Rationality and desirability

A foundational study

Doctoral Dissertation submitted to the
Faculty of Informatics of the Università della Svizzera Italiana
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

presented by
Arianna Casanova Flores

under the supervision of
Luca Maria Gambardella and Marco Zaffalon

May 2023

Dissertation Committee

Fabio Crestani	Università della Svizzera italiana, Switzerland
Ernst C. Wit	Università della Svizzera italiana, Switzerland
Matthias C.M. Troffaes	Durham University, UK
Nic Wilson	University College Cork, Ireland

Research Advisor

Luca Maria Gambardella

Co-Advisor

Marco Zaffalon

PhD Program Director

Walter Binder and Silvia Santini

I certify that except where due acknowledgement has been given, the work presented in this thesis is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; and the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program.

Arianna Casanova Flores
Lugano, 2 May 2023

To my beloved family

Abstract

The past century witnessed significant achievements in classical scientific fields like decision theory, social choice theory, and Bayesian probability, all rooted in the common principle of consistency as a foundation for rationality, i.e, in order to be rational, a calculus needs to deliver *coherent* inferences, or a subject needs to maintain *coherent* preferences among options.

Bayesian probability is the arena where this idea has been faced more vividly, starting with the work of de Finetti, 1937. De Finetti indeed clearly saw that rationality, in the mentioned interpretation of coherence, and probability, were just the same thing. Another arena where coherence has been identified with a science field is logic: consistency indeed is the subject matter of logic.

A turning point in the interplay of logic, probability, and coherence, has been Peter Williams' definition of *desirability* [Williams, 1975], an extension of de Finetti's theory of probability made to deal with imprecision, as originated by incompleteness or other reasons.

Desirability's main tools, namely the *coherent sets of gambles*, represent acceptable bets for rational agents adhering to specific axioms. These sets have connections to logic and probability, resembling *closed theories* and encompassing various generalizations of probability, such as lower and upper probabilities, convex sets of distributions, and more [Walley, 1991, 2000; Quaeghebeur, 2014]. Furthermore, recent works by Zaffalon and Miranda [2017, 2021] have extended desirability to decision making and demonstrated its correspondence with traditional decision-making preferences.

In our thesis, we delve deeper into the potential of desirability, pursuing three primary research lines. First, we propose desirability as a framework for aggregating opinions, thus unifying different forms of opinions based on coherence. This approach provides novel insights into traditional results and simplifies comparisons among different formalisms [Arrow, 1951; Feldman and Serrano, 2006].

Second, we examine the relationship between desirability and *information algebras* [Kohlas, 2003], revealing that desirability can be seen as an instance of

these algebraic structures for inference. This insight presents a novel algebraic analysis of desirability and enriches it with the inference tools provided by these structures.

Third, we explore relaxations of desirability's foundational axioms to allow a more realistic interpretation of gambles. Specifically, we analyse different sets of axioms and we re-interpret them as binary (usually nonlinear) classification problems. Then, borrowing ideas from machine learning, we define *feature mappings* allowing us to reformulate the above nonlinear classification problems as linear ones in higher-dimensional spaces.

Acknowledgements

First of all, I would express my deepest gratitude to my advisor, prof. Luca Maria Gambardella, for giving me the opportunity to join a leading and vibrant research institution such as Dalle Molle Institute for Artificial Intelligence USI-SUPSI (IDSIA), and for having guided and supported me throughout my PhD journey. I am also immensely grateful to my co-advisor, prof. Marco Zaffalon, whose unwavering dedication, expertise, and mentorship have been critical to my success. I am deeply appreciative of the countless hours he has devoted to advising me on my research, meticulously reviewing my work, and offering invaluable feedback.

I am also extremely grateful for the collaboration and support of the other professors who have contributed to my research. Prof. Enrique Miranda, Prof. Juerg Kohlas, and Prof. Alessio Benavoli have been instrumental in shaping my research, providing invaluable insights and feedback, and sharing their expertise in their respective fields. Their contributions have greatly enriched my work, and I am honored to have had the opportunity to learn from them.

I am deeply indebted to the other members of my thesis committee, Prof. Fabio Crestani, Prof. Ernst C. Wit, Prof. Matthias C.M. Troffaes and Dr. Nic Wilson, for their invaluable inputs and feedback during the proposal of my dissertation. Their diverse perspectives and expertise have enriched my work and contributed to its overall quality. I am grateful for their time and commitment to providing me with constructive and insightful feedback.

I am also grateful to all members of IDSIA to promote such a supportive and intellectually stimulating environment that provided an invaluable contribution to my personal and academic growth.

Last but not least, I would like to express my heartfelt gratitude to my family and friends for their tenacious support throughout my academic journey. Their love, encouragement, and unwavering belief in me have sustained me through the highs and lows of my PhD journey.

Contents

Chapter 1

Introduction

The core of the present thesis is *desirability* or the *theory of coherent sets of desirable gambles*. It was first introduced by Williams [1975] as a generalization of de Finetti's theory of probability to deal with imprecise information. Then, it was addressed in much more detail by Walley [1991].

In this chapter, we focus specifically on introducing its fundamentals and detailing its plural role as a powerful uncertainty formalism that can be regarded also as a general theory of decision making and eventually interpreted as a logic.

For further details, we refer to Walley [1991], Quaeghebeur [2014], Troffaes and de Cooman [2014], Zaffalon and Miranda [2017, 2021].

1.1 Desirability

De Finetti established a theory of probability entirely based on an idea of consistency, called *coherence*. In his work, he showed that familiar axioms of probability can be justified by imposing only a consistency principle:[1] avoiding being exposed to a sure loss on prices fixed on a set of odds about an uncertain experiment (see de Finetti, 1937). This approach has been further developed giving rise to *desirability* or the *theory of coherent sets of desirable gambles* (see Williams, 1975; Walley, 1991; Quaeghebeur, 2014).

To present its basic tools, we need to first introduce *gambles*. Consider a non-empty set Ω describing the possible mutually exclusive outcomes of some experiment. We call it *possibility space*.

[1] It's worth noting that de Finetti rejects the adoption of Kolmogorov's axiom of *countable additivity*, instead opting for using finitely additive models only. He firmly believed that countable additivity lacked theoretical justification and was irrelevant for practical purposes.

Definition 1 (Gamble). *Given a possibility space Ω, a* gamble *on Ω is a bounded real-valued function $f : \Omega \to \mathbb{R}$.*

A gamble is interpreted as an uncertain reward in a linear utility scale.[2] Finding a gamble f *desirable* or *acceptable* is regarded as a commitment to receive $f(\omega)$ whatever ω occurs. The next example will further clarify the concept.

Example 1. *Let us consider Alice, a detective who is investigating a murder case. In this context, let us consider Ω denoting the possibility space of the answers to the following question: 'How tall is the murderer?' (e.g., $[1.5, 2]m$).*

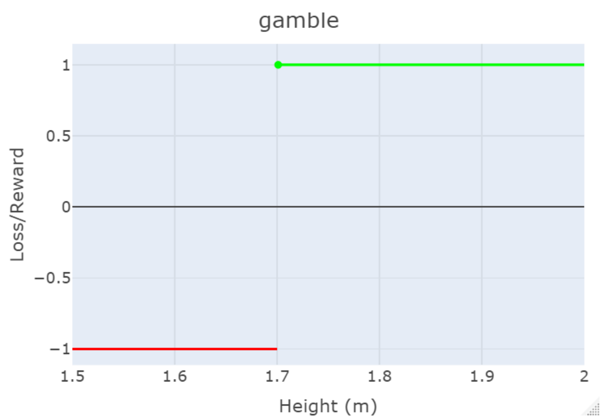

Figure 1.1. A gamble.

If Alice is disposed to accept the gamble f represented in Fig. 1.1, she would commit herself to receive, for instance, 1 utile (generic unity of measure of utility) if the murderer is taller than $1.7m$ ($\omega \geq 1.7m$), and lose 1 utile if the murderer is lower than $1.7m$.

To simplify the notation, we can also partition the set Ω into the subsets $Tall := \{\omega \in \Omega : \omega \geq 1.7m\}$ and $Low := Tall^C = \Omega \setminus Tall$. Using this notation, $f = \mathbb{I}_{Tall} - \mathbb{I}_{Low}$, where we denote with $\mathbb{I}_B$ the indicator function of a set $B \subseteq \Omega$.

Gambles exist that a rational agent should always be disposed to accept or, respectively, reject. Non-negative, non-vanishing gambles should always be desirable since they can increase the wealth of an agent without the risk of decreasing it. Similarly, non-positive non-vanishing gambles should never be acceptable

[2] Linearity means that if we are willing to take gambles f_1 and f_2, then we are also willing to take gambles $\lambda_1 f_1 + \lambda_2 f_2$ for any real constants $\lambda_1, \lambda_2 \geq 0$ not both equal to 0 (check D3, D4 in Definition 2). The standard interpretation of gambles that aligns with this assumption presumes that their rewards are represented by lottery tickets that can be either won or lost [Walley, 1991, Section 2.2].

for opposite reasons. As a consequence of the linearity of the utility scale moreover, a rational agent disposed to accept the transactions represented by gambles f and g, should also be disposed to accept the transactions $\lambda f + \mu g$ for every $\lambda, \mu \geq 0$ not both equal to 0.

These rationality criteria can be summarised in the following notion of *coherent set of desirable gambles* or, for short, *coherent set of gambles* (see Quaeghebeur, 2014, Section 1.2.3). In what follows, we use the notation $\mathscr{L}(\Omega)$ to denote the set of all gambles defined on Ω, $\mathscr{L}^+(\Omega)$ to denote the set of non-negative, non-vanishing gambles $\mathscr{L}^+(\Omega) := \{f \in \mathscr{L}(\Omega) : f \geq 0, f \neq 0\}$ and $\mathscr{L}^-(\Omega)$ to denote the set of non-positive gambles $\mathscr{L}^-(\Omega) := \{f \in \mathscr{L}(\Omega) : f \leq 0\}$. To simplify the notation, whenever possible we omit the possibility space Ω. Thus we write $\mathscr{L}, \mathscr{L}^+, \mathscr{L}^-$ in place of $\mathscr{L}(\Omega)$, $\mathscr{L}^+(\Omega), \mathscr{L}^-(\Omega)$ respectively.

Definition 2 (**Coherence for sets of gambles**)**.** *We say that a subset $\mathscr{D}$ of $\mathscr{L}$ is a* coherent *set of desirable gambles or, for short, a* coherent *set of gambles, if and only if $\mathscr{D}$ satisfies the following properties:*

D1. $\mathscr{L}^+ \subseteq \mathscr{D}$ *[Accepting Partial Gains],*

D2. $0 \notin \mathscr{D}$ *[Avoiding Status Quo],*

D3. $f, g \in \mathscr{D} \Rightarrow f + g \in \mathscr{D}$ *[Additivity],*

D4. $f \in \mathscr{D}, \lambda > 0 \Rightarrow \lambda f \in \mathscr{D}$ *[Positive Homogeneity].*

Thus, geometrically, coherent sets of gambles correspond to convex cones.

It's worth emphasizing that, in line with this definition, coherent sets of gambles *avoid partial loss*: $\mathscr{D} \cap \mathscr{L}^- = \emptyset$. Whether or not to include the 0 gamble is, however, a matter of convention. There is, in fact, an alternative version of this definition that incorporates it [Quaeghebeur, 2014, Section 1.4.2].

For further reference, we also introduce the set:

$$\mathbb{D}(\Omega) := \{\mathscr{D} \subseteq \mathscr{L}(\Omega) : \mathscr{D} \text{ is coherent}\}, \tag{1.1}$$

which we abbreviate to $\mathbb{D}$ if no ambiguity is possible.

