we first show that a_1 is not l-coherent for $\mathbf{G}$, and similarly for a_2 and a_2. Therefore, we need to verify that for every $\sigma_1, \sigma_2, \sigma_3$ in $\mathbb{R}_{\geqslant}$, there exist valuations v_1, v_2, v_3 such that

1. $\sigma_2(v_1(a_2) - q) + \sigma_3(v_1(a_3) - q) \geqslant \sigma_1(v_1(a_1) - q)$,
2. $\sigma_2(v_2(a_2) - q) \geqslant \sigma_1(v_2(a_1) - q)$, and
3. $\sigma_3(v_3(a_3) - q) \geqslant \sigma_1(v_3(a_1) - q)$.

We reason by cases as follows.

1. If we consider the homomorphism $v_1 \colon \alpha_i \mapsto 0$, then $\sigma_2(v_1(a_2)-q)+\sigma_3(v_1(a_3)-q) \geqslant \sigma_1(v_1(a_1)-q)$ holds, i.e. $\sigma_2(-q) + \sigma_3(-q) \geqslant \sigma_1(-q)$, only if $\sigma_2 + \sigma_3 \leqslant \sigma_1$. If we consider the homomorphism $v_2 \colon \alpha_1 \mapsto 1, \alpha_2 \mapsto 1, \alpha_3 \mapsto 0$, then $\sigma_2(1-q) + \sigma_3(1-q) \geqslant \sigma_1(1-q)$ holds only if $\sigma_2 + \sigma_3 \geqslant \sigma_1$ (since $q \in (1/3, 1/2]$, then $1 - q \in [1/2, 2/3)$). Therefore, for every value of $\sigma_1, \sigma_2, \sigma_3$ we can find a homomorphism v such that (1) holds.
2. If we consider the homomorphism $v_1 \colon \alpha_i \mapsto 0$, then $\sigma_2(v_1(a_2) - q) \geqslant \sigma_1(v_1(a_1) - q)$ holds, i.e. $\sigma_2(-q) \geqslant \sigma_1(-q)$, only if $\sigma_2 \leqslant \sigma_1$. If we consider the homomorphism $v_2 \colon \alpha_1 \mapsto 0, \alpha_2 \mapsto 1, \alpha_3 \mapsto 0$, then $\sigma_2(1-q) \geqslant \sigma_1(1-q)$ only if $\sigma_2 \geqslant \sigma_1$. Therefore, for every value of $\sigma_1, \sigma_2, \sigma_3$ we can find a homomorphism v such that (2) holds.
3. If we consider the homomorphism $v_1 \colon \alpha_i \mapsto 0$, then $\sigma_3(v_1(a_3) - q) \geqslant \sigma_1(v_1(a_1) - q)$ holds, i.e. $\sigma_3(-q) \geqslant \sigma_1(-q)$, only if $\sigma_3 \leqslant \sigma_1$. If we consider the homomorphism $v_2 \colon \alpha_1 \mapsto 1, \alpha_2 \mapsto 0, \alpha_3 \mapsto 0$, then $\sigma_3(1-q) \geqslant \sigma_1(1-q)$ holds only if $\sigma_3 \geqslant \sigma_1$. Therefore, for every value of $\sigma_1, \sigma_2, \sigma_3$ we can find a homomorphism v such that (3) holds.

Analogous arguments hold for a_2 and a_3, hence β can be extended to a lower probability.

In conclusion, it is useful to point out explicitly that the duality which arises from the definitions of uncertainty measures discussed in Remark 2.11 has a natural counterpart in coherence.

Remark 2.20 (Dual coherence)**.** It is indeed clear that $\beta : a_1 \mapsto \beta_1, \ldots, a_n \mapsto \beta_n$ is coherent (bf-coherent, l-coherent, respectively), then the *dual* book $\beta' : \neg a_1 \mapsto 1 - \beta_1, \ldots, \neg a_n \mapsto 1 - \beta_n$ extends to a probability function (plausibility function, upper probability, respectively). This simple observation allows to immediately define notions of *pl-coherence* (for extensions to plausibility functions) and *u-coherence* (as regards to extensions to upper probabilities).

3 Axiomatic-wise and coherence-wise comparisons

We are now ready to address the key question of this paper, namely the comparison of the two key non-additive measures of uncertainty, namely belief functions and lower probabilities. In Subsection 3.1 we put forward an axiomatic comparison whereas in Subsection 3.2 we set up a geometric framework in which all the various notions of coherence of interest are comparable. This allows us to get, in Subsection 3.3, the main result of this paper, which identifies rather mild conditions under which belief functions and lower probability are coherence-wise indistinguishable despite being axiomatically distinct.

3.1 An axiomatic-wise comparison

As anticipated in Remark 2.9 above, belief functions and lower probabilities are related through probability functions. Say that $P \in \mathbb{P}$ *is compatible with a belief function Bel* if

$$Bel(a) \leqslant P(a) \leqslant Pl(a), \tag{11}$$

where Pl is a *plausibility function* defined for $a \in A$ by $Pl(a) = 1 - Bel(\neg a)$. Relation (11) lends itself to giving Belief functions a natural "imprecise improbability" interpretation. Just note that the length of the interval $[Bel(a), Pl(a)]$ is maximal when $Bel(a) = 0$ and $Pl(a) = 1$, and minimal when $Bel(a) = Pl(a)$. Note also that if this holds for all a, then Bel is a probability function.

Now, the set of probability functions which bound a belief function

$$\mathbb{P}(Bel) = \{P \in \mathbb{P} \mid P \text{ is compatible with } Bel\},$$

is called the *credal set of* Bel, and for any a, $Bel(a)$ is the lower envelope of $\mathbb{P}(Bel)$:

$$Bel(a) = \min_{P \in \mathbb{P}(Bel)} P(a). \tag{12}$$

As pointed out in [50, Section 6], this interpretation is justified *only* for a single belief function, i.e. one defined by a specific mass function m. This warns against considering in general belief functions as the lower bounds of some unknown true probability function. Nonetheless (2) and (12) highlight the close connection between the two. Indeed by constructing a suitable compatible belief function it can be seen that every belief function is a lower probability. But, as Williams [57] noted, the converse does not hold (see also [50, 36]). The conditions under which a lower probability function is a Belief function can be pinned down precisely as follows.

Remark 3.1. A lower probability $\underline{P}$ on an algebra $\mathbf{A}$ is a belief function if and only if $\underline{P}$ satisfies (B2), namely

$$\underline{P}\left(\bigvee_{i=1}^{n} a_i\right) \geqslant \sum_{i=1}^{n} \sum_{\{J \subseteq \{1,\ldots,n\}:|J|=i\}} (-1)^{i+1} \underline{P}\left(\bigwedge_{j \in J} a_j\right) \tag{13}$$

for all $n = 1, 2, \ldots$.

As an immediate consequence we get the following Corollary, which gives a minimal algebraic requirement to distinguish belief functions and lower probabilities. It will be useful to justify our main result and its consequences in Section 3.

Corollary 3.2. *Let* $\mathbf{A}$ *be a Boolean algebra. Then, every lower probability on* $\mathbf{A}$ *is a belief function if and only if* $\mathbf{A}$ *has two atoms.*

Proof. ($\Leftarrow$) Assume α_1, α_2 be the unique atoms of $\mathbf{A}$ and let $\underline{P}$ be a lower probability on $\mathbf{A}$. Discarding trivial cases, let us focus on the non-trivial events of $\mathbf{A}$: α_1 and α_2. Then, $\underline{P}(\alpha_1 \vee \alpha_2) = \underline{P}(\top) = 1$ and $\underline{P}(\alpha_1) + \underline{P}(\alpha_2) - \underline{P}(\alpha_1 \wedge \alpha_2) = \underline{P}(\alpha_1) + \underline{P}(\alpha_2) - \underline{P}(\bot) = \underline{P}(\alpha_1) + \underline{P}(\alpha_2)$. Moreover, α_1 and α_2 are disjoint, so $\underline{P}(\alpha_1) + \underline{P}(\alpha_2) \leqslant 1$, whence (13) is satisfied.

($\Rightarrow$) Assume that $\mathbf{A}$ has more than two atoms. Then the claim follows from Example 3.12 below where we show that in every algebra with three atoms one can define a lower probability that is not a belief function and obviously the same argument applies to any structure whose atoms are more than three. $\square$

The axiomatic relation between the uncertainty measures of interest in this work is depicted in Figure 3.1.