A coherent set of gambles represents the set of gambles desirable for an agent respecting a set of rationality axioms. Nevertheless, the agent can alternatively evaluate the desirability of a smaller set of gambles and then employ the *natural extension* operation to infer which additional gambles must also be deemed desirable based on axioms D1, D3, D4. This approach is guaranteed to avoid any other commitments.

Definition 3 (Natural extension for sets of gambles). *Given a set* $\mathscr{K} \subseteq \mathscr{L}$*, we call* $\mathscr{E}(\mathscr{K}) := \mathrm{posi}(\mathscr{K} \cup \mathscr{L}^+)$ *its* natural extension*, where*

$$\mathrm{posi}(\mathscr{K}') := \left\{ \sum_{j=1}^{r} \lambda_j f_j : f_j \in \mathscr{K}', \lambda_j > 0, r \geq 1 \right\},$$

for every set $\mathscr{K}' \subseteq \mathscr{L}$.[3]

If $\mathscr{E}(\mathscr{K})$ is coherent, it is the smallest coherent set containing $\mathscr{K}$:

$$\mathscr{E}(\mathscr{K}) = \bigcap \{\mathscr{D}' \in \mathbb{D} : \mathscr{K} \subseteq \mathscr{D}'\}.$$

It is coherent, in particular, if and only if $0 \notin \mathscr{E}(\mathscr{K})$, i.e., if it is not possible to derive the 0 gamble from $\mathscr{K}$ simply applying rules D1,D3,D4 in the most conservative way.

Example 2. *Returning to the previous example, suppose now Alice declares (only) to be disposed to accept* f*, i.e.,* $\mathscr{K} = \{f\}$*. The minimal set of acceptable gambles deducible for her applying rules D1–D4 is*

$$\mathscr{E}(\mathscr{K}) := \mathrm{posi}(\{f\} \cup \mathscr{L}^+) = \mathrm{posi}(\{\mathbb{I}_{Tall} - \mathbb{I}_{Low}\} \cup \mathscr{L}^+) =$$
$$= \{\lambda(\mathbb{I}_{Tall} - \mathbb{I}_{Low}) + f : \lambda > 0, f \geq 0\} \cup \{f : f \geq 0, f \neq 0\},$$

which is in particular a coherent set of gambles.

Coherent sets of gambles that are not proper subsets of other coherent sets are called *maximal*.

Proposition 1 (Maximal coherent set of gambles). *A coherent set of gambles* $\mathscr{D}$ *is* maximal *if and only if*

$$(\forall f \in \mathscr{L} \setminus \{0\})\; f \notin \mathscr{D} \Rightarrow -f \in \mathscr{D}. \tag{1.2}$$

We shall employ the notation M for maximal coherent sets of gambles to differentiate them from the general case of coherent ones. We use as well $\mathbb{M}(\Omega)$ or $\mathbb{M}$ to denote the set of all the maximal coherent sets of gambles. Maximal coherent sets of gambles satisfy moreover the following properties that will be used in Chapter 3:

1. any coherent set of gambles $\mathscr{D}$ is a subset of a maximal one, which we call *maximal superset* of $\mathscr{D}$;

[3] A variation of the posi operator, also called *conic hull* operator, is the *convex hull* operator, which will be frequently used later on:

$$(\forall \mathscr{K} \subseteq \mathscr{L})\; \mathrm{ch}(\mathscr{K}) := \left\{ \sum_{j=1}^{r} \lambda_j f_j : f_j \in \mathscr{K}, \lambda_j \geq 0, \sum_{j=1}^{r} \lambda_j = 1, r \geq 1 \right\}.$$

2. any coherent set of gambles is the intersection of all its maximal supersets.

For the proofs of these properties, see [de Cooman and Quaeghebeur, 2012, Theorem 3, Corollary 4].

Axioms D1–D4 listed above actually derive from a primitive definition of coherence more generally applicable in those situations where the attention of the agent is restricted to a subset $Q \subseteq \mathscr{L}$ of gambles [Walley, 1991, Section. 3.7.8; Miranda and Zaffalon, 2010].

Definition 4 (Coherence relative to a set of gambles). *Consider a set $Q \subseteq \mathscr{L}$. We say that a set of gambles R is* coherent relative to Q *if $0 \notin \mathscr{E}(R)$ and $Q \cap \mathscr{E}(R) \subseteq R$. In case Q coincides with $\mathscr{L}$, we simply say that R is coherent.*

This definition is clearly consistent with Definition 2. In what follows, however, we shall mostly focus on coherent sets of gambles. We shall point out when this is not the case.

In what follows, we may also wish to focus on one particular *aspect* of an experiment. The typical situation is when the experiment's possibility space has a product structure and we want to concentrate only on one component, ignoring the others. In those cases, the interesting gambles are the ones depending only on the compelling component of the possibility space. This concept can be generalised considering the following definition.

Definition 5 (Measurable gambles). *Given a partition $\mathscr{P}$ of Ω, we say that a gamble f on Ω is* $\mathscr{P}$-measurable *if and only if it is actually a function on $\mathscr{P}$:*

$$(\forall B \in \mathscr{P})(\forall \omega, \omega' \in B)\, f(\omega) = f(\omega').$$

We shall denote by $\mathscr{L}_{\mathscr{P}}(\Omega)$ the subset of $\mathscr{L}(\Omega)$ given by the $\mathscr{P}$-measurable gambles. Note that there is a one-to-one correspondence between $\mathscr{L}_{\mathscr{P}}(\Omega)$ and $\mathscr{L}(\mathscr{P})$. Indeed, to every $f \in \mathscr{L}_{\mathscr{P}}(\Omega)$, we can associate the gamble $f^{\downarrow\mathscr{P}}(B) := f(\omega)$ for every $\omega \in B$ and every $B \in \mathscr{P}$. Vice versa, if we introduce the partition $\mathscr{T}$ composed by the singletons of Ω, i.e., $\mathscr{T} := \{\{\omega\} : \omega \in \Omega\}$, to every gamble $g \in \mathscr{L}(\mathscr{P})$ we can associate the gamble $g^{\uparrow\mathscr{T}}(\omega) := g(B)$ for every $\omega \in \Omega$ such that $\omega \in B$ for some $B \in \mathscr{P}$. Clearly we have $f = (f^{\downarrow\mathscr{P}})^{\uparrow\mathscr{T}}$ and $g = (g^{\uparrow\mathscr{T}})^{\downarrow\mathscr{P}}$.

Given a coherent set of gambles, we call its subset of $\mathscr{P}$-marginal gambles, its *$\mathscr{P}$-marginal set of gambles*.

Definition 6 (Marginal set of gambles). *Let $\mathscr{D} \subseteq \mathscr{L}(\Omega)$ be a coherent set of gambles and consider a partition $\mathscr{P}$ of Ω. The* $\mathscr{P}$-marginal *of $\mathscr{D}$ is the set $\mathscr{D}_{\mathscr{P}} := \mathscr{D} \cap \mathscr{L}_{\mathscr{P}}(\Omega)$.*

In other situations, we may be interested in obtaining an uncertainty model for those cases where the experiment's outcome belongs to a *conditioning event* $B \subseteq \Omega$. In these situations, we can focus on gambles that are *contingent* on B occurring, i.e., gambles such that if B does not occur, no payoff is received—*status quo* is maintained—.

Definition 7 (Conditional gambles). *Given a non-empty set $B \subseteq \Omega$, we say that a gamble f on Ω is* conditional on B *if and only if it is zero outside B: $f = \mathbb{I}_B f$.*

We shall denote by $\mathscr{L}(\Omega)|B$ the subset of $\mathscr{L}(\Omega)$ made of gambles that are conditional on $B \subseteq \Omega$. Note that, similar to before, there is a one-to-one correspondence between $\mathscr{L}(\Omega)|B$ and $\mathscr{L}(B)$.

Given a coherent sets of gambles $\mathscr{D}$, its intersection with B-conditional gambles is called its B-conditional set.

Definition 8 (Conditional set of gambles). *Let $\mathscr{D} \subseteq \mathscr{L}(\Omega)$ be a coherent set of gambles and consider a non-empty set $B \subseteq \Omega$. The B-*conditional *of $\mathscr{D}$ is the set $\mathscr{D}|B := \mathscr{D} \cap \mathscr{L}(\Omega)|B$.*

From the coherence of $\mathscr{D}$, it follows in particular that $\mathscr{D}_{\mathscr{P}}$ is coherent relative to $\mathscr{L}_{\mathscr{P}}$ for every partition $\mathscr{P}$ of Ω and $\mathscr{D}|B$ is coherent relative to $\mathscr{L}|B$ for every non-empty set $B \subseteq \Omega$. Analogously, we can derive the coherence of the correspondent sets in $\mathscr{L}(\mathscr{P})$ and $\mathscr{L}(B)$ respectively.

Moving to order theory, sets of gambles form a partially ordered set with respect to inclusion $(\mathrm{P}(\mathscr{L}), \subseteq)$.[4] The natural extension operator, in particular, is a *closure operator* on this poset $(\mathrm{P}(\mathscr{L}), \subseteq)$, i.e., a map $Cl : \mathrm{P}(\mathscr{L}) \to \mathrm{P}(\mathscr{L})$ satisfying the following properties [Davey and Priestley, 2002, Definition. 7.1]:

CL1. $(\forall \mathscr{K} \subseteq \mathscr{L})\ \mathscr{K} \subseteq Cl(\mathscr{K})$;

CL2. $(\forall \mathscr{K}, \mathscr{K}' \subseteq \mathscr{L})\ \mathscr{K} \subseteq \mathscr{K}' \Rightarrow Cl(\mathscr{K}) \subseteq Cl(\mathscr{K}')$;

CL3. $(\forall \mathscr{K} \subseteq \mathscr{L})\ Cl(Cl(\mathscr{K})) = Cl(\mathscr{K})$.

If we now focus only on the poset of coherent sets of gambles $(\mathbb{D}, \subseteq)$, we can notice that, again by standard order theory [Davey and Priestley, 2002], it induces a meet-semilattice where meet is intersection. It is possible to define also a join for a family of coherent sets of gambles $\{\mathscr{D}_j\}_{j \in J}$ where J is an index set, if and only if they have an upper bound among coherent sets:

$$\bigvee_{j \in J} \mathscr{D}_j := \bigcap \{\mathscr{D}' \in \mathbb{D} : \bigcup_{j \in J} \mathscr{D}_j \subseteq \mathscr{D}'\}.$$

[4] We indicate with $\mathrm{P}(\mathscr{L})$ the power-set of $\mathscr{L}$.

Notice in particular that if $\mathscr{E}(\bigcup_{j\in J}\mathscr{D}_j)$ is coherent, we have:

$$\bigvee_{j\in J}\mathscr{D}_j = \mathscr{E}(\bigcup_{j\in J}\mathscr{D}_j).$$

To obtain a complete lattice, we need to add $\mathscr{L}(\Omega)$ to $\mathbb{D}(\Omega)$. We denote the resulting set with $\Phi(\Omega) := \mathbb{D}(\Omega)\cup\{\mathscr{L}(\Omega)\}$. In what follows, we can simply refer to it with Φ when there is no possible ambiguity. In particular, $(\Phi,\subseteq)$ induces a complete lattice where meet is intersection and join is defined for any family of sets $\{\mathscr{D}_j\}_{j\in J}$ with $\mathscr{D}_j \in \Phi$ for every $j \in J$, as:

$$\bigvee_{j\in J}\mathscr{D}_j := \bigcap\{\mathscr{D}' \in \Phi : \bigcup_{j\in J}\mathscr{D}_j \subseteq \mathscr{D}'\}. \tag{1.3}$$

Starting from this definition of join, we can construct another closure operator on $(\mathrm{P}(\mathscr{L}),\subseteq)$ similar to the natural extension operator:

$$(\forall \mathscr{K} \subseteq \mathscr{L})\ \mathscr{C}(\mathscr{K}) := \bigcap\{\mathscr{D}' \in \Phi : \mathscr{K} \subseteq \mathscr{D}'\}. \tag{1.4}$$

Notice that, given $\mathscr{K} \subseteq \mathscr{L}$:

- if $0 \notin \mathscr{E}(\mathscr{K})$, $\mathscr{C}(\mathscr{K}) = \mathscr{E}(\mathscr{K})$;
- if $0 \in \mathscr{E}(\mathscr{K})$, $\mathscr{C}(\mathscr{K}) = \mathscr{L}$ and it is possible to have $\mathscr{E}(\mathscr{K}) \neq \mathscr{L} = \mathscr{C}(\mathscr{K})$.[5]

We refer to de Cooman [2005] for a similar order-theoretic view of desirability.