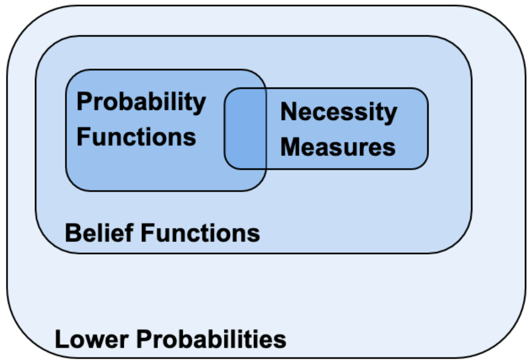


Figure 2: Graphical Representation of the axiomatic comparison between probability, necessity measures, belief functions and lower probabilities. Our main result shows that this is not matched in coherence.

3.2 A geometric framework for coherence

The framework for our results is rooted, in addition to de Finetti's seminal contributions, in the logico-geometric perspective put forward by Paris in [44]. In retrospect, this paper played an important role in the coming of age of the Dutch Book method, as the title says. In this Subsection, we lay down the geometric tools, and relevant notation/terminology, which will be used in the rest of the paper. We cast in this framework the well-known extension results yielded by geometrically coherent assignments, collected in Theorem 3.4.

Let $\Psi = \{a_1, \ldots, a_n\}$ be a finite set of events (i.e., elements of a finite Boolean algebra $\mathbf{A}$). Let us denote by $\mathbb{V} = \{v_1, \ldots, v_t\}$ the finite set of all possible homomorphisms of $\mathbf{A}$ to the boolean chain on the two-element set $\{0, 1\}$. For every $j = 1, \ldots, t$, call $\mathbf{e}_j$ the binary vector

$$\mathbf{e}_j = (v_j(a_1), \ldots, v_j(a_n)) \in \{0, 1\}^n. \tag{14}$$

Given this basic construction, and using an approach similar to Paris's we can characterize in geometric terms the extendability problem for books on Ψ to finitely additive probability measures, normalized necessity measures and belief functions. The additional notions we need are the Euclidean closed convex hull $\overline{\mathrm{co}}(X)$ of a subset $X \subseteq \mathbb{R}^t$ (which reduces to $\mathrm{co}(X)$ in case X is finite) and the less common *tropical* convex hull $\mathrm{co}_{\wedge,+}(X)$ of X (see [13]).

Definition 3.3. Let $\mathbf{x}_1, \ldots, \mathbf{x}_t \in [0, 1]^n$. The *tropical hull* of the $\mathbf{x}_j$'s is the subset $\mathrm{co}_{\wedge,+}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ of all points $\mathbf{y}$ of $[0, 1]^n$ for which there exist parameters $\lambda_1, \ldots, \lambda_t \in [0, 1]$ such that $\bigwedge_{j=1}^{t} \lambda_j = 0$ and

$$\mathbf{y} = \bigwedge_{j=1}^{t} \lambda_j + \mathbf{x}_j.$$

The symbol $\wedge$ stands for the minimum and $+$ for the ordinary addition in the tropical semiring $(\mathbb{R}, \wedge, +)$. Given $\lambda \in [0, 1]$ and $\mathbf{x} \in [0, 1]^n$, $\lambda + \mathbf{x} = (\lambda + x_1, \ldots, \lambda + x_n)$ and the $\bigwedge$ operator is defined component-wise.

For $\mathbf{e}_1, \ldots, \mathbf{e}_t$ being defined as above from the formulas a_i's in Ψ, let us consider the following sets:

1. $\mathscr{P}_\Psi = \mathrm{co}(\mathbf{e}_1, \ldots, \mathbf{e}_t)$;
2. $\mathscr{N}_\Psi = \mathrm{co}_{\wedge,+}(\mathbf{e}_1, \ldots, \mathbf{e}_t)$;

3. $\mathscr{B}_\Psi = \overline{\text{co}}(\mathscr{N}_\Psi)$, where, in this case, being $\mathscr{N}_\Psi$ usually uncountable, $\overline{\text{co}}$ denotes the topological closure of the Euclidean convex hull co.

Theorem 3.4 (Geometric extension results [12, 22, 25])**.** *Let $\Psi = \{a_1, \ldots, a_n\}$ be a finite set of events. Then a book $\beta : \Psi \to [0, 1]$ extends to a*

$$\left\{ \begin{array}{c} \textit{Probability measure} \\ \textit{Necessity measure} \\ \textit{Belief function} \end{array} \right\} \textit{ if and only if } (\beta(a_1), \ldots, \beta(a_n)) \in \left\{ \begin{array}{c} \mathscr{P}_\Psi \\ \mathscr{N}_\Psi \\ \mathscr{B}_\Psi \end{array} \right\}.$$

In general, $\mathscr{P}_\Psi$ and $\mathscr{N}_\Psi$ are both strictly included in $\mathscr{B}_\Psi$ (i.e., $\mathscr{P}_\Psi \subset \mathscr{B}_\Psi$ and $\mathscr{N}_\Psi \subset \mathscr{B}_\Psi$) and this is expected because belief functions are strictly more general than both probabilities and normalized necessity measures. In the present work, we investigate, via coherence, whether it is possible to distinguish uncertainty theories when we consider the more general setting of lower probabilities. In particular, we study if coherence is sufficiently robust to distinguish lower probabilities from belief functions.

Let us notice that, for every subset of events Ψ as above, the sets $\mathscr{P}_\Psi$, $\mathscr{N}_\Psi$ and $\mathscr{B}_\Psi$ are polyhedra of $[0, 1]^n$. More precisely, $\mathscr{P}_\Psi$ and $\mathscr{B}_\Psi$ are polytopes, i.e. convex polyhedra in the usual Euclidean sense, while $\mathscr{N}_\Psi$ is convex in the tropical sense specified in Definition 3.3 above, but it is not convex in the standard Euclidean model of tropical geometry. Thus, in that standard model, $\mathscr{N}_\Psi$ is represented by a polyhedron.

As for lower probabilities, the situation is similar but not fully understood. Denote by $\mathscr{L}_\Psi$ the set of all books on $\Psi \subseteq A$ that extend to a lower probability on the Boolean algebra $\mathbf{A}$. Although $\mathscr{L}_\Psi$ is known to be a polytope [46], a characterization of its extremal points is not fully understood. In [10] (see in particular §9 of the same paper) the authors make this point particularly clear. However, for the sake of the present paper, and in particular for the results of this Section and of Section 4, such a full description is not needed.

The next proposition makes clear a first relation between $\mathscr{L}_\Psi$ and $\mathscr{P}_\Psi$.

Proposition 3.5. *Let $\mathbf{A}$ be a finite Boolean algebra and $\Psi = \{a_1, \ldots, a_n\} \subseteq A$. A book β on Ψ belongs to $\mathscr{L}_\Psi$ if and only if there are $\beta_1, \ldots, \beta_n \in \mathscr{P}_\Psi$ such that, for all $a_i \in \Psi$, $\beta(a_i) = \min\{\beta_j(a_i) \mid j = 1, \ldots, n\}$.*

Proof. The right-to-left direction is trivial. Let us hence assume that β extends to a lower probability $\underline{P}$. Let $\mathcal{M}(\underline{P}) = \{P \mid P(a) \geqslant \underline{P}(a), \forall a \in A\}$ as in Section 2.2 and then, for all $a_i \in \Psi$,

$$\underline{P}(a_i) = \min\{P(a_i) \mid P \in \mathcal{M}(\underline{P})\}.$$

For all $P \in \mathscr{M}(\underline{P})$, call β_P the (necessarily coherent) book on Ψ obtained from P by restriction. Then, obviously,

$$\beta(a_i) = \min\{\beta_P(a_i) \mid P \in \mathscr{M}(\underline{P})\}.$$

Finally, since Ψ is finite, for every $a_i \in \Psi$ fix a book $\beta_{P(i)}$ among the β_P's such that

$$\beta_{P(i)}(a_i) = \beta(a_i) = \min\{\beta_P(a_i) \mid P \in \mathcal{M}(\underline{P})\}.$$

For every i, $\beta_{P(i)}$ exists. Then the claim follows since $\beta(a_i) = \min\{\beta_P(a_i) \mid P = P(i)\}$. In other words $\beta = \min\{\beta_{P(1)}, \ldots, \beta_{P(n)}\}$. □

The relation between belief functions and lower probabilities described in Section 3.1 is faithfully captured within our geometric framework.