We conclude this section introducing two simplified variants of desirability that will be useful later on. Specifically, the concepts we introduce correspond to Definition 3.7.8. and Definition 3.7.3. of Walley, 1991.

Definition 9 (**Strict desirability**). *A set of gambles $\mathscr{K}$ is said to be a* coherent set of strictly desirable gambles *if and only if it is coherent and it satisfies*

$$(\forall f \in \mathscr{K} \setminus \mathscr{L}^{+})(\exists \delta > 0)\ f - \delta \in \mathscr{K}.^{6} \tag{1.5}$$

We use the notation $\mathscr{D}^{+}$ for coherent sets of strictly desirable gambles. Notice moreover that, with a certain abuse of notation, we use λ to denote the constant gamble defined as $f(\omega) = \lambda$ for every $\omega \in \Omega$.

Definition 10 (**Almost desirability**). *We say that a subset $\mathscr{K}$ of $\mathscr{L}$ is a* coherent set of almost desirable gambles *if and only if $\mathscr{K}$ satisfies the following properties:*

D1'. $f \in \mathscr{L}$ *and* $\inf f > 0 \Rightarrow f \in \mathscr{K}$ *[Accepting Sure Gains],*

[5] Consider for example $\mathscr{K} = \{0\}$.

[6] In this context, we present a slightly different definition of a coherent set of strictly desirable gambles, as compared to Definition 3.7.8 of Walley, 1991. Nevertheless, it's important to note that these two definitions are equivalent.

D2'. $f \in \mathscr{K} \;\Rightarrow\; \sup f \geq 0$ *[Avoiding Sure Loss],*

D3'. $f, g \in \mathscr{K} \Rightarrow f + g \in \mathscr{K}$ *[Additivity],*

D4'. $f \in \mathscr{K}, \lambda > 0 \Rightarrow \lambda f \in \mathscr{K}$ *[Positive Homogeneity],*

D5'. $f \in \mathscr{L}$ *and* $f + \delta \in \mathscr{K}$ *for all* $\delta > 0 \;\Rightarrow\; f \in \mathscr{K}$ *[Closure].*[7]

We use the notation $\overline{\mathscr{D}}$ for coherent sets of almost desirable gambles.

Axioms D1–D4 introduced in Definition 2 represent the consistency conditions for gambles that are considered *really* desirable by an agent. A coherent set of almost desirable gambles also includes all the gambles that are a limit (under the supremum norm) of desirable gambles—also called *almost desirable*—though some of them, as the null gamble, are not really desirable. It is therefore closed under the supremum-norm topology and it corresponds to the relative closure of a coherent set of desirable gambles [Quaeghebeur, 2014, Section 1.6.4; Walley, 1991, Appendix F.]. Coherent sets of strictly desirable gambles instead correspond to relative interiors of coherent sets of gambles plus the non-negative non-vanishing gambles and, excluding these latter gambles, are open under the supremum-norm topology [Quaeghebeur, 2014, Section 1.6.4; Walley, 1991, Appendix F.]. The simplified border structure of coherent sets of strictly and almost desirable gambles makes them equivalent to other uncertainty models such as the *coherent lower and upper previsions* that we analyze in the next subsection. Coherent sets of gambles, however, are more general and can deal more effectively with the problem of conditioning on sets of measure zero [Walley, 1991, Appendix F4].

Coherent sets of strictly desirable gambles are in particular coherent and form a subfamily of the coherent sets of gambles. In what follows, we indicate with $\mathbb{D}^+(\Omega)$ or $\mathbb{D}^+$, the set of all the coherent sets of strictly desirable gambles. Similarly to coherent sets, we can then add $\mathscr{L}(\Omega)$ to $\mathbb{D}^+(\Omega)$ and obtain $\Phi^+(\Omega) := \mathbb{D}^+(\Omega) \cup \{\mathscr{L}(\Omega)\}$, which we abbreviate to Φ^+ when no ambiguity is possible.

Coherent sets of almost desirable gambles instead are not coherent in the sense of Definition 2 (they contain the zero gamble). However, they are subjected to a weaker notion of coherence, since they cannot contain strictly negative gambles. In what follows, we indicate with $\overline{\mathbb{D}}(\Omega)$ or $\overline{\mathbb{D}}$ the set of all the coherent sets of almost desirable gambles.

[7] It is useful to note that, in this definition, $\inf f > 0$ can be equivalently relaxed to $\inf f \geq 0$ and $\lambda > 0$ to $\lambda \geq 0$. This alternative definition differs from the original one but it clearer conceives the idea of the structure of a coherent set of almost desirable gambles (see the discussion below the definition and the one opening Chapter 4).

For future references, we also point out that, given a set $\mathscr{K} \subseteq \mathscr{L}(\Omega)$, the closure of its natural extension in the supremum-norm topology $\overline{\mathscr{E}(\mathscr{K})}$ corresponds, when it does not contain strictly negative gambles, to the smallest coherent set of almost desirable gambles containing $\mathscr{K}$ [Walley, 1991, Section 3.7.4].

1.1.1 Desirability and probability

Coherent sets of desirable, strictly desirable, and almost desirable gambles have a tight connection with probability theory. In particular, from each of these sets, it is possible to construct a *coherent lower* and a *coherent upper prevision*, which correspond respectively to lower and upper expectation functionals.

Lower and upper previsions, including coherent lower and upper previsions as specific instances, fall under the umbrella term of *imprecise probabilities* alongside sets of gambles. This term is employed in a broad sense to encompass all mathematical models which measure chance or uncertainty without providing precise numerical probabilities. Notably, lower and upper previsions represent some of the most versatile mathematical models for handling uncertainty, even more general than probability measures, *possibility measures*, *belief functions*, etcetera [Walley, 2000].[8] Lower/Upper previsions also have a clear behavioural interpretation. More formally, a *lower prevision* $\underline{P}$ is a function with values in $\mathbb{R} \cup \{+\infty\}$ defined on some class of gambles $dom(\underline{P})$, called the *domain* of $\underline{P}$. Its values $\underline{P}(f)$ represent the supremum buying price an agent is willing to spend for $f \in dom(\underline{P})$. Analogously, it is possible to define an *upper prevision* $\overline{P} : dom(\overline{P}) \to \mathbb{R} \cup \{-\infty\}$ whose values represent the infimum selling price an agent is disposed to set for gambles in its domain.[9]

Definition 11 (Lower and upper prevision). *Let us consider a non-empty subset $\mathscr{K}$ of $\mathscr{L}$ representing the set of gambles that an agent finds desirable. The agent's* lower prevision *(operator) $\underline{P} : dom(\underline{P}) \to \mathbb{R} \cup \{+\infty\}$ is defined as*

$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathscr{K}\} \tag{1.6}$$

for every $f \in dom(\underline{P})$. The agent's upper prevision *(operator) $\overline{P} : dom(\overline{P}) \to$*

[8] They, however, are still less general than coherent sets of gambles, as explained later on in this subsection.

[9] In the literature, lower previsions are typically defined as functions with values in $\mathbb{R}$. However, in this thesis, to extend their applicability to the limit set $\mathscr{L}$ as required in Chapter 3, we broaden the definition of a lower prevision to include the value $+\infty$ within its range. It's important to highlight that a similar, more general definition of lower prevision has been previously proposed in Troffaes and de Cooman, 2014, Theorem 4.33.

$\mathbb{R} \cup \{-\infty\}$ is instead defined as

$$\overline{P}(f) := -\underline{P}(-f) \tag{1.7}$$

for every $f \in dom(\overline{P})$, where $dom(\underline{P}), dom(\overline{P}) \subseteq \mathscr{L}$.

In what follows we will concentrate only on lower previsions since, from them, we can completely derive the upper ones.

In the definition above we have not made explicit the dependence on $\mathscr{K}$. When it is important, we can indicate a lower prevision $\underline{P}$ as the outcome of a function σ applied to a set of gambles $\mathscr{K}$ and write $\underline{P} = \sigma(\mathscr{K})$.

If $\mathscr{K} \subseteq \mathscr{L}$ is a coherent set of desirable/strictly desirable/almost desirable gambles, its associated lower prevision is called *coherent*. Coherent lower previsions $\underline{P}$ are characterised by having $\mathscr{L}$ as their domain and satisfying the following properties.[10]

LP1. $(\forall f \in \mathscr{L})\ \underline{P}(f) \geq \inf_{\omega \in \Omega} f(\omega)$,

LP2. $(\forall f \in \mathscr{L},\ \lambda > 0)\ \underline{P}(\lambda f) = \lambda \underline{P}(f)$ [Positive Homogeneity],

LP3. $(\forall f, g \in \mathscr{L})\ \underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$ [Superlinearity].

In what follows, we denote the set of all the *coherent lower previsions* as $\underline{\mathbb{P}}(\Omega)$ or $\underline{\mathbb{P}}$.

The correspondence between coherent sets of strictly desirable gambles and coherent lower previsions is one-to-one. The same is true for the coherent sets of almost desirable gambles. Given a coherent lower prevision $\underline{P}$ indeed, the set:

$$\mathscr{D}^+ := \{f \in \mathscr{L} : \underline{P}(f) > 0\} \cup \mathscr{L}^+, \tag{1.8}$$

is a coherent set of strictly desirable gambles and moreover induces $\underline{P}$ through Eq. (1.6). Alternatively, the set:

$$\overline{\mathscr{D}} := \{f \in \mathscr{L} : \underline{P}(f) \geq 0\}, \tag{1.9}$$

is a coherent set of almost desirable gambles and induces again $\underline{P}$ through Eq. (1.6).

We can therefore define the maps τ^+ and $\overline{\tau}$ as the inverses of the map σ restricted respectively to coherent sets of strictly and almost desirable gambles:

$$(\forall \underline{P} \in \underline{\mathbb{P}}) \quad \tau^+(\underline{P}) := \{f \in \mathscr{L} : \underline{P}(f) > 0\} \cup \mathscr{L}^+, \quad \overline{\tau}(\underline{P}) := \{f \in \mathscr{L} : \underline{P}(f) \geq 0\}.$$

These considerations cannot be extended to arbitrary coherent sets of desirable gambles. Several different coherent sets of desirable gambles $\mathscr{D}$ may induce,

[10] Analogously to sets of gambles, our definition of coherent lower prevision derives from a more primitive concept involving lower previsions possibly defined on different domains [Walley, 1991, Section 2.3.2, 2.5.1]. However, to ensure consistency with sets of gambles, we restrict ourselves to the above definition.

in fact, the same coherent lower prevision $\underline{P} = \sigma(\mathscr{D})$ by means of Eq. (1.6). The former are therefore an uncertainty formalism even more general than coherent lower previsions. As a consequence, all of the coherent sets of desirable gambles inducing the same coherent lower prevision induce in turn the same coherent set of almost desirable gambles $\overline{\mathscr{D}} := \overline{\tau}(\underline{P})$ and the same coherent set of strictly desirable gambles $\mathscr{D}^+ := \tau^+(\underline{P})$, which generate again the same lower prevision $\underline{P} = \sigma(\mathscr{D})$. The latter coherent set of strictly desirable gambles can be recovered directly from $\mathscr{D}$ by

$$\mathscr{D}^+ := \tau^+(\sigma(\mathscr{D})) = \{f \in \mathscr{L} : (\exists \delta > 0)\, f - \delta \in \mathscr{D}\} \cup \mathscr{L}^+. \tag{1.10}$$

An important class of coherent lower previsions are the ones corresponding to de Finetti's previsions, the *linear* ones.