Proposition 3.6. *Let $\mathbf{A}$ be a finite Boolean algebra and $\Psi = \{a_1, \ldots, a_n\} \subseteq A$, then*

$$\mathscr{B}_\Psi \subseteq \mathscr{L}_\Psi.$$

Proof. To prove our claim, it is sufficient to show that all the vertices of $\mathscr{B}_\Psi$ belong to $\mathscr{L}_\Psi$. The vertices of $\mathscr{B}_\Psi$ are either $\mathbf{e}_1, \ldots, \mathbf{e}_t$ or vectors recovered by taking the component-wise minimum between any subset of $\mathbf{e}_1, \ldots, \mathbf{e}_t$. Therefore, by Proposition 3.5, all the vertices of $\mathscr{B}_\Psi$ are in $\mathscr{L}_\Psi$, and the claim holds. □

Finally, let us notice that in case $\Psi = A$, i.e., the set of events we are considering coincides with the domain of the finite Boolean algebra we dealing with, then $\mathscr{P}_A$, $\mathscr{N}_A$, $\mathscr{B}_A$ and $\mathscr{L}_A$ are all polyhedra that respectively correspond to the sets of probabilities, necessity measure, belief functions and lower probabilities on $\mathbf{A}$.

3.3 A coherence-wise comparison

We now compare the various definitions of coherence arising from the Dutch Book method recalled in Section 2. More precisely we will carry out a comparison of the corresponding extendability theorems. A first result in this direction, where books are considered only on the whole algebra **4** of two atoms, is Corollary 3.2. This establishes that on **4** every lower probability is a belief function.

Note that this does not apply in general to other uncertainty measures. It is indeed easy to see that on **4**, probabilities, necessity measures and belief functions can be distinguished. In other words, and adopting the notation introduced in Subsection 3.2, $\mathscr{P}_4 \cap \mathscr{N}_4 \neq \mathscr{P}_4$, $\mathscr{P}_4 \cap \mathscr{N}_4 \neq \mathscr{N}_4$ (meaning that on **4** there are probabilities that are not necessity measures and vice-versa) and $\mathscr{P}_4, \mathscr{N}_4 \subset \mathscr{B}_4$ (i.e., there are belief functions on **4** that neither are probabilities, nor necessity measures). By a cardinality argument, the same relations hold among $\mathscr{P}_A$, $\mathscr{N}_A$ and $\mathscr{B}_A$ for all Boolean algebra $\mathbf{A}$ with more than 4 elements. For later use, let us hence state the following result that complements Corollary 3.2.

Proposition 3.7. *For every Boolean algebra* $\mathbf{A}$ *with cardinality* $|A| > 4$, $\mathscr{P}_A \cap \mathscr{N}_A \neq \mathscr{P}_A$, $\mathscr{P}_A \cap \mathscr{N}_A \neq \mathscr{N}_A$ *and* $\mathscr{P}_A, \mathscr{N}_A \subset \mathscr{B}_A \subset \mathscr{L}_A$.

If, instead of full measures on an algebra, we consider books on sets of events, a first trivial, yet suggestive, observation is the following: if $\mathbf{A}$ is any Boolean algebra and $\Psi = \{a\}$, then $\mathscr{P}_\Psi = \mathscr{N}_\Psi = \mathscr{B}_\Psi = \mathscr{L}_\Psi$. The same result applies trivially to sets of events of the form $\Psi = \{\bot, a, \top\}$ or $\Psi = \{\bot, a\}$ since all uncertainty measures considered so far are *normalized*, whence they all assign 0 to $\bot$ and 1 to $\top$. For this reason it will be important to define when a set of events is *adequate* for the analysis we propose. Those sets will be formally defined below.

In light of the above easy remark, we now aim at investigating the robustness of extension theorems for uncertainty measures and, by doing so, at understanding up to which extent the above Proposition 3.7 generalizes to adequate sets of events.

The next example has inspired the result described by our main result, namely Theorem 3.10.

Example 3.8. Let $\mathbf{A}$ be the Boolean algebra of 8 elements and 3 atoms $\{\alpha_1, \alpha_2, \alpha_3\}$ and consider the non-trivial set of events $\Psi = \{a_1, a_2, a_3\} \subset A$ where $a_1 = \alpha_1 \vee \alpha_2$, $a_2 = \alpha_2 \vee \alpha_3$ and $a_3 = \alpha_1 \vee \alpha_3$. The algebra $\mathbf{A}$ has 3 homomorphisms to $\{0,1\}$. Computing the points $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ as in (14), we get

$$\mathbf{e}_1 = (1,0,1);\ \mathbf{e}_2 = (1,1,0);\ \mathbf{e}_3 = (0,1,1).$$

The polyhedra $\mathscr{P}_\Psi$, $\mathscr{N}_\Psi$, $\mathscr{B}_\Psi$ and $\mathscr{L}_\Psi$ are hence as in Figure 3.

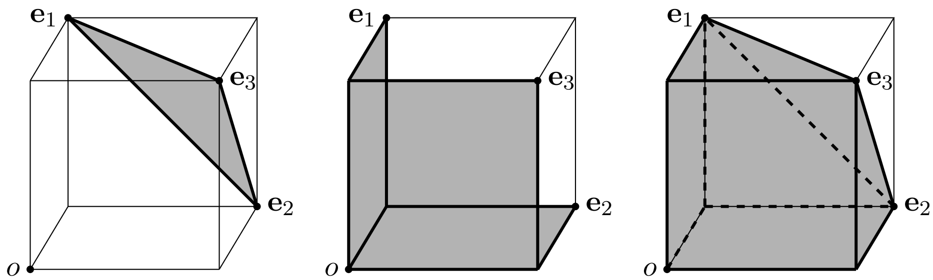


Figure 3: (left-most) The polytope $\mathscr{P}_\Psi$; (center) The tropical polytope $\mathscr{N}_\Psi$; (right-most) The polytopes $\mathscr{B}_\Psi$ and $\mathscr{L}_\Psi$.

Notice that, although Ψ does not coincide with the whole algebra $\mathbf{A}$, it allows to distinguish those books that are either extendible to a probability or a normalized necessity, from those extendible to belief functions or lower probabilities. Indeed both $\mathscr{P}_\Psi$ and $\mathscr{N}_\Psi$ are strict subsets of $\mathscr{B}_\Psi$ and $\mathscr{L}_\Psi$. Interestingly, in this specific example, $\mathscr{B}_\Psi$ and $\mathscr{L}_\Psi$ coincide.

We now define the notion of "adequate" set of events Ψ which allows us to discard those cases that we already know do not allow us to distinguish $\mathscr{B}_\Psi$ from $\mathscr{L}_\Psi$.

Definition 3.9. Let $\mathbf{A}$ be a Boolean algebra. A non-empty subset Ψ of A is *adequate* if Ψ is a strict subset of $A \backslash \{\bot, \top\}$ and the subalgebra $\mathbf{A}_\Psi$ of $\mathbf{A}$ generated by Ψ has at least three atoms.

Our main result shows that partial assignments on events exist for which it is impossible to tell whether they are coherent in the sense of lower probability theory but fail coherence according to belief functions. In logical terms, this suggests that there are non-negligible limits to the expressive power of coherence. In other words, we can show that for every algebra $\mathbf{A}$ with at least three atoms there exists $\Psi \subset A$ s.t. the convex hull characterising the assignments β on Ψ extendible to probability measures over $\mathbf{A}$ is not included in the convex hull characterising the assignments extendible to necessity measures over $\mathbf{A}$ and vice versa. I.e., there are $\beta_1, \beta_2 \colon \Psi \to [0,1]$ s.t. $\beta_1 \in \mathscr{P}_\Psi$ but $\beta_1 \notin \mathscr{N}_\Psi$, and $\beta_2 \in \mathscr{N}_\Psi$ but $\beta_2 \notin \mathscr{P}_\Psi$. In the same setting, we could expect a similar behaviour also for the more general uncertainty measures of belief functions and lower probabilities. However, as Theorem 3.10 shows, this is not the case. In fact, we will show that the convex hull characterising the assignments on Ψ that are bf-coherent ($\mathscr{B}_\Psi$) coincides with the convex hull characterising the assignments on Ψ that are l-coherent ($\mathscr{L}_\Psi$). However, if $\beta \in \mathscr{B}_\Psi = \mathscr{L}_\Psi$, then the corresponding extensions $\beta' \in \mathscr{B}_A$ and $\beta'' \in \mathscr{L}_A$ might not be the same.

Theorem 3.10 (When $\mathscr{B}_\Psi = \mathscr{L}_\Psi$)**.** *For every algebra $\mathbf{A}$ with at least three atoms there exists an adequate subset Ψ of A such that $\mathscr{P}_\Psi \cap \mathscr{N}_\Psi \neq \mathscr{P}_\Psi$ and $\mathscr{P}_\Psi \cap \mathscr{N}_\Psi \neq \mathscr{N}_\Psi$, but $\mathscr{B}_\Psi = \mathscr{L}_\Psi$.*

Proof. Let us assume without loss of generality that $\alpha_1, \ldots, \alpha_n$ ($n \geqslant 3$) are the atoms of $\mathbf{A}$ and let us fix the subset Ψ of A made of the following elements: $a_1 = \alpha_1 \vee \alpha_2$, $a_2 = \alpha_1 \vee \alpha_3$ and $a_3 = \alpha_2 \vee \alpha_3$. Clearly Ψ is adequate in the sense of Definition 3.9.

First, let us show that $\mathscr{P}_\Psi \cap \mathscr{N}_\Psi \neq \mathscr{P}_\Psi$ and $\mathscr{P}_\Psi \cap \mathscr{N}_\Psi \neq \mathscr{N}_\Psi$.

By Proposition 3.6, $\mathscr{B}_\Psi \subseteq \mathscr{L}_\Psi$. Thus, let β be a book in $\mathscr{L}_\Psi$. We want to prove that $\beta \in \mathscr{B}_\Psi$. Let $\underline{P}$ be a lower probability on $\mathbf{A}$ such that, for all $i = 1, \ldots, 3$, $\underline{P}(a_i) = \beta(a_i)$. Let us also assume that $\underline{P}$ is not a probability, that is to say, that β does not belong to $\mathscr{P}_\Psi$, otherwise, the claim would be trivial.

Now we prove the following.

Fact 3.11. $\beta \in \mathscr{M} = \mathrm{co}(\min\{\mathbf{e}_1, \mathbf{e}_2\}, \min\{\mathbf{e}_2, \mathbf{e}_3\}, \min\{\mathbf{e}_1, \mathbf{e}_3\},$
$\min\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\})$.

Proof. (of Fact 3.11). Assume, by way of contradiction, that $\beta \notin \mathscr{M}$. Thus, $\beta \in [0,1]^3 \setminus \mathscr{M}$, that is to say, $\beta \in \mathrm{co}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \max\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\})$. In other words, there exist $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ (with $\lambda_4 > 0$) such that $\sum_i \lambda_i = 1$ and

$$\beta = \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 + \lambda_4 \max\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}.$$

The expression above equals $\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \lambda_3 \mathbf{e}_3 + \max\{\lambda_4 \mathbf{e}_1, \lambda_4 \mathbf{e}_2, \lambda_4 \mathbf{e}_3\}$ and since $a + \max\{b, c\} = \max\{a + b, a + c\}$, one has

$$\beta = \max\{\beta_1, \beta_2, \beta_3\}$$

where $\beta_1 = (\lambda_1 + \lambda_4)\mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \lambda_3 \mathbf{e}_3$, $\beta_2 = \lambda_1 \mathbf{e}_1 + (\lambda_2 + \lambda_4)\mathbf{e}_2 + \lambda_3 \mathbf{e}_3$, $\beta_3 = \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + (\lambda_3 + \lambda_4)\mathbf{e}_3$. Thus, $\beta_1, \beta_2, \beta_3 \in \mathscr{P}_\Psi$. Letting P_i, for $i = 1, 2, 3$ such that P_i extends β_i, we conclude that β extends to an upper probability. Therefore, by assumption β extends to a lower probability. In addition, β extends to an upper probability, thus β extends to a probability that is absurd by a previous hypothesis. □

Now, we go back to the proof of the main claim and we prove that $\min\{\mathbf{e}_1, \mathbf{e}_2\}$, $\min\{\mathbf{e}_2, \mathbf{e}_3\}$, $\min\{\mathbf{e}_1, \mathbf{e}_3\}$, $\min\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \in \mathscr{N}_\Psi$. The claim is indeed easy to show by direct computation. For instance, check that $\min\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = (0, 0, 0)$ is $(N(a_1), N(a_2), N(a_3))$ where N is the necessity measure computed as in (4) and given by the normalized possibility distribution $\pi : \alpha_t = 1$ for all $t = 1, \ldots, n$.

Therefore β is a convex combination of points belonging to $\mathscr{N}_\Psi$. Hence it extends to a belief function, concluding the proof of Theorem 3.10. □

Notice that the above result does not say that if β extends to a lower probability $\underline{P}$, then $\underline{P}$ is a belief function. All it shows is that if β on events a_1, a_2, a_3 extends to lower probability $\underline{P}$, then there exists a belief function Bel that agrees with $\underline{P}$ on the a_i's but need not agree elsewhere. The following example clarifies this.

Example 3.12. Take again the setup of Example 2.16, now with the set of events $\Psi = \{a_1, a_2, a_3\}$ defined as in the proof of Theorem 3.10: $a_1 = \alpha_1 \vee \alpha_2$, $a_2 = \alpha_2 \vee \alpha_3$, and $a_3 = \alpha_1 \vee \alpha_3$. Let us consider also the book $\beta : a_i \mapsto q$ for every $i = 1, 2, 3$.

Since $q \leqslant 1/2$, $q \leqslant 1 - q$ and hence $\underline{P}(a_1) = \underline{P}(a_2) = \underline{P}(a_3) = q$. Thus, the lower probability $\underline{P}$ defined as in (10) extends β.

Furthermore, $\underline{P}$ is not a belief function. Indeed, $\underline{P}(a_1) + \underline{P}(a_2) + \underline{P}(a_3) - \underline{P}(a_1 \wedge a_2) - \underline{P}(a_2 \wedge a_3) - \underline{P}(a_1 \wedge a_3) + \underline{P}(a_1 \wedge a_2 \wedge a_3)$. Now, $a_1 \wedge a_2 \wedge a_3 = \perp$, whence $\underline{P}(a_1 \wedge a_2 \wedge a_3) = 0$ and, by definition of a_i, $\underline{P}(a_1 \wedge a_2) = \underline{P}(a_2 \wedge a_3) = \underline{P}(a_1 \wedge a_3) = 0$. Therefore, since $q > 1/3$, the above expression reduces to $\underline{P}(a_1) + \underline{P}(a_2) + \underline{P}(a_3) = 3q > 1 = \underline{P}(a_1 \vee a_2 \vee a_3)$ showing that $\underline{P}$ does not satisfy (13).

However, the belief function Bel whose mass assignments is $m(\{\alpha_1\}) = m(\{\alpha_2\}) = m(\{\alpha_3\}) = q/2$, $m(\{\alpha_1, \ldots, \alpha_t\}) = 1 - \frac{3}{2}q$ and $m(X) = 0$ otherwise, extends the same book β to the Boolean algebra $\mathbf{A}$.

Let us consider the set of events $\Psi' = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ where $a_1 = \alpha_1$, $a_2 = \alpha_2$, $a_3 = \alpha_3$ and a_4, a_5, and a_6 are defined as in the proof of Theorem 3.10: $a_4 = \alpha_1 \vee \alpha_2, a_5 = \alpha_2 \vee \alpha_3$, and $a_6 = \alpha_1 \vee \alpha_3$. The vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ relative to Ψ' are defined as follows:

$$\mathbf{e}_1 = (1, 0, 0, 1, 0, 1);\ \mathbf{e}_2 = (0, 1, 0, 1, 1, 0);\ \mathbf{e}_3 = (0, 0, 1, 0, 1, 1).$$

Thus, the vertices of $\mathscr{B}_{\Psi'}$ are $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\min\{\mathbf{e}_1, \mathbf{e}_2\} = (0, 0, 0, 1, 0, 0)$, $\min\{\mathbf{e}_2, \mathbf{e}_3\} = (0, 0, 0, 0, 1, 0)$, $\min\{\mathbf{e}_1, \mathbf{e}_3\} = (0, 0, 0, 0, 0, 1)$ and $\min\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = (0, 0, 0, 0, 0, 0)$. We can then verify that the point $(0, 0, 0, q, q, q) \notin \mathscr{B}_{\Psi'}$ while it belongs to $\mathscr{L}_{\Psi'}$.