Definition 12 (Linear previsions). *Consider a coherent lower prevision $\underline{P}$. Consider also its conjugate upper prevision $\overline{P}$. If $\underline{P}(f) = \overline{P}(f)$ for some $f \in \mathscr{L}$, we call the common value, the* prevision *of f and we denote it by $P(f)$. If this happens for all $f \in \mathscr{L}$, we call the functional P a* linear prevision*.*

Let us denote the set of all the linear previsions as $\mathbb{P}(\Omega)$ or $\mathbb{P}$. From the above definitions, it follows that linear previsions are in particular linear functionals on $\mathscr{L}$.

If $\mathscr{K}$ is a maximal coherent set of gambles, its associated lower prevision $\sigma(\mathscr{K})$ is a linear prevision. For this reason, we can give the following definitions.

Definition 13 (Maximal coherent set of strictly desirable gambles). *Given a linear prevision P, we call the set:*

$$M^+ = \tau^+(P) := \{f \in \mathscr{L} : P(f) > 0\} \cup \mathscr{L}^+(\Omega) = \{f \in \mathscr{L} : -P(-f) > 0\} \cup \mathscr{L}^+(\Omega).$$

maximal coherent set of strictly desirable gambles.

Definition 14 (Maximal coherent set of almost desirable gambles). *Given a linear prevision P, we call the set:*

$$\overline{M} = \overline{\tau}(P) := \{f \in \mathscr{L} : P(f) \geq 0\} = \{f \in \mathscr{L} : -P(-f) \geq 0\}. \tag{1.11}$$

maximal coherent set of almost desirable gambles.

The subsequent lemmas, which have been demonstrated in Appendix A, establish the consistency of these definitions.

Lemma 1. *Maximal coherent sets of strictly desirable gambles are not strictly included in any other coherent set of strictly desirable gambles.*

Lemma 2. *Maximal coherent sets of almost desirable gambles are not strictly included in any other coherent set of almost desirable gambles.*

Coherent lower previsions correspond to lower expectation operators. To analyze this alternative representation in more detail, let us first introduce the following definition.

Definition 15 (**Dominance**). *Given two lower previsions $\underline{P}, \underline{Q}$ defined respectively on $dom(\underline{P}), dom(\underline{Q}) \subseteq \mathscr{L}(\Omega)$, we say that $\underline{Q}$* dominates *$\underline{P}$, if $dom(\underline{P}) \subseteq dom(\underline{Q})$ and $\underline{P}(f) \leq \underline{Q}(f)$ for all $f \in dom(\underline{P})$.*

Each coherent lower prevision $\underline{P}$ is the lower envelope of its set of dominating linear previsions:

$$(\forall f \in \mathscr{L})\ \underline{P}(f) := \inf\{P(f) : P \in \mathscr{M}(\underline{P})\},$$

where

$$\mathscr{M}(\underline{P}) := \{P \in \mathbb{P} : (\forall f \in \mathscr{L})\ P(f) \geq \underline{P}(f)\} \tag{1.12}$$

is, in particular, convex and closed under the *weak* topology* [Walley, 1991, Theorem 3.6.1].[11] There is indeed a one-to-one correspondence between the coherent lower previsions $\underline{P}$ and the non-empty weak*-compact convex sets of linear previsions $\mathscr{M}$: $\underline{P}$ is the lower envelope of its corresponding $\mathscr{M} = \mathscr{M}(\underline{P})$ and $\mathscr{M}$ is the set $\mathscr{M}(\underline{P})$ of all the linear previsions dominating $\underline{P}$.

The restriction of a linear prevision P to the set of all events, obtained by identifying indicators $\mathbb{I}_B$ with their corresponding events B, is simply called *probability* and corresponds to a *finitely additive probability measure* or *probability charge* [Troffaes and de Cooman, 2014]. Linear previsions are then completely determined by the values they assume on events and correspond to expectations [Troffaes and de Cooman, 2014, Section 8.4].[12] On finite spaces, in particular, they are equivalent to the usual expectation operators associated with probability mass functions.

A coherent lower prevision $\underline{P}$ can thus be equivalently represented as the lower envelope of a set of expectation operators $\mathscr{M}(\underline{P})$ which, in turn, can be regarded as a set of finitely additive probabilities (a so-called *credal set*). The same result can be obtained by considering only the extreme points of $\mathscr{M}(\underline{P})$, i.e., $\mathrm{ext}(\mathscr{M}(\underline{P}))$ [Walley, 1991, Section 3.6.2].

For further reference, we also recall that, starting from a coherent set of gambles $\mathscr{D}$, it is possible to directly calculate the set of linear previsions dominating $\underline{P} := \sigma(\mathscr{D})$ as:

$$\mathscr{M}(\underline{P}) = \mathscr{M}(\sigma(\mathscr{D})) = \{P \in \mathbb{P} :\ (\forall f \in \mathscr{D})\ P(f) \geq 0\}, \tag{1.13}$$

[11] It is the smallest topology such that all the evaluation functionals given by $f(P) := P(f)$, where $f \in \mathscr{L}$, are continuous [Walley, 1991, Appendix D3].

[12] This can be expressed, for example, using a *Dunford integral* [Troffaes and de Cooman, 2014, Section 8.7].

see [Walley, 1991, chapter 3.8.4]. The same is valid for a coherent set of almost desirable gambles. Hence, we can use the equivalent notations $\mathscr{M}(\mathscr{D}), \mathscr{M}(\overline{\mathscr{D}})$ to indicate $\mathscr{M}(\sigma(\mathscr{D})), \mathscr{M}(\sigma(\overline{\mathscr{D}}))$ respectively.

In Chapter 4, we will limit ourselves to finite possibility spaces. In that context, we will consider coherent sets $\overline{\mathscr{E}(\mathscr{K})}$ constructed from finite sets $\mathscr{K}$. These sets are called *finitely generated* and are in particular characterised by credal sets with finite sets of extreme points [Walley, 1991, Chapter 4.2].[13] This notion will be then used and generalised to other concepts of coherence in Section 4.3 and Section 4.4.

Example 3. *Let us consider the set* $\mathscr{D} := \mathscr{E}(\{f\})$ *defined in Example 2. We can construct from it the coherent lower prevision* $\underline{P} := \sigma(\mathscr{D})$.

For every linear prevision $P \in \mathbb{P}$, *we have:*

$$(\forall f \in \mathscr{D})\, P(f) \geq 0 \iff P(Tall) \geq P(Low),$$

where $P(B) := P(\mathbb{I}_B)$ *for every* $B \subseteq \Omega$. *Therefore, by equation (1.13), we have:*

$$\mathscr{M}(\underline{P}) = \mathscr{M}(\sigma(\mathscr{D})) = \{P \in \mathbb{P} : P(Tall) \geq P(Low)\}.$$

Thus, $\underline{P}$ *models the beliefs of an agent thinking the murderer is more (or at least equally) probable to be tall than low.*

Axioms LP1, LP2, LP3 characterize coherence of a lower prevision defined on the whole set $\mathscr{L}$. However, similarly to what happens for sets of gambles, a notion of coherence can be established also for lower previsions defined on different sets Q [Walley, 1991, Section 2]. In particular, if Q is a linear space, a lower prevision $\underline{P} : Q \to \mathbb{R} \cup \{+\infty\}$ is said to be *coherent relative to* Q if and only if it respects LP1, LP2, LP3 for gambles in Q [Walley, 1991, Section 2.3.3].

Analogously, a notion of marginalisation and conditioning can be established also for coherent lower previsions. For further reference, we only remind the definition of the former.

Definition 16 (Marginal lower prevision). *Let* $\underline{P}$ *be a coherent lower prevision and consider a partition* $\mathscr{P}$ *of* Ω. *The* $\mathscr{P}$-marginal *of* $\underline{P}$ *is the restriction* $\underline{P}_{\mathscr{P}}$ *of* $\underline{P}$ *to* $\mathscr{P}$*-measurable gambles.*

The $\mathscr{P}$- marginal lower prevision $\underline{P}_{\mathscr{P}}$ of a coherent lower prevision $\underline{P}$ is defined on $\mathscr{L}_{\mathscr{P}}(\Omega)$ and it is coherent relative to $\mathscr{L}_{\mathscr{P}}(\Omega)$. Moreover, given a coherent set of gambles $\mathscr{D}$ and its correspondent coherent lower prevision $\underline{P}$, $\underline{P}_{\mathscr{P}} = \sigma(\mathscr{D} \cap \mathscr{L}_{\mathscr{P}})$ [Miranda and Zaffalon, 2023, pp.192]. For further details, see Walley [1991].

[13] The definition given in [Walley, 1991, Chapter 4.2] of a finitely generated coherent set is more general. For the aim of this thesis, however, we only need to recall this simpler concept.

We conclude this subsection by introducing the *natural extension operator* $\underline{E}$ for lower previsions. Similarly to what happens for gambles, if $\underline{E}(\underline{P})$ is coherent, it corresponds to the minimal coherent lower prevision dominating $\underline{P}$:

$$(\forall f \in \mathscr{L})\ \underline{E}(\underline{P})(f) = \min\{\underline{P}' \in \underline{\mathbb{P}} : \underline{P} \leq \underline{P}'\}(f).$$

Here however, as for sets of gambles, we consider a slightly different operator $\underline{E}^*$ having for lower previsions the same role of the operator $\mathscr{C}$ for sets of gambles. It therefore coincides with $\underline{E}$ when it results in a coherent lower prevision, i.e., when it is applied to a lower prevision for which there exists at least a coherent lower prevision dominating it. Otherwise, it corresponds to $\sigma(\mathscr{L})$, where $\sigma(\mathscr{L})(f) := +\infty$ for every gamble $f \in \mathscr{L}$.[14]

Definition 17. *Given a lower prevision $\underline{P}$ on a domain $dom(\underline{P}) \subseteq \mathscr{L}$, we define*

$$\underline{E}^*(\underline{P}) := \begin{cases} \underline{E}(\underline{P}) & \text{if } (\exists \underline{P}' \in \underline{\mathbb{P}})\ \underline{P} \leq \underline{P}', \\ \sigma(\mathscr{L}) & \text{otherwise.} \end{cases}$$

In view of the following Chapter 3, again similarly to sets of gambles, it is convenient to consider $\underline{\Phi}(\Omega) := \underline{\mathbb{P}}(\Omega) \cup \{\sigma(\mathscr{L}(\Omega))\}$. We can also refer to it with $\underline{\Phi}$, if there is no possible ambiguity. For further reference, we also enlarge the map τ^+ such that it is defined on $\underline{\Phi}$ with values in Φ^+, by defining $\tau^+(\sigma(\mathscr{L})) := \mathscr{L}$.

1.1.2 Desirability and preference relations

In the past century, the works of von Neumann and Morgestern [1947]; Anscombe and Aumann [1963]; Savage [1972], through their justification of rational decision making via expected utility maximization, have represented a major achievement in the development of an axiomatic treatment of preferences. This view has had an enormous impact in a wide range of research fields being at the basis of any process of decision making.

Although von Neumann, Morgenstern, Anscombe, and even Aumann recognised the impracticality of demanding that an agent can always compare alternatives, the axiomatization of rational decision-making with *incomplete* preferences emerged much later through the works of of [Seidenfeld et al., 1995; Bewley, 2002; Nau, 2006; Galaabaatar and Karni, 2013], built upon the analytical framework of Anscombe and Aumann.