Remark 3.13. As we pointed out in Section 3.2, an exhaustive description of the polytope $\mathscr{L}_\Psi$ is not known yet. More precisely, although $\mathscr{L}_\Psi$ is generated by its extremal points by Krein-Milman Theorem [19, Theorem 1.2], those latter, that are extremal lower probabilities coherent on Ψ, are not known in general. However, if Ψ is a set of events for which $\mathscr{B}_\Psi = \mathscr{L}_\Psi$, the extremal lower probabilities that are coherent on Ψ coincides with the extremal belief functions that are coherent on the same events. In other words, under that hypothesis and from the results recalled in Section 3.2, extremal lower probabilities that are coherent on Ψ are coherent necessity measures on Ψ, i.e., $\mathrm{ext}(\mathscr{L}_\Psi) \subseteq \mathscr{N}_\Psi$.

4 A metalogical representation

In this final section, we will put forward a framework to represent in logical terms the notions of coherence and, more in general, the uncertainty theories we considered so far. More in details we will see how the geometric approaches developed in the previous sections allow to bridge coherence and uncertainty theories on one side, and propositional logic and deductive reasoning on the other. A major role, in this sense, will be played by propositional Riesz logic [16] that we will briefly recall in the next Subsection 4.1, while in Subsection 4.2 we will present the connection between such formalism and coherence.

4.1 Riesz consequence relation $\models_{\mathsf{R}}$

A *Riesz space* is a vector space further endowed with a lattice order $\leqslant$ that is compatible with the vector spaces operations, i.e., a *vector lattice*. *Riesz* propositional logic R that we will briefly present in this section, has been firstly introduced in [16] as the extension of Łukasiewicz logic [41] by a uncountable family of unary connectives ∇_r (for $r \in [0,1]$) axiomatized in such a way that, from the algebraic viewpoint, they are necessarily interpreted as the scalar product in a(n interval of a) vector lattice.

Giving an exhaustive description for the logic R is out of the scope of the present paper, and we invite the interested reader to consult the literature on this subject (cf. [16]). However, what is of key importance for our logical analysis of coherence, is to recall in what R is able to speak and reason about polyhedral geometry (see [15] and [14, §3.3]). On this latter aspect, we will focus the rest of this subsection. In order to improve readability, we will assume the reader to be familiar with the basic of Łukasiewicz logic and MV-algebras. For a more exhaustive introduction about this subject we invite the reader to consult [5].

Consider the real unit interval $[0,1]$ and the algebraic language $(\oplus, \neg, 0)$ of type $(2,1,0)$ and where, for all $x, y \in [0,1]$, $x \oplus y = \min\{1, x+y\}$ and $\neg x = 1 - x$. Furthermore, for every $r \in [0,1]$ let $p_r : [0,1] \to [0,1]$ be the unary operation $x \mapsto p_r(x) = rx$. The algebra $[0,1]_{RMV} = ([0,1], \oplus, \neg, \{p_r\}_{r\in[0,1]}, 0)$ is the prototypical example of a *Riesz MV-algebra* and it is called the *standard* Reisz MV-algebra.

Definition 4.1. A system $\mathbf{A}$ on a non-empty domain A and in the algebraic language $(\oplus, \neg, \{p_r\}_{r\in[0,1]}, 0)$ is a *Riesz MV-algebra* if $\mathbf{A}$ belongs to $\mathbb{V}([0,1]_{RMV}) = \mathbb{RMV}$, the algebraic variety generated by $[0,1]_{RMV}$.

For the logical translation of coherence that we will present in this final section, a major role will be played by special RMV-algebras: the finitely generated free algebras. These structures are, up to isomorphism, the Lindenbaum-Tarski algebra of Riesz logic that we will briefly present below, and they can be further characterized as follows.

Example 4.2 ([15, Theorem 1.3])**.** Let k be finite. A function $f : [0,1]^k \to [0,1]$ is said to be a *Riesz function* if f is continuous, piecewise linear and each piece has coefficients from $\mathbb{R}$. Then, for every finite $k \in \mathbb{N}$, let $R(k)$ the set of all Reisz functions on $[0,1]^k$. The k-generated free Riesz MV-algebra is, up to isomorphism, the algebra $\mathcal{R}(k) = (R(k), \oplus, \neg, \{p_r\}_{r\in[0,1]}, 0)$ where operations are defined by the point-wise application of those of the standard algebra $[0,1]_{RMV}$.

Riesz logic $\vdash_{\mathsf{R}}$ is the algebraizable logic, in the sense of Blok and Pigozzi [4], whose equivalent algebraic semantics is $\mathbb{RMV}$. The algebraizability of R w.r.t. $\mathbb{RMV}$ gives us that *formulas* of the former can be equivalently regarded as terms of the latter. In what follows we will hence say that $\hat\varphi$ (possibly $\hat\varphi(x_1, \dots, x_k)$ if we want to point out the propositional variables occurring in it)[1] is a *formula of* R, meaning that $\hat\varphi$ is a term in the algebraic language of Riesz-algebra (on variables $x_1, \dots, x_k$). As we

[1] In this last section we will use the notation $\hat\varphi$, $\hat\psi$ etc., to distinguish formulas of R from the elements of an arbitrary Boolean algebras that, as in the previous sections, we will denote by lowercase Greek letters φ, ψ, etc.

recalled above, the Lindenbaum-Tarski algebra of R on k propositional variable $\mathcal{L}(k)$ coincides, up to isomorphism, with the k-generated free Riesz MV-algebra $\mathcal{R}(k)$ of Example 4.2.

The isomorphism between $\mathcal{L}(k)$ and $\mathcal{R}(k)$, tells us that if $\hat{\varphi}(x_1, \ldots, x_k)$ is a formula of Riesz logic, its equivalence class (modulo equi-provability in R) in $\mathcal{L}(k)$ can be regarded as a Riesz function $f_{\hat{\varphi}} : [0,1]^k \to [0,1]$ and, vice versa, for every Riesz function $f \in R(k)$, there exists a (possibly not unique) formula $\hat{\varphi}_f(x_1, \ldots, x_k)$ whose equivalence class in $\mathcal{L}(k)$ can be isomorphically associated to f.

Definition 4.3. For every finite k and every formula $\hat{\varphi}(x_1, \ldots, x_k)$, we will write $\mathcal{M}od(\hat{\varphi}) = \{(a_1, \ldots, a_k) \in [0,1]^k \mid f_{\hat{\varphi}}(a_1, \ldots, a_k) = 1\}$. This set will be called the set of *models* of $\hat{\varphi}$ or, equivalently, the *one-set* of $f_{\hat{\varphi}}$.

The semantic consequence relation of R will be denoted by $\models_{\mathsf{R}}$. Hence, if $\hat{\varphi}(x_1, \ldots, x_k)$ and $\hat{\psi}(x_1, \ldots, x_k)$ are formulas, $\hat{\varphi} \models_{\mathsf{R}} \hat{\psi}$ means that $\mathcal{M}od(\hat{\varphi}) \subseteq \mathcal{M}od(\hat{\psi})$. In other words, $f_{\hat{\psi}}(a_1, \ldots, a_k) = 1$ in $[0,1]_{RMV}$ for all those $(a_1, \ldots, a_k) \in [0,1]^k$ such that $f_{\hat{\varphi}}(a_1, \ldots, a_k) = 1$ in $[0,1]_{RMV}$; $\hat{\varphi} \dashv\models_{\mathsf{R}} \hat{\psi}$ stands for $\hat{\varphi} \models_{\mathsf{R}} \hat{\psi}$ and $\hat{\psi} \models_{\mathsf{R}} \hat{\varphi}$ and it hence indicates that $\mathcal{M}od(\hat{\varphi}) = \mathcal{M}od(\hat{\psi})$.

The next result summarises known and useful facts about Riesz logic and its close connection with polyhedral geometry.

Proposition 4.4 ([15, Theorem 3.3]). *(1) For every formula $\hat{\varphi}(x_1, \ldots, x_k)$, $\mathcal{M}od(\hat{\varphi})$ is a polyhedron of $[0,1]^k$;*

(2) For every polyhedron $\mathscr{P}$ of $[0,1]^k$ there exists a formula $\hat{\rho}_{\mathscr{P}}(x_1, \ldots, x_k)$ such that $\mathscr{P} = \mathcal{M}od(\hat{\rho}_{\mathscr{P}})$.

Recall that the formula $\hat{\rho}_{\mathscr{P}}$ of the above Proposition 4.4 is not necessarily unique. However, in what follows, we will speak about *the* formula $\hat{\rho}_{\mathscr{P}}$ such that $\mathscr{P} = \mathcal{M}od(\hat{\rho}_{\mathscr{P}})$ intending that we have chosen one among all those formulas that satisfy the above claim.

The last useful definition that is necessary to recall is that of $\mathbb{R}$-function between polyhedra.

Definition 4.5. Let $\mathscr{P} \subseteq [0,1]^k$ and $\mathscr{Q} \subseteq [0,1]^n$ be polyhedra. We say that a map $\eta : \mathscr{P} \to \mathscr{Q}$ is a $\mathbb{R}$-*map* if there exists Riesz functions $f_1, \ldots, f_n \in \mathcal{R}(k)$ such that, for all $x \in \mathscr{P}$,

$$\eta(x) = (f_1(x), \ldots, f_n(x)) \in \mathscr{Q}.$$

For what we will show in the next subsection is important to notice that the projection maps of any polyhedron $\mathscr{P} \subseteq [0,1]^k$ to a lower dimension $[0,1]^n$ (for $n \leqslant k$) are elementary examples of $\mathbb{R}$-maps.

4.2 Coherence through $\models_{\mathsf{R}}$

The results presented in Subsection 3.2, and Theorem 3.4 in particular, show that for every finite set of events $\Psi = \{\psi_1, \ldots, \psi_k\}$, the sets $\mathscr{P}_\Psi$, $\mathscr{N}_\Psi$, $\mathscr{B}_\Psi$ and $\mathscr{L}_\Psi$ of books that extends to probability functions, necessity measures, belief functions and lower probabilities respectively, are polyhedra of $[0,1]^k$. The following result provides preliminary results on the logical description of coherence via Riesz logic R. Its proof is an immediate consequence of Proposition 4.4 and the definition of $\models_{\mathsf{R}}$. In what follows we will adopt the notation used in Proposition 4.4 (2).

Corollary 4.6. *For every finite set of events $\Psi = \{a_1, \ldots, a_k\}$ and for every book $\beta : \Psi \to [0,1]$, the following conditions hold:*

1. *$(\beta(a_1), \ldots, \beta(a_k)) \in \mathscr{C}_\Psi$ iff $\hat{\rho}_{\{\beta\}} \models_{\mathsf{R}} \hat{\rho}_{\mathscr{C}_\Psi}$ and for every $\mathscr{C} \in \{\mathscr{P}, \mathscr{N}, \mathscr{B}, \mathscr{L}\}$;*
2. *$\hat{\rho}_{\mathscr{P}_\Psi} \models_{\mathsf{R}} \hat{\rho}_{\mathscr{B}_\Psi}$; $\hat{\rho}_{\mathscr{N}_\Psi} \models_{\mathsf{R}} \hat{\rho}_{\mathscr{B}_\Psi}$; $\hat{\rho}_{\mathscr{B}_\Psi} \models_{\mathsf{R}} \hat{\rho}_{\mathscr{L}_\Psi}$.*

Notice that, while claim (1) in the corollary above follows from the fact that points are special examples of polyhedra, claim (2) is a consequence of what we observed in Section 3. Precisely that coherence for lower probabilities is a more general notion than coherence for belief functions and this latter, in turn, is more general than coherence for necessity measures and probability functions.

Also the main result of Section 3 on the existence of adequate sets of events for which one cannot distinguish between books that are extendible to belief functions or lower probabilities, Theorem 3.10, can be rephrased in the following terms.

Proposition 4.7. *For every Boolean algebra* $\mathbf{A}$ *with at least three atoms there exists an adequate subset* Ψ *of* A *such that:*

1. $\hat{\rho}_{\mathscr{P}_\Psi} \not\models \hat{\rho}_{\mathscr{P}_\Psi \cap \mathscr{N}_\psi}$ *and* $\hat{\rho}_{\mathscr{N}_\Psi} \not\models \hat{\rho}_{\mathscr{P}_\Psi \cap \mathscr{N}_\psi}$*;*
2. $\hat{\rho}_{\mathscr{B}_\Psi} \dashv\models_{\mathsf{R}} \hat{\rho}_{\mathscr{L}_\Psi}$.

So far, we have seen how the geometric description of coherence can be straightforwardly described by metalogical properties of Reisz logic. Now, we end this section by showing a less trivial interpretation that allows to regard polyhedra of coherent books as projections of full measures defined on finite algebras.

In what follows, let us use the symbol $\mathscr{C}$ to be any among $\{\mathscr{P}, \mathscr{N}, \mathscr{B}, \mathscr{L}\}$ and, once we have fixed a $\mathscr{C} \in \{\mathscr{P}, \mathscr{N}, \mathscr{B}, \mathscr{L}\}$ we will say that a map d from a Boolean algebra $\mathbf{A}$ to $[0,1]$ is a $\mathscr{C}$-measure as a general nomenclature for: d is a finitely additive probability measure (in case $\mathscr{C} = \mathscr{P}$), d is a necessity measure (if $\mathscr{C} = \mathscr{N}$), d is a belief function (for $\mathscr{C} = \mathscr{B}$), and d is a lower probability (if $\mathscr{C} = \mathscr{L}$).

From the next result, whose proof is a direct consequence of the extension Theorem 3.4 and the definition of $\mathscr{L}_\Psi$, we will start denoting by π_Ψ the projection map of $[0,1]^A$ to $[0,1]^\Psi$.

Proposition 4.8. *For every finite Boolean algebra* $\mathbf{A}$ *and every subset* Ψ *of* A*,*

$$\mathscr{C}_\Psi = \pi_\Psi(\mathscr{C}_A)$$

Therefore, each polytope $\mathscr{C}_\Psi$ is the projection of $\mathscr{C}_A$ on the axes indexed by $a_1, \ldots, a_k$.

From what we observed at the end of Subsection 4.1 projections of polyhedra to lower dimensional cubes are $\mathbb{R}$-maps. Therefore, by Definition 4.5, for every $\Psi \subseteq A$ and for every $\mathscr{C}_\Psi \subseteq [0,1]^k$, there exist Riesz functions $f_1, \ldots, f_k : [0,1]^A \to [0,1]$ such that the projection map $\pi_\Psi : \mathscr{C}_A \subseteq [0,1]^{2^n} \to \mathscr{C}_\Psi \subseteq [0,1]^k$ acts as follows: for every $(a_1, \ldots, a_{2^n}) \in \mathscr{C}_A$,

$$\pi_\Psi(a_1, \ldots, a_{2^n}) = (f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_{2^n})).$$

Thus, $(f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_{2^n})) \in \mathscr{C}_\Psi$. More details on what such projections look like from the logico-algebraic perspective will be given in the proof of the next result that provides a more precise logical reading of the previous Proposition 4.8.

Proposition 4.9. *Let* $\mathbf{A}$ *be the finite Boolean algebra of cardinality* 2^n *and let* $\Psi = \{a_1, \ldots, a_k\} \subseteq A$*. Then, there are Riesz functions* $f_1, \ldots, f_k : [0,1]^A \to [0,1]$ *such that* $(b_1, \ldots, b_k) \in \mathscr{C}_\Psi$ *if and only if there exists* $(a_1, \ldots, a_{2^n}) \in \mathscr{C}_A$ *such that*

$$(b_1, \ldots, b_k) = (f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_{2^n})).$$

Proof. Let us fix, without loss of generality, the enumeration $a_1, \ldots, a_{2^n}$ for the elements of A in such a way that the first k element, in the natural order, determine the event set Ψ (clearly $k \leqslant 2^n$ since $\Psi \subseteq A$). For every $i = 1, \ldots, k$, let f_i be the map from $[0,1]^A$ to $[0,1]$ defined as follows: for all $g : A \to [0,1]$, $f_i(g) = g(a_i)$. In other words, once identified the functions from A to $[0,1]$ as stings of length 2^n of elements of $[0,1]$ as $(a_1, \ldots, a_{2^n})$, $f_i(a_1, \ldots, a_{2^n}) = a_i$. Thus, f_i is a Riesz homomorphism of $[0,1]^A$ to $[0,1]$.