[14] It's worth noting that we could have circumvented the introduction of $\underline{E}^*$ by adopting the convention that the minimum of the empty set is $+\infty$. Nevertheless, to maintain parallelism with sets of gambles for which we introduced the operator $\mathscr{C}$ alongside the standard natural extension operator $\mathscr{E}$, we chose to introduce this additional operator $\underline{E}^*$ for lower previsions.

Preferences and desirability have essentially lived two separate lives. Preferences indeed can deal both with probabilities and utilities, while desirability is born as a theory of uncertainty alone. Recently, however, Zaffalon and Miranda enriched desirability with a set of possible prizes, thus generalising it to deal also with utility considerations. At this stage, they prove it to be essentially equivalent to the traditional axiomatization of incomplete preferences à la Anscombe-Aumann.

In Zaffalon and Miranda [2017], in particular, they present their results limited to a finite set of prizes. In Zaffalon and Miranda [2021], they extend their construction to the general case. In their very latest work Miranda and Zaffalon [2023] moreover, they consider again the problem of modeling general rational decision-making with desirability by providing a more direct solution based on the generalisations of its founding axioms. This new version of desirability, in particular, permits to solve problems related to the traditional axiomatization of preferences, such as the ones pointed out by the well-known *Allais paradox* [Allais, 1953]. The same idea has also been considered in our previous works Casanova, Benavoli and Zaffalon [2021]; Casanova et al. [2023]. Here, we only focus on some examples of axiomatisations weaker than traditional desirability but we assign them an operational meaning based on their reformulation as classification problems. More details will be given in the next Chapter 4.

The remaining part of this subsection will be instead dedicated to illustrate in more details results provided in Zaffalon and Miranda [2017, 2021], which will play an important role in our discussion.

In order to deal with preferences, Zaffalon and Miranda first enrich desirability with a set of possible prizes. Let us consider therefore a new space $\mathscr{Z} := \Omega \times \mathscr{X}$, where Ω represents the usual space of possibilities and $\mathscr{X}$ represents a set of prizes. We assume that all the pairs of elements in $\Omega \times \mathscr{X}$ are possible or, which is equivalent, that Ω and $\mathscr{X}$ are *logically independent*.

The treatment of preferences in Zaffalon and Miranda [2021] relies on the notion of *conditional horse lottery*.

Definition 18 (Conditional horse lottery). *A* conditional horse lottery *is a function* $p : (\mathscr{Z} := \Omega \times \mathscr{X}) \to [0,1]$.

Conditional horse lotteries, also called *acts* for simplicity, are a generalisation of the well-known *horse lotteries* of Anscombe and Aumann [1963], more suitable to deal with infinitely many prizes [Zaffalon and Miranda, 2021, Section 5]. Traditional horse lotteries indeed correspond to particular conditional horse lotteries that place a probability over prizes for each $\omega \in \Omega$. The process through which an agent is rewarded in the context of conditional horse lotteries is somewhat more intricate and relies on the idea of having compound lotteries. For more details see Zaffalon and Miranda, 2021, Section 4.

Let us indicate the set of all acts with $\mathscr{H}$. In what follows, it will be useful to consider the *zero act*.

Definition 19 (Zero act). *Let $0 \in \mathscr{H}$ denote the zero act, which is defined by $0(\omega, x) := 0$, for all $\omega \in \Omega$, $x \in \mathscr{X}$.*

An agent will prefer some acts to others, depending on their beliefs on Ω and their attitude towards prizes in $\mathscr{X}$. Particularly interesting are *coherent* preference relations.

Definition 20 (Coherent preference relation). *A preference relation $\succ$ over conditional horse lotteries is a subset of $\mathscr{H} \times \mathscr{H}$. It is said to be coherent if it satisfies the next four axioms:*

A1. $(\forall p \in \mathscr{H} \setminus \{0\})\ p \succ 0$ *[Worst Act];*

A2. $(\forall p \in \mathscr{H})\ p \not\succ p$ *[Irreflexivity];*

A3. $(\forall p, q, r \in \mathscr{H})\ p \succ q \succ r \Rightarrow p \succ r$ *[Transitivity];*

A4. $(\forall p, q, r \in \mathscr{H})\ p \succ q \iff (\forall \alpha \in (0, 1])\ \alpha p + (1-\alpha)r \succ \alpha q + (1-\alpha)r$ *[Mixture Independence].*

If also the next axiom is satisfied, then we say that the coherent preference relation is weakly Archimedean.

A0. $(\forall p, q \in \mathscr{H} : \neg(p \gneq q)),\ p \succ q \Rightarrow (\exists \alpha \in (0, 1))\ \alpha p \succ q$ *[Weak Archimedeanity].*[15]

Now we can list Zaffalon and Miranda [2021]'s main results. Before recalling them, we need to remind the reader a further definition.

Definition 21 (Linear preference relation). *We say that a relation $\succ \subseteq \mathscr{H} \times \mathscr{H}$ is linear if $p \succ q \Rightarrow r \succ s$ for all $r, s \in \mathscr{H}$ such that $p - q = r - s$.*

Linear preference relations are not necessarily coherent. However, it is possible to discuss additional axioms they can satisfy in relation to those of desirability. First, we need to recall the following result that correspond to Lemma 3 of Zaffalon and Miranda, 2021.

[15] With $p \gneq q$ meaning $p \geq q$, $p \neq q$.

Lemma 3. *Let $\mathscr{L}_1(\mathscr{X}) := \{f \in \mathscr{L}(\mathscr{X}) : \sup|f| \leq 1\}$ and $\mathscr{L}_1^+(\mathscr{X}) := \mathscr{L}_1(\mathscr{X}) \cap \mathscr{L}^+(\mathscr{X})$. Sets of gambles in $\mathscr{L}_1(\mathscr{X})$ and linear preference relations in $\mathscr{H} \times \mathscr{H}$ are in a one-to-one correspondence, where given a set of gambles $R \subseteq \mathscr{L}_1(\mathscr{X})$, we consider the preference relation $\succ$ as $p \succ q \iff p - q \in R$ and given $\succ$ we define $R := \{f \in \mathscr{L}_1(\mathscr{X}) : (\exists p, q \in \mathscr{H})\, f = p - q, p \succ q\}$.*

Now, we can recall their main result that correspond to Lemma 4 of Zaffalon and Miranda, 2021.

Theorem 1. *Let R and $\succ$ be respectively a set of gambles in $\mathscr{L}_1(\mathscr{X})$ and the corresponding linear preference relation. Then:*

1. *Relation $\succ$ has worst outcome* 0 *if and only if R* accepts partial gains relative to *$\mathscr{L}_1(\mathscr{X})$, i.e., if and only if $\mathscr{L}_1^+(\mathscr{X}) \subseteq R$;*

2. *Relation $\succ$ is irreflexive if and only if R* avoids status quo*, i.e., if and only if $0 \notin R$;*

3. *Relation $\succ$ is transitive if and only if R satisfies:*
$$(\forall f, g \in R : f = p - q,\ g = q - r \text{ for some } p, q, r \in \mathscr{H})\ f + g \in R; \quad (1.14)$$

4. *Relation $\succ$ satisfies the mixture independence axiom if and only if R satisfies* positive homogeneity relative to *$\mathscr{L}_1(\mathscr{X})$, i.e., if and only if, if $f \in R$ and $\lambda > 0$ such that $\lambda f \in \mathscr{L}_1(\mathscr{X})$ then $\lambda f \in R$;*

5. *Relation $\succ$ is weakly Archimedean if and only if R satisfies:*
$$\begin{gathered}(\forall f \in R \setminus \mathscr{L}_1^+(\mathscr{X}), g \in \mathscr{L}_1(\mathscr{X}) : \max(0, -f) \leq g \leq \min(1, 1 - f)) \\ \exists \alpha \in (0,1) : \alpha f - (1 - \alpha) g \in R.\end{gathered} \quad (1.15)$$

The following lemma instead corresponds to Lemma 5 of Zaffalon and Miranda, 2021.

Lemma 4. *Let R be a set of gambles in $\mathscr{L}_1(\mathscr{X})$. If it satisfies* positive homogeneity relative to *$\mathscr{L}_1(\mathscr{X})$, then it satisfies also Eq. (1.14) if and only if it is* additive relative to *$\mathscr{L}_1(\mathscr{X})$:*
$$(\forall f, g \in R : f + g \in \mathscr{L}_1(\mathscr{X}))\ f + g \in R.$$

By the previous results, it follows that if $\succ$ is a coherent preference relation on $\mathscr{H} \times \mathscr{H}$, then the correspondent set $R \subseteq \mathscr{L}_1(\mathscr{X})$ is coherent relative to $\mathscr{L}_1(\mathscr{X})$. The converse is also true. A further one-to-one correspondence can be established between coherent preference relations on $\mathscr{H} \times \mathscr{H}$ and properly coherent sets of gambles, as stated by the next results that summarise the main outcomes of Theorem 6 and 7 of Zaffalon and Miranda, 2021.

Theorem 2. *If a set $R \subseteq \mathscr{L}_1(\mathscr{X})$ accepts partial gains relative to $\mathscr{L}_1(\mathscr{X})$, avoids status quo, satisfies Eq. (1.14) and positive homogeneity relative to $\mathscr{L}_1(\mathscr{X})$, it is coherent relative to $\mathscr{L}_1(\mathscr{X})$ and the set $\mathscr{D}_R := \mathscr{E}(R)$ is coherent. If the former satisfies also Eq. (1.15), then $\mathscr{D}_R$ is a coherent set of strictly desirable gambles.*

Vice versa, given a coherent set of gambles $\mathscr{D} \subseteq \mathscr{L}(\mathscr{X})$, the set

$$R_{\mathscr{D}} := \{\frac{f}{\sup|f|} : 0 \neq f \in \mathscr{D}\} \subseteq \mathscr{L}_1(\mathscr{X})$$

accepts partial gains relative to $\mathscr{L}_1(\mathscr{X})$, avoids status quo, satisfies Eq. (1.14) and positive homogeneity relative to $\mathscr{L}_1(\mathscr{X})$. Hence, it is coherent relative to $\mathscr{L}_1(\mathscr{X})$. If $\mathscr{D}$ is also a coherent set of strictly desirable gambles, then $R_{\mathscr{D}}$ satisfies Eq. (1.15).

Corollary 1. *Coherent preference relations on $\mathscr{H} \times \mathscr{H}$ and coherent sets of gambles in $\mathscr{L}(\mathscr{X})$ are in a one-to-one correspondence, where:*

- *given a coherent preference relation $\succ$ on $\mathscr{H} \times \mathscr{H}$, the set*
$$\mathscr{D} := \mathscr{E}(\{f \in \mathscr{L}_1(\mathscr{X}) : (\exists p, q \in \mathscr{H})\, f = p - q, p \succ q\})$$
is a coherent set of gambles;

- *given a coherent set of gambles $\mathscr{D}$, the preference relation $\succ$ defined as*
$$p \succ q \iff p - q \in R_{\mathscr{D}} := \{\frac{f}{\sup|f|} : 0 \neq f \in \mathscr{D}\} \subseteq \mathscr{L}_1(\mathscr{X})$$
is a coherent preference relation on $\mathscr{H} \times \mathscr{H}$.

Moreover, the same correspondence links coherent weakly Archimedean preference relations on $\mathscr{H} \times \mathscr{H}$ and coherent sets of strictly desirable gambles in $\mathscr{L}(\mathscr{X})$.

It is possible to notice, in particular, that the mixture independence axiom satisfied by a coherent preference is linked with the linearity axioms of desirability, i.e., those based on the linearity assumption of the scale in which rewards of gambles are measured.

The mixture independence axiom has been highly criticised in literature since there are famous examples where people often violate it, see for example the well known *Allais paradox* [Allais, 1953]. This was also the starting point for *non-expected utilities* theories, such that *rank dependent utility* (see Quiggin, 1982), *weighted expected utility* (see Chew and MacCrimmon, 1979a,b) and *prospect theory* (see Kahneman and Tversky, 2013), which provide a solution for these paradoxical situations.

We can conclude therefore that even though desirability is a very powerful theory of decision making, the linearity assumption on the scale used to measure gambles' rewards limits its capability to represent general rational decision making. In Chapter 4, examples of different generalisations of the axiomatic definition of desirability are provided to relax the above-mentioned linearity assumption. An even more general solution is instead provided in Miranda and Zaffalon [2023], where desirability is extended to the nonlinear case by letting the utility scale be represented via a general closure operator. It is then proven that the Allais parodox finds a solution in this more general theory.

At now, we have established an equivalence result for desirability and preferences at the level of sets of desirable gambles. However, as usually done in literature, it is also possible to directly work with a probability-utility representation. This means essentially to deduce from a coherent set of gambles the corresponding lower prevision. More precisely, given a coherent set of gambles $\mathscr{D} \subseteq \mathscr{L}(\Omega \times \mathscr{X})$, we obtain its corresponding coherent lower prevision $\underline{P} := \sigma(\mathscr{D})$ for all $f \in \mathscr{L}(\Omega \times \mathscr{X})$, using Eq.(1.6). $\underline{P}(f)$ is interpreted as the lower expectation of f taken with respect to the probabilities and utilities that are implicit in the definition of $\mathscr{D}$, see Section 1.1.1. Unless $\underline{P}$ is subject to the property of *state independence* [16] however, considerations of probability and utility cannot be disentangled. In the general case of state dependence, it is needed to use the 'joint' model $\underline{P}$ directly. A lower prevision can be deduced also from sets of gambles satisfying axioms of *nonlinear coherence*. Properties of such lower previsions are analysed in Chapter 4 and in Miranda and Zaffalon [2023].

1.1.3 Desirability and logic

The idea of coherence behind desirability can be captured more vividly by exploiting a parallel with logic. That desirability and logic are intimately related is already quite intuitively clear by considering the former as a paradigm for *conservative (probabilistic) inference*. Such a conservative inference can be represented by considering desirability statements: given an agent who considers a set of gambles $\mathscr{K}$ to be desirable, we can infer which other gambles are desirable by calculating its natural extension $\mathscr{E}(\mathscr{K})$, i.e., by following the *coherence rules* D1, D3, D4 [de Cooman et al., 2023]. In more formal terms, in Troffaes and de Cooman [2014], a parallelism between the inference mechanism behind natural extension and the one of classical propositional logic was established proving in particular that the former subsumes the latter. Still on the relation between

[16] *State independence* is the condition that allows us to have separate models for beliefs and values. See the following Definition 34 for the precise case. The latter definition can then be extended to multiple probability-utility pairs unless both Ω and $\mathscr{X}$ are infinite. In this case, we have to rely on the asymmetric notion of *irrelevance*, see Zaffalon and Miranda, 2021, Appendix A.1.1.

imprecise probabilities and propositional logic, another result in the same work proves the latter can be formally embedded into the theory of coherent lower previsions.

The parallel between desirability and logic can be clarified in more general terms by considering the abstract *algebraic* view of logic introduced in the works of Tarski (see, e.g., Tarski, 1983, Chapter 5). According to this view, we can define a *logical system* simply as a pair $(\boldsymbol{L}, \boldsymbol{Cl})$, composed by a language $\boldsymbol{L}$ and a *closure (or consequence) operator* $\boldsymbol{Cl}$ on $\boldsymbol{L}$, i.e., a mapping $\boldsymbol{Cl} : \mathrm{P}(\boldsymbol{L}) \to \mathrm{P}(\boldsymbol{L})$ respecting properties analogous to CL1–CL3 for elements of $\mathrm{P}(\boldsymbol{L})$, see for example Kohlas, 2003, Lemma 6.20.[17] A theory $\mathbf{T}$ in a logical system $(\boldsymbol{L}, \boldsymbol{Cl})$ is defined as any *closed* subset of $\boldsymbol{L}$, i.e., any $\boldsymbol{T} \subseteq \boldsymbol{L}$ such that $\boldsymbol{T} = \boldsymbol{Cl}(\boldsymbol{T})$. A theory $\boldsymbol{T}$ is called *consistent* if and only if it is a proper subset of $\boldsymbol{L}$.

In general, the nature of the language $\boldsymbol{L}$ varies depending on the specific kind of logical system one considers. For example, in propositional logic $\boldsymbol{L}$ includes all the *well-formed formulae* obtained by composing atoms $P, Q, R, \dots$ from a possibly infinite collection of symbols through the Boolean connectives $\neg$, $\wedge \vee$, $\to$ [Hodges, 2001].

A closure operator $\boldsymbol{Cl}$ is usually specified through a *proof system*, i.e., a finite set of rules that permit to construct the elements of $\boldsymbol{Cl}(\boldsymbol{K})$ for any given $\boldsymbol{K} \subseteq \boldsymbol{L}$. For propositional logic, for example, different proof systems have been proposed [Buss, 1998], which include rules like the well-known *modus ponendo ponens* (from P and $P \to Q$ infer Q). In the context of desirability, a proof system can be established by adhering to the conventional rules of coherence, as demonstrated in Wilson and Moral [1994]. In this paper, the authors undertake a formal redefinition of desirability as a logical system comprising a language $\boldsymbol{L}$, a closure operator $\boldsymbol{Cl}$ defined through a proof system, and an additional layer of enrichment through a *semantics*. A semantic provides a further interpretation of the elements of the language that is usually employed for proving the internal consistency of the system (for a more detailed explanation of the role of semantics in logic, see Hodges [1997]).

In our study, we likewise view desirability as a logical system $(\boldsymbol{L}, \boldsymbol{Cl})$, where $\boldsymbol{L}$ corresponds specifically to the set of all gambles, as proposed in Wilson and Moral [1994]. The role of the $\boldsymbol{Cl}$ operator can be assigned to either the natural extension or the operator $\mathscr{C}$ introduced in Eq. (1.4). When $\mathscr{C}$ is considered as the closure operator, coherent sets of gambles can be identified as the consistent theories within the system.

[17] In contrast to Tarski [1983], we adopt a broader definition of closure operator. This definition is now widely accepted in the literature since it permits to encompass a wider range of logic systems, see for example Jansana, 2022, Section 1.

It's important to note that in this thesis, we do not delve into the formal details of a proof system or semantics. Instead, the logical treatment of desirability serves primarily as a means to explore the possible connections between desirability, viewed as a logical system in the algebraic Tarskian sense, and the algebraic formalism of the *information algebras*, as discussed in Casanova, Kohlas and Zaffalon [2021]; Kohlas et al. [2021]; Casanova et al. [2022a,b] and Chapter 3 of this thesis.

Information algebras are general algebraic structures introduced to manage knowledge or information and subsuming many formalisms in computer science both inspired by probability and logic [Kohlas, 2003]. Notably, they also arise from particular *information systems*, i.e., tuples $(\boldsymbol{L}, \boldsymbol{Cl}, \mathfrak{S})$ composed by a language $\boldsymbol{L}$, a closure operator $\boldsymbol{Cl}$ and a family of sublanguages $\mathfrak{S}$, i.e., subsets of the language satisfying certain properties.[18] The similarities existing between logical and information systems allow us to induce instances of these structures from desirability. This result enriches the view of the latter as an algebraic structure and permits to further generalise the treatment of some well-known problems approached with desirability, see for example the *marginal problem* in Section 3.1.3 and Section 3.2.2. An interesting further development of this research line would consist in proving a sort of converse statement, i.e., that information algebras can be included in a generalised version of desirability (notably, a 'minimal' one obtained by eliminating the linearity assumption underlying their defining axioms, see Casanova, Benavoli and Zaffalon [2021]; Miranda and Zaffalon [2023]). The latter would result in a very general framework for conservative inference encompassing (and bridging) both probability calculus and the various formalisms that can be modelled via information algebras, including propositional logic and various kinds of other logical formalisms of great relevance for artificial intelligence [Wilson and Mengin, 1999].

We conclude this subsection by reminding the reader that an order-theoretic structure similar in spirit to information algebras and underlying many of the models for representing beliefs in the literature was also discovered by de Cooman [2005]. In particular, this structure subsumes classical propositional logic and other belief models, such as some based on imprecise probabilities. Proper connections between information algebras induced by desirability and belief structures as inroduced by de Cooman [2005] still need to be established.

[18] A related notion of information system in domain theory has been introduced by Scott, see for example Davey and Priestley [2002]. Here, however, we refer to the definition given in [Kohlas, 2003, Section 6.4].

1.2 Summary

In this chapter we introduced the main definitions and tools of *desirability*, an extension of de Finetti's Bayesian theory made to deal with imprecision in probabilities.

Given a *possibility space* Ω for an experiment, i.e., a set denoting all its mutually exclusive outcomes, desirability models a subject's uncertainty about the experiment by focusing on *gambles*, i.e., bets on Ω, which an agent is disposed to accept. Sets of *acceptable* gambles are considered *rational* when they adhere to a specific set of axioms and, in this case, they are referred to as *coherent sets of gambles*. Central to desirability is the *natural extension* operator, which deduces additional acceptable gambles to satisfy these axioms without further commitments.

Uncertainty expressed by coherent sets of gambles can be formalized using probabilistic models on the possibility space. These models involve *lower* and *upper previsions*, representing lower and upper envelopes of expectation operators calculated with respect to *finitely additive probabilities*, known as *credal sets*. Coherent lower and upper previsions can also be equivalently represented by their corresponding credal sets or the extreme points of these sets. While coherent sets of gambles are more general than lower and upper previsions, two slightly different variants can be introduced: *coherent sets of strictly desirable gambles* and *coherent sets of almost desirable gambles*, which are, in fact, equivalent.

Although initially conceived as a theory of uncertainty, desirability has recently expanded its scope to include considerations of value. When desirability models are extended to incorporate a set of potential prizes, they effectively align with traditional incomplete preferences widely employed in decision-making contexts. Desirability therefore results to be both a very powerful uncertainty theory and a general tool for decision making.

From a logical perspective, the natural extension operator functions as a deductive inference tool or a *closure operator* for desirability. This perspective enables us to interpret coherent sets of gambles as closed and consistent sets within a logical system and formally prove important parallelisms between desirability, and *imprecise probabilities* in general, and classical propositional logic. In particular, indeed, classical propositional logic can be formally embedded in the theory of coherent lower previsions.

In the following sections, we will apply these concepts to present the findings of my PhD research. Further details regarding the upcoming discussion are provided below.

1.3 Contributions

Henceforth, the main contributions of the present work are annotated with reference to the publications in which they are presented.

1. In Casanova et al. [2020]; Casanova, Miranda and Zaffalon [2021], we formulate a general problem of *opinions aggregation* where both individual and group's opinions are modeled through coherent sets of gambles. This approach leads to a new tool to compare different formulation of the problem as given in the literature, and provide a new perspective of traditional results. In particular, we studied it in connection with *social choice* and *opinion pooling*, two research fields dealing with opinions aggregation but historically approaching it with different mathematical instruments. In this context, the capability of desirability to model incomplete information turns out to be a key factor to escape from classical impossibility results, such as the ones imposed by the *Arrowian* framework, and permit the preservation of valuable properties, such as probabilistic independence, through the aggregation process.

2. In Casanova, Kohlas and Zaffalon [2021]; Kohlas et al. [2021]; Casanova et al. [2022a,b], we provide a way to bridge desirability and *information algebras*. The latter are general algebraic structures introduced to manage information at an abstract level, providing also basic operations and architectures for inference. Both desirability and information algebras can be analysed from a logical perspective. On the one hand, coherent sets of gambles can be interpreted as closed theories in a logical system. On the other hand, information algebras can be alternatively represented as particular *information systems*. Starting from these observations, we proceed by verifying that the former formally induce instances of the latter. The obtained result enriches the view of desirability as an algebraic (logic-like) structure and enhances it with the inference tools of information algebras. Conversely, desirability provides general instances of information algebras able to manipulate very different types of information.