Now, if $(b_1, \ldots, b_k) \in \mathscr{C}_\Psi$, by definition there exists a $\mathscr{C}$-measure $d : A \to [0,1]$ that extends it. That is to say, $(b_1, \ldots, b_k) \in \mathscr{C}_\Psi$ if and only if the vector $(d(a_1), d(a_2), \ldots, d(a_{2^n})) \in \mathscr{C}_A$ and for all $i = 1, \ldots, k$, $b_i = d(a_i)$. Therefore, for all $i = 1, \ldots, k$, by definition of f_i, one has

$$b_i = d(a_i) = f_i(d(a_1), \ldots, d(a_k)).$$

Conversely, if $(b_1, \ldots, b_{2^n}) \in \mathscr{C}_A$, then there exists a $\mathscr{C}$-measure $d : \mathbf{A} \to [0,1]$ such that, for all $j = 1, \ldots, 2^n$, $b_j = d(a_j)$. By definition of f_i, and by our previous assumption on the a_i'a for $i = 1, \ldots, k$,

$$(f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_{2^n})) = (b_1, \ldots, b_k) = (d(a_1), \ldots, d(a_k)).$$

The latter is coherent being the restriction to Ψ of a $\mathscr{C}$-measure d. Thus we finally get $(f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_{2^n})) \in \mathscr{C}_\Psi$ and the claim is settled. □

Our next result provides a uniform logical representation of the extension theorems for probabilities, necessity measures, belief functions and lower probabilities. More precisely, the next contains two results: the first one describes, in logical terms, the claim of the above Propositions 4.8 and 4.9 by interpreting geometric projections as logical substitutions[2]; the second one bridges the claim of Theorem 3.4 that characterizes extension results via the geometry of coherence and that of Corollary 4.6 that links the geometry of coherence and deductions in Riesz logic. Moreover, it is worth noticing that the characterization we present in the next theorem extends [21, Theorem 5.5] in the scope of uncertainty theories that allow for such logical description.

Theorem 4.10. *Let* $\mathbf{A}$ *be a finite Boolean algebra and* $\Psi \subseteq A$*. Then there exists a substitution* σ_Ψ *such that*

(1) $\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi}) \dashv\models_{\mathsf{R}} \hat{\rho}_{\mathscr{C}_A}$.

Furthermore, for all $\beta : \Psi \to [0,1]$ *the following conditions are equivalent:*

(2) β *extends to a* $\mathscr{C}$*-measure;*

(3) $\hat{\rho}_{\{\beta\}} \models_{\mathsf{R}} \hat{\rho}_{\mathscr{C}_\Psi}$*;*

Proof. Let us notice that the equivalence between (2) and (3) immediately follows from Theorem 3.4 and Corollary 4.6 (1).

As to prove (1), by Proposition 4.9, there are Riesz functions $f_1, \ldots, f_k$ (where $k = |\Psi|$) such that,

$$\mathscr{C}_\Psi = \{(f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_n)) \mid (a_1, \ldots, a_{2^n}) \in \mathscr{C}_A\},$$

where $2^n = |A|$.

Therefore, if $x_1, \ldots, x_k$ are the variable occurring in $\hat{\rho}_{\mathscr{C}_\Psi}$, define σ_Ψ to be the substitution that maps each variable x_i to the term $\hat{\varphi}_{f_i}(y_1, \ldots, y_{2^n})$, the Riesz formula that corresponds to f_i. More precisely, σ_Ψ maps $\hat{\rho}_{\mathscr{C}_\Psi}(x_1, \ldots, x_k)$ to

$$\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi})(y_1, \ldots, y_{2^n}) = \hat{\rho}_{\mathscr{C}_\Psi}(\hat{\varphi}_{f_1}(y_1, \ldots, y_{2^n}), \ldots, \hat{\varphi}_{f_k}(y_1, \ldots, y_{2^n}))$$

Then the claim follows. Indeed, by Proposition 4.9 and Proposition 4.4(2), one has:

$$\begin{array}{rll} (a_1, \ldots, a_{2^n}) \in \mathcal{M}od(\hat{\rho}_{\mathscr{C}_A}) & \text{iff} & (a_1, \ldots, a_{2^n}) \in \mathscr{C}_A \\ & \text{iff} & (f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_n)) \in \mathscr{C}_\Psi \\ & \text{iff} & (f_1(a_1, \ldots, a_{2^n}), \ldots, f_k(a_1, \ldots, a_n)) \in \mathcal{M}od(\hat{\rho}_{\mathscr{C}_\Psi}) \\ & \text{iff} & (a_1, \ldots, a_{2^n}) \in \mathcal{M}od(\hat{\rho}_{\mathscr{C}_\Psi}(\hat{\varphi}_{f_1}, \ldots, \hat{\varphi}_{f_k}) \\ & \text{iff} & (a_1, \ldots, a_{2^n}) \in \mathcal{M}od(\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi}). \end{array}$$

[2]Recall that in the framework of algebraic logic, given a language $\mathcal{L}$, a *logical substitution* (or simply a *substitution*) is a map σ that assigns to every propositional variable x of $\mathcal{L}$, a formula $\hat{\varphi}$ of the same language $\mathcal{L}$. Equivalently, substitutions are endomorphisms of the Lindenbaum-Tarski algebra of $\mathcal{L}$, once restricted to the variables of $\mathcal{L}$.

Therefore, $(a_1, \ldots, a_{2^n}) \in \mathcal{M}od(\hat{\rho}_{\mathscr{C}_A})$ iff $(a_1, \ldots, a_{2^n}) \in \mathcal{M}od(\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi}))$. In other words, $\mathcal{M}od(\hat{\rho}_{\mathscr{C}_A}) = \mathcal{M}od(\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi}))$ and hence, by definition of $\dashv\models_{\mathsf{R}}$, it follows that $\sigma_\Psi(\hat{\rho}_{\mathscr{C}_\Psi}) \dashv\models_{\mathsf{R}} \hat{\rho}_{\mathscr{C}_A}$. □

5 Conclusions and Future Work

We have put forward a logico-geometric framework which allows us to investigate the notion of coherence at a considerable level of generality and detail. Within this framework we have put forward i) a comparison between the geometric representations of coherence for finitely-additive measures with their non-additive counterparts, and ii) a comparison between non-additive measures themselves. Our key finding is that non-additive measures which *can* be distinguished axiomatically may not be distinguishable coherence-wise. The outcomes of our geometrical comparisons are also recovered within a logical representation of coherence provided by means of suitably defined Riesz consequence relations.

Two questions which we think are worthy of further investigation arise in the framework put forward in this paper.

In Chapter 3 of [12], Bruno de Finetti establishes the equivalence between the Dutch Book method discussed above and the method of (proper) *scoring rules*. This latter defines coherence as the minimisation of expected loss under a well-defined penalty function known as the Brier Score. As de Finetti points out, the equivalence between the two seemingly different criteria has a geometric explanation, as they both boil down to pinning the convex hull of the n-dimensional linear space $\mathcal{S}$ which arises from assigning values in $[0, 1]$ to a set of n events Ψ. A natural question then is to extend the coherence-wise comparison of non-additive uncertainty measures carried out in this paper to the scoring rules method. Our preliminary investigations on this show that the answer is not straightforward. In [8], we put forward a scoring rule for belief functions and show that a characterisation of coherence as minimisation of expected loss under that rule in the context of tropical geometry. To carry out the analogue of the present comparison a suitable scoring rule for lower probabilities must be put forward. Some earlier results of Seidenfeld Schervish and Kadane [48] point out several difficulties in doing this. In addition, with the exception of what we have noted in Remark 3.13, a full description of the extreme points of $\mathscr{L}_\Psi$ is not presently known. Hence more research in this direction is needed to put forward an analysis similar to the present one in terms of scoring rules.

The metalogical framework of Section 4 allows us to pursue an analogy which has been used casually in this paper. The idea is to take the properties of the betting game as fixing the "intended semantics" of an uncertainty representation. To unfold the analogy, recall that in classical logic, model inclusion is the intended semantics of the classical consequence relation "$\models$". Similarly, provability is the intended semantics of its intuitionistic counterpart, and so on. The role of intended semantics is chiefly to provide guidelines for the *material adequacy* of the formal definition, to borrow Tarkis's own expression. So in the case of de Finetti's Dutch Book argument, this justifies "incurring sure loss" as a blatantly undesirable outcome for a bookmaker. Hence coherence is defined in such a way to avoid that. This suggests the following question: are there basic properties that can be identified as crucial in moving from one uncertainty measure to another, pretty much the way in which the family of normal modal logics arise from adding suitable conditions to the distribution axiom K? In other words, can we use the logical framework introduced above to put forward a modular approach enabling us to match, whenever possible, distinct notions of coherence with distinct properties of the underlying decision problem? We know from our main result that this is not going to match fully, and generally, the axiomatic-wise distinct measures of uncertainty, but nothing in principle prevents the remaining cases to be described metalogically along the lines of Theorem 4.10.