3. In Casanova, Benavoli and Zaffalon [2021]; Casanova et al. [2023], we explore possible generalisations of desirability.

   Desirability is a powerful theory of uncertainty and decision making. At the core of its axiomatic formulation, however, lies a linearity assumption on the scale used to measure gambles' rewards that limits its modeling capabilities and conflicts with a general representation of rational decision making. For these main reasons, we analyse possible generalizations of the founding axioms of desirability and make the theory able to deal directly with bets whose rewards are not necessarily expressed in a linear currency. For each set of axioms examined, we also provide a practical tool to check them for a generic agent providing finite sets of acceptable and rejectable gambles. The latter is based on an alternative interpretation of the problem as a binary (usually nonlinear) classification task, thus yielding a different interpretation of desirability. Finally, by borrowing ideas from machine learning, we also propose possible feature mappings that transform the nonlinear classification problems into linear ones defined in higher-dimensional spaces.

1.4 Organization of the thesis

This thesis opens with Chapter 1, where main definitions and results of desirability are introduced. Then, three chapters follow covering the research lines presented above. In particular, in Chapter 2 a general problem of opinions aggregation is formulated with instruments of desirability; in Chapter 3, basic tools of desirability are shown to provide general instances of information algebras; in Chapter 4, examples of weaker axiomatisations of standard desirability are provided and reformulated as classification problems. Chapter 5 finally outlines some insights for possible future investigations concerning the various lines of research developed.

Chapter 2

Social choice and opinion pooling with desirability

Social choice theory concerns the definition of 'social' functions, called *social welfare functions* [Feldman and Serrano, 2006; Weymark, 1984], to best aggregate preferences of a group of rational voters. *Arrow's impossibility theorem* (see Arrow, 1951) is generally acknowledged as the basis of modern social choice theory; it establishes limits to what is rationally possible to do while avoiding dictatorial solutions to the aggregation problem.

To understand its statement, let us return to the framework of Example 1. Let us consider therefore the possibility space $[1.5, 2]m$, corresponding to the set of the possible values for the height of a murderer. Suppose now to partition it considering three alternatives: (a) height $\in [1.85, 2]m$; (b) height $\in [1.70, 1.85)m$; and (c) height $\in [1.5, 1.70)m$. Assume moreover that two journalists, Bill and Clark, and the detective Alice of Example 1 have different opinions on the matter.

Opinions of rational agents in social choice are expressed through *complete* and *transitive* preference relations over a set of alternatives: a *complete* preference relation allows to always choose between any two options; a *transitive* one implies that if an alternative x is preferred to y that in turn is preferred to z, then x is preferred to z. Transitivity, in particular, is often regarded as a non-negotiable requirement as it appears to express a fundamental property of preferences. We suppose therefore that Alice, Bill and Clark express their views by ranking the three options a,b,c as in Figure 2.1 (each set of rankings is called a *profile* for the group).

What Arrow's theorem states is that it is impossible to have a social welfare function (i.e., a function that aggregates a profile into a group opinion) leading again to a complete and transitive preference relation over alternatives, which

	a	b	c
Alice	1	2	3
Bill	3	1	2
Clark	2	3	1

Figure 2.1. A profile.

satisfies in addition some arguably necessary desiderata (*unlimited maximal domain*.[1], i.e., all complete and transitive preference relations for the voters are allowed; *weak Pareto*, i.e., given alternatives x, y, if every voter prefers alternative x over y, then the group prefers x over y; *independence of irrelevant alternatives* or *i.i.a.*, i.e., if every voter's preference between x and y remains unchanged, then the group's preference between x and y will also remain unchanged) and that is not a *dictatorship*, i.e., it does not give the possibility to a single voter, the *dictator*, to always determine the group's preferences. To escape this paradox, various theorists have suggested weakening some of the rationality criteria required. For example, eliminating completeness of the group preferences, we can be led in some cases to the existence of: an *oligarchy*, i.e., a group of persons that always determine the group's preferences or a *democracy*, where the whole society forms the oligarchy (see Weymark, 1984; Pini et al., 2008).

The questions addressed by opinion pooling are similar to those of social choice. The former, however, is expressed in a probabilistic framework: opinion pooling deals with the aggregation of probabilistic opinions on some events of interest (see Lindley et al., 1979). Fig. 2.2, for example, illustrates a situation where Alice, Bill and Clark express different probabilistic opinions about the murderer's height and the goal is to aggregate them.

	P(a)	P(b)	P(c)
Alice	0.8	0.15	0.05
Bill	0.2	0.5	0.3
Clark	0.34	0.31	0.35

$\longrightarrow P(a), P(b), P(c)$ of the group?

Figure 2.2. Different probabilistic opinions about three alternatives.

Similarly to social choice, opinion pooling proposes an axiomatic approach for defining a pooling function. In particular, different classes of pooling func-

[1] In literature it is indicated more simply with the term *unlimited domain*. We add the term *maximal* to distinguish it to the more general case in which we allow all transitive rankings, not necessarily complete.

tions and their properties are studied, such as *linear* (see Aczél and Wagner, 1980), *geometric*, or *multiplicative* (see Dietrich, 2010). Precise probabilistic approaches to opinion pooling however can present some difficulties, such as the impossibility to preserve independence of events. Considering group opinions represented by imprecise probabilities can solve the problem [Stewart and Quintana, 2018]. A strong appeal for imprecise probability in opinion pooling was advocated long ago also by Walley along with a deep analysis of the subject in Walley [1982], where he discusses in detail a number of properties that an aggregation rule for imprecise beliefs may or may not satisfy.

Although social choice and opinion pooling have a number of distinct features, they both regard the aggregation of *opinions*, here intended in broad sense, of a number of experts, expressed as preferences over a number of alternatives or uncertainty models about some experiment, into a global opinion that, at least, do not generate inconsistencies, i.e., do not generate intransitive preference relations or numbers not respecting axioms of probability. This gives us the idea of a general framework for opinions' aggregation based on desirability. It indeed permits to unify many other frameworks under a same idea of consistency. More technically, it allows the aggregation of opinions expressed in different forms including (incomplete) preference relations and (sets of) probability measures. Moreover, since traditionally the ability to model incompleteness is proven to be the key to escape from traditional impossibility results and preserve important properties in the aggregation process, desirability results very apt to the aim. Finally, sets of gambles permit us to simultaneously deal with considerations of beliefs and values, to consider any domain and possibility space and are easier to work with and more general than sets of probability measures or coherent lower previsions. On this basis, in Casanova et al. [2020]; Casanova, Miranda and Zaffalon [2021], we exploit desirability as a framework for a general opinions' aggregation problem.

In the next sections we discuss our results in more detail. In particular, in Section 2.1, we explore the impact of rationality axioms typically expected from aggregation functions in the literature. In this regard, we give an alternative formulation of the limits provided by Arrow's theorem and its generalisation in desirability terms and we analyse their validity in our context. Our analysis leads us to the conclusion that the only democratic aggregation rule that satisfies the requirements imposed is Walley [1982]'s *unanimity rule*.

Turning our attention to Section 2.2, we shift our focus on *coherence*, i.e., on the minimal conditions necessary to achieve a group opinion that is *consistent* with those of the individual agents in our framework. We find that the *weak Pareto* axiom can serve this purpose effectively. Specifically, we demonstrate that aggregation functions satisfying this axiom encompass in particular *convex opinion pooling*, a specialized instance of Walley's unanimity rule that allows for the aggregation of a set of probability distributions into a credal set. This not only reinforces the rationality of the unanimity rule as an aggregating function but also emphasizes the proper role of convex opinion pooling, which is already recognized for its ability to preserve critical properties throughout the aggregation process. Our analysis culminates with an examination of the aggregation of expert opinions based on probability and (state-independent) utility, a common scenario in practical applications. In this case as well, we identify conditions that can lead to dictatorship.

The proof of all the discussed material can be found in Appendix B.

2.1 Desirability reformulation of social choice

Given the flexibility and generality of desirability, in our works Casanova et al. [2020]; Casanova, Miranda and Zaffalon [2021] we chose to analyse a problem of opinions' aggregation by representing both the agents' and the group's points of view as coherent sets of gambles defined on $\mathscr{Z} := \Omega \times \mathscr{X}$, as defined in Section 1.1.2. In this way we can also deal with opinions involving not only beliefs but also utility considerations without encountering any added complications, as we technically treat $\mathscr{Z}$ just as a possibility space.

In this context, we first analyze whether and how results predicted by Arrow's theorem and its generalisations transform in our context. Notably, to examine this issue, we give a re-formulation, similar in spirit, of the main concepts of social choice with desirability.

We start by re-defining individuals' profiles. To this end, let us consider a set of n 'individuals' $\mathscr{V}$, whose opinions are represented by coherent sets of gambles. Individuals' profiles can be defined as follows.

Definition 22 (Profiles). *Let $\mathbb{D}(\mathscr{Z})$ be the set of coherent sets of gambles on $\mathscr{Z}$, as defined in Section 1. $\mathbb{D}^n(\mathscr{Z})$, the n-times Cartesian product of $\mathbb{D}(\mathscr{Z})$, is the set of possible* profiles *of individual sets of gambles.*

A profile of coherent sets of gambles is thus a vector $[\mathscr{D}_i]_{i\in\mathscr{V}} \in \mathbb{D}^n(\mathscr{Z})$, where $\mathscr{D}_i$ is the coherent set of gambles representing the opinion of individual i. To keep the notation simple, we shall often denote it simply by $[\mathscr{D}_i]$.

Now we re-define the concept of *social welfare function*, which we rename in this context as *social rule*, and all the properties it may satisfy (for the original concepts and definitions see for example Weymark, 1984).

Definition 23 (Social rule). *A* social rule Γ *is a function from a set* $\mathscr{A} \subseteq \mathbb{D}^n(\mathscr{X})$ *to a coherent set of gambles (social coherent set of gambles). Its domain* $\mathscr{A}$ *is called the* admissible set of profiles *for* Γ.

Social coherent sets of gambles remain unsubscripted while subscripts distinguish individuals' coherent sets of gambles.

Example 4. *Consider the opinions of three agents Alice, Bill and Clark expressed by means of the coherent sets of gambles* $\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3$ *on a set* $\mathscr{X}$ *with at least three elements. The following are four instances of social rules.*

- $(\forall \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3 \in \mathbb{D}(\mathscr{X}))\ \Gamma_1(\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3) = \mathscr{D}_1$*;*
- $(\forall \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3 \in \mathbb{D}(\mathscr{X}))\ \Gamma_2(\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3) = \mathscr{D}_2 \cap \mathscr{D}_3$*;*
- $(\forall \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3 \in \mathbb{D}(\mathscr{X}))\ \Gamma_3(\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3) = M_2(\mathscr{D}_2)$*, where* $M_2(\mathscr{D}_2)$ *is a maximal coherent set of gambles that includes* $\mathscr{D}_2$*;*
- $(\forall \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3 \in \mathbb{D}(\mathscr{X}))\ \Gamma_4(\mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3) = \mathscr{E}(\mathscr{D}_2 \setminus (\mathscr{D}_1 \cup \mathscr{D}_3))$.