Acknowledgements We are very grateful to the organisers and the audiences of ISIPTA 2021, ECSQARU 2021 and PROGIC 2021, where our preliminary results have been presented and discussed. Authors are grateful to the referees for their work and to Serafina Lapenta for her useful comments and suggestions concerning Riesz logic. Corsi and Hosni acknowledge funding by the Department of Philosophy "Piero Martinetti" of the University of Milan under the Project "Departments of Excellence

2018-2022" awarded by the Ministry of Education, University and Research (MIUR). Flaminio acknowledges partial support by the MOSAIC project (EU H2020- MSCA-RISE- 2020 Project 101007627) and also by the Spanish project PID2019-111544GB-C21 funded by MCIN/AEI/10.13039/501100011033. Hosni also acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, grant LA 4093/3-1).

References

[1] B. Anger and J. Lembcke. Infinitely subadditive capacities as upper envelopes of measures. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 68(3):403–414, 1985.

[2] T. Augustin, F. P. Coolen, G. De Cooman, and M. C. Troffaes. *Introduction to imprecise probabilities.* John Wiley & Sons, 2014.

[3] N. Bingham. Finite Additivity Versus Countable Additivity. *Electronic Journal for History of Probability and Statistics*, 6, 2010.

[4] W. J. Blok and D. Pigozzi. *Algebraizable logics*, volume 77. American Mathematical Soc., 1989.

[5] R. L. Cignoli, I. M. d'Ottaviano, and D. Mundici. *Algebraic foundations of many-valued reasoning*, volume 7. Springer Science & Business Media, 2013.

[6] R. Cori and D. Lascar. *Mathematical Logic: Part 1: Propositional Calculus, Boolean Algebras, Predicate Calculus, Completeness Theorems.* OUP Oxford, 2000.

[7] R. Cori, D. Lascar, and D. H. Pelletier. *Mathematical Logic.* Oxford University Press, 1994.

[8] E. A. Corsi, T. Flaminio, and H. Hosni. Scoring rules for belief functions and imprecise probabilities: A comparison. In J. Vejnarová and N. Wilson, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 16th European Conference, ECSQARU 2021, Prague, Czech Republic, September 21-24, 2021, Proceedings*, volume 12897 of *Lecture Notes in Computer Science*, pages 301–313. Springer, 2021.

[9] E. A. Corsi, T. Flaminio, and H. Hosni. When belief functions and lower probabilities are indistinguishable. In *ISIPTA 2021 - Proceedings of Machine Learning Research (147)*, pages 83–89, 2021.

[10] J. De Bock and G. De Cooman. Extreme lower previsions. *Journal of Mathematical Analysis and Applications*, 421(2):1042–1080, 2015.

[11] B. De Finetti. Sul significato soggettivo della probabilita. *Fundamenta mathematicae*, 17(1):298–329, 1931.

[12] B. De Finetti. *Theory of probability*, volume 1. John Wiley & Sons, 1974.

[13] M. Develin and B. Sturmfels. Tropical convexity. *arXiv preprint math/0308254*, 2003.

[14] A. Di Nola, S. Lapenta, and I. Leuştean. An analysis of the logic of riesz spaces with strong unit. *Annals of Pure and Applied Logic*, 169(3):216–234, 2018.

[15] A. Di Nola, G. Lenzi, and G. Vitale. Riesz–mcnaughton functions and riesz mv-algebras of nonlinear functions. *Fuzzy Sets and Systems*, 311(1–14), 2017.

[16] A. Di Nola and I. Leuştean. Łukasiewicz logic and Riesz spaces. *Soft Computing*, 18(12):2349–2363, 2014.

[17] D. Dubois and H. Prade. Possibility theory: Approach to computerized processing of uncertainty, plennm n. 4. *Plenum Press, New York*, 1988.

[18] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics*, pages 643–669, 1961.

[19] G. Ewald. *Combinatorial convexity and algebraic geometry*, volume 168. Springer Science & Business Media, 1996.

[20] R. Fagin, J. Y. Halpern, and N. Megiddo. A logic for reasoning about probabilities. *Information and computation*, 87(1-2):78–128, 1990.

[21] T. Flaminio. Three characterizations of strict coherence on infinite-valued events. *The Review of Symbolic Logic*, 13(3):593–610, 2020.

[22] T. Flaminio and L. Godo. A note on the convex structure of uncertainty measures on MV-algebras. In *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, pages 73–81. Springer, 2013.

[23] T. Flaminio, L. Godo, and H. Hosni. On the logical structure of de Finetti's notion of event. *Journal of Applied Logic*, 12(3):279–301, 2014.

[24] T. Flaminio, L. Godo, and H. Hosni. Coherence in the aggregate: a betting method for belief functions on many-valued events. *International Journal of Approximate Reasoning*, 58:71–86, 2015.

[25] T. Flaminio, L. Godo, and E. Marchioni. Geometrical aspects of possibility measures on finite domain MV-clans. *Soft Computing*, 16(11):1863–1873, 2012.

[26] G. Gerla. Inferences in probability logic. *Artificial Intelligence*, 70(1-2):33–52, 1994.

[27] S. Givant and P. Halmos. *Introduction to Boolean algebras.* Springer Verlag, 2009.

[28] R. Haenni, J.-W. Romeijn, G. Wheeler, and J. Williamson. *Probabilistic logics and probabilistic networks*, volume 350. Springer Science & Business Media, 2010.

[29] P. Hájek. Basic fuzzy logic and bl-algebras. *Soft computing*, 2(3):124–128, 1998.

[30] P. R. Halmos. *Measure theory.* Springer Science & Business Media, 1950.

[31] P. R. Halmos. *Lectures on Boolean algebras.* Springer Verlag, 1974.

[32] P. R. Halmos and S. Givant. *Introduction to Boolean algebras.* Springer, 2009.

[33] J. Y. Halpern. *Reasoning about uncertainty.* MIT press, 2003.

[34] J.-Y. Jaffray. Coherent bets under partially resolving uncertainty and belief functions. *Theory and decision*, 26(2):99–105, 1989.

[35] T. Jech. Measures on boolean algebras. *Fundamenta Mathematicae*, 239(2):177–183, 2017.

[36] H. E. Kyburg. Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, 31(3):271–293, 1987.

[37] H. E. Kyburg and C. M. Teng. *Uncertain inference.* Cambridge University Press, 2001.

[38] I. Levi. *The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability, and Chance.* MIT Press, 1980.

[39] P. Marquis, O. Papini, and H. Prade. *A Guided Tour of Artificial Intelligence Research: Volume I: Knowledge Representation, Reasoning and Learning.* Springer Nature, 2020.

[40] E. Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.

[41] D. Mundici. *Advanced Łukasiewicz calculus and MV-algebras*, volume 35. Springer Science & Business Media, 2011.

[42] N. J. Nilsson. Probabilistic logic. *Artificial intelligence*, 28(1):71–87, 1986.

[43] J. Paris. *The uncertain reasoner's companion: A mathematical perspective.* Cambridge University Press, 1994.

[44] J. Paris. A note on the Dutch Book method. In *ISIPTA*, volume 1, pages 301–306, 2001.

[45] J. Paris and A. Vencovska. *Pure Inductive Logic*. Cambridge University Press, 2015.

[46] E. Quaeghebeur. Characterizing the set of coherent lower previsions with a finite number of constraints or vertices. In Grünwald, Peter and Spirtes, Peter, editor, *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 466–473, 2010.

[47] H. P. Sankappanavar and S. Burris. *A course in Universal Algebra*. Springer-Velag, New York, 1981.

[48] T. Seidenfeld, M. J. Schervish, and J. B. Kadane. Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning*, 53(8):1248–1261, 2012.

[49] G. Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.

[50] G. Shafer. Constructive probability. *Synthese*, pages 1–60, 1981.

[51] C. Smith. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society. Series B (Methodological)*, 23(1):1–37, 1961.

[52] M. Troffaes and G. de Cooman. *Lower Previsions*. Wiley, 2014.

[53] P. Wakker. *Prospect theory: for risk and ambiguity*. Cambridge University Press, 2010.

[54] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Wiley, 1991.

[55] P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman & Hall, 1991.

[56] P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125–148, 2000.

[57] P. M. Williams. On a new theory of epistemic probability. *The British Journal for the Philosophy of Science*, 29(4):375–387, 1978.

[58] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

[59] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*, 1(1):3–28, 1978.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Milan, 30.09.2021
On behalf of all the co-authors of the paper

Esther Anna Corsi