Next we consider a number of additional properties that a social rule may satisfy. From our formulation, such a rule turns the, possibly imprecise, assessments of a number of individuals into an aggregated assessment, which may be imprecise too. As a particular case of interest, we may consider the one where the aggregated set represents precise assessments:

Definition 24 (Completeness). *A social rule* Γ *satisfies* completeness *if and only if* $\Gamma([\mathscr{D}_i])$ *is a maximal coherent set of gambles for every admissible profile* $[\mathscr{D}_i]$*, and it satisfies* strict completeness *if and only if* $\Gamma([\mathscr{D}_i])$ *is a maximal coherent set of strictly desirable gambles for every admissible profile* $[\mathscr{D}_i]$.

Another important assumption we shall consider is that Γ can be applied to any profile of individuals representing imprecise or precise assessments.

Definition 25 (Unlimited domain). *A social rule* Γ *satisfies* unlimited domain *if and only if* $\mathscr{A} = \mathbb{D}^n(\mathscr{X})$*, where* $\mathscr{A}$ *denotes the admissible set of profiles for* Γ*. It is said to satisfy* unlimited maximal domain *when* $\mathscr{A} = \mathbb{M}^n(\mathscr{X})$*, the n-times Cartesian product of* $\mathbb{M}(\mathscr{X})$.

The social rules in Example 4 satisfy unlimited domain by definition. However, only Γ_3 is complete.

Next we consider *independence of irrelevant alternatives*.

Definition 26 (Independence of irrelevant alternatives). *A social rule with $\mathscr{A}$ as its admissible set of profiles is* independent of irrelevant alternatives *if and only if* $(\forall f \in \mathscr{L})(\forall [\mathscr{D}_i], [\mathscr{D}'_i] \in \mathscr{A})$

$$(\forall i \in \mathscr{V},\ f \in \mathscr{D}_i \Leftrightarrow f \in \mathscr{D}'_i) \Rightarrow (f \in \Gamma([\mathscr{D}_i]) \Leftrightarrow f \in \Gamma([\mathscr{D}'_i])).$$

The interpretation of this property is similar to the one given in traditional social choice: whether a gamble f belongs to the aggregated coherent set depends only on which individuals are endorsing f.

If we consider again the social rules in Example 4, we observe that Γ_1, Γ_2 satisfy independence of irrelevant alternatives, because their definition depends only on which sets in the profile include the gamble; but Γ_3 and Γ_4 do not. To see this for Γ_3, consider a profile $[\mathscr{D}_i]$ where $\mathscr{D}_2$ is not maximal, and take f such that $f, -f \notin \mathscr{D}_2$. If for instance $f \in \Gamma_3([\mathscr{D}_i])$, then we can consider another profile $[\mathscr{D}'_i]$ where $\mathscr{D}'_1 = \mathscr{D}_1$ and $\mathscr{D}'_3 = \mathscr{D}_3$ and $\mathscr{D}'_2$ is a maximal set of gambles that includes $-f$. It follows that $-f \in \Gamma_3[\mathscr{D}'_i]$, and as a consequence $f \notin \Gamma_3[\mathscr{D}'_i]$, thus violating independence of irrelevant alternatives. For Γ_4 instead, consider a gamble $f \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$, fix $\epsilon > 0$ such that $f - \epsilon \notin (\mathscr{L}^- \cup \{0\})$ and let $\mathscr{D}_1 = \mathscr{D}_3 = \mathscr{E}(\{f\})$, $\mathscr{D}_2 = \mathscr{E}(\{f - \epsilon\})$. Then $f \in \Gamma_4([\mathscr{D}_i])$. If instead we take the profile $\mathscr{D}'_1 = \mathscr{D}'_2 = \mathscr{D}'_3 = \mathscr{E}(\{f\})$, then we get that $\Gamma_4([\mathscr{D}'_i]) = \mathscr{L}^+$, whence $f \notin \Gamma_4([\mathscr{D}'_i]$ even though in both cases f belongs to all the sets in the profile. Therefore, Γ_4 does not satisfy independence of irrelevant alternatives.

Then we consider *weak Pareto*.

Definition 27 (Weak Pareto). *A social rule* Γ *satisfies* weak Pareto *if and only if*

$$(\forall [\mathscr{D}_i] \in \mathscr{A})\ \cap_{i \in \mathscr{V}} \mathscr{D}_i \subseteq \Gamma([\mathscr{D}_i]).$$

Note that since the intersection of a family of coherent sets of gambles is again a coherent set of gambles, this definition is consistent. We can also see that all the social rules in Example 4, except Γ_4, satisfy weak Pareto. To see a profile on which the latter does not satisfy weak Pareto, consider the profile $[\mathscr{D}'_i]$ used to show that Γ_4 does not satisfy independence of irrelevant alternatives.

To define *dictatorship, oligarchy* and *democracy*, we need the following concepts.

Definition 28 (Almost decisive set of individuals). *Given a social rule* Γ*, a set of individuals* $\mathscr{G} \subseteq \mathscr{V}$ *is* almost decisive for a gamble f *when*

$$(\forall [\mathscr{D}_i] \in \mathscr{A})\ f \in \cap_{i \in \mathscr{G}} \mathscr{D}_i \text{ and } f \notin \cup_{i \notin \mathscr{G}} \mathscr{D}_i \Rightarrow f \in \Gamma([\mathscr{D}_i]).$$

It is called almost decisive *when it is almost decisive for every gamble* f*.*

Note that when $\mathscr{G} = \mathscr{V}$ almost decisiveness reduces to Γ satisfying weak Pareto. With respect to the social rules in the Example 4, for Γ_1 the first individual is almost decisive, while the second and the third are not; for Γ_2, the set $\{2,3\}$ is almost decisive, but $\{2\}$ or $\{3\}$ separately are not; and for Γ_3 and Γ_4, the second individual is almost decisive, while the first and the third are not.

A slightly stronger notion is the following:

Definition 29 (**Decisive set of individuals**). *Given a social rule* Γ*, a set of individuals* $\mathscr{G} \subseteq \mathscr{V}$ *is* decisive *for a gamble* f *when*

$$(\forall[\mathscr{D}_i] \in \mathscr{A})\ f \in \cap_{i\in\mathscr{G}}\mathscr{D}_i \Rightarrow f \in \Gamma([\mathscr{D}_i]).$$

It is called decisive *when it is decisive for every gamble* f.

Notice that this means that if $\mathscr{G}$ is decisive, then any other $\mathscr{G}' \supset \mathscr{G}$ is also decisive.

In our Example 4, the first individual is decisive for Γ_1 and the second is decisive for Γ_3, thus implying that also any other group of individuals containing the first and the second individual respectively is decisive. Additionally, for Γ_2, the group $\mathscr{G} := \{2,3\}$ is decisive. Notice that for Γ_4 the second individual is almost decisive without being decisive (to see it, consider the profile $[\mathscr{D}_i']$ used to show that Γ_4 does not satisfy independence of irrelevant alternatives).

As one extreme version of decisiveness, we have *dictatorship*, i.e., the case where the decisive group consists of only one individual.

Definition 30 (**Dictatorship**). *An individual* $i \in \mathscr{V}$ *is a* dictator *for a social rule* Γ *if and only if* $\{i\}$ *is decisive.*

Returning to our Example 4, the first individual is a dictator for Γ_1 as well as the second is a dictator for Γ_3.

Notice that, if i is dictator for a social rule Γ, then $\mathscr{D}_i \subseteq \Gamma([\mathscr{D}_j])$ for any profile $[\mathscr{D}_j]$. The two sets need not coincide: dictatorship means that those gambles that are considered desirable by individual i must also be considered desirable in the overall assessment, but the latter may include other gambles, meaning that $\Gamma([\mathscr{D}_j])$ may be a strict superset.

When more than one person form the decisive group, we can arrive instead to an *oligarchy*.

Definition 31 (**Oligarchy**). *Given a social rule* Γ*, a set of individuals* $\mathscr{G} \subseteq \mathscr{V}$ *is an* oligarchy *if and only if:*

O1. *$\mathscr{G}$ is decisive;*

O2. $(\forall f \in \mathscr{L}(\mathscr{X}))(\forall[\mathscr{D}_i] \in \mathscr{A})((\exists i \in \mathscr{G})\ f \in \mathscr{D}_i) \Rightarrow -f \notin \Gamma([\mathscr{D}_i])$.

In other words, those gambles that are deemed desirable by all individuals of the oligarchy should also be considered desirable by the group; and the choices of the group should not contradict those of any member of the oligarchy.

Under mild conditions, there can be at most one oligarchy:

Lemma 5. *For any social rule satisfying unlimited domain or unlimited maximal domain, there can be at most one oligarchy.*

The polar opposite to dictatorship, i.e., when the oligarchy consists of the whole society, it is called *democracy*.

Having concluded the phase of re-definition of the main concepts of social choice and, in particular, of the Arrowian framework, we can move to the main results of our works in this context [Casanova et al., 2020; Casanova, Miranda and Zaffalon, 2021].

First of all, we have established a version of Arrow's theorem in our context (for a version of Arrow's theorem using preference relations see for example Feldman and Serrano, 2006, Section 13) and verified its limiting results are still valid in our context.

Theorem 3 (**Arrow's theorem**)**.** *Assume that* $|\mathscr{X}| \geq 3$. *Any social rule that satisfies:*

- *completeness,*
- *unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*

makes one (unique) individual a dictator.

Regarding the social rules in Example 4, we can observe that if we restrict the set of admissible profiles of Γ_3 to the ones composed only by maximal sets of gambles, then the resulting rule satisfies also independence of irrelevant alternatives. Hence it satisfies all the hypothesis of Theorem 3, giving rise to the dictatorship of the second individual.

Next we have shown that, under the assumption of unlimited maximal domain, independence of irrelevant alternatives is also necessary for the existence of a dictator.

Proposition 2. *Assume that* $|\mathscr{X}| \geq 3$. *Any social rule that satisfies unlimited maximal domain and that makes one (unique) individual a dictator, must satisfy independence of irrelevant alternatives.*

To see that a similar result does not hold under the assumption of unlimited domain, note that in the rule Γ_3 of Example 4 the second individual is a dictator. However, the rule does not satisfy independence of irrelevant alternatives.

A traditional way to generalize Arrow's theorem obtaining an oligarchy is to eliminate the hypothesis of completeness from the group preference relation and, possibly, also from the individuals' profile [Weymark, 1984, Corollary 2]. In the same spirit, we obtained an oligarchy removing the hypothesis of completeness from the social rules of Theorem 3 and, possibly, requiring unlimited domain in place of unlimited maximal domain.

Theorem 4. *Assume that* $|\mathscr{X}| \geq 3$. *For any social rule* Γ *that satisfies:*

- *unlimited domain or unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*

there exists a unique oligarchy.

Regarding the social rules in Example 4, we notice that Γ_2 satisfies all the hypotheses of Theorem 4. Hence, it gives rise to an oligarchy. It is possible to observe that, in this example, the social rule Γ_2 just picks those gambles that are deemed desirable by all members of the oligarchy: $\Gamma_2([\mathscr{D}_i]) = \cap_{i \in \{2,3\}} \mathscr{D}_i$ (we may refer to these as *strong oligarchies*). We show that this is indeed a result valid for every social rule satisfying the assumptions of Theorem 4.

Proposition 3. *Let* Γ *be a social rule satisfying unlimited domain (resp., unlimited maximal domain), independence of irrelevant alternatives, and weak Pareto, and let* $\mathscr{G}$ *be its associated oligarchy that follows from Theorem 4. Then* $(\forall [\mathscr{D}_i] \in \mathscr{A})\ \Gamma([\mathscr{D}_i]) = \cap_{i \in \mathscr{G}} \mathscr{D}_i$.

Again in the same spirit of traditional social choice results, we show that by adding the property of *anonymity* to the axioms of Theorem 4, we obtain the polar opposite to Arrow's dictatorship, i.e., a democracy.

Definition 32 (Anonymity). *A social rule* Γ *satisfies* anonymity *if and only if for every permutation* σ *of* $\mathscr{V}$ *and for every profile* $[\mathscr{D}_i]$, *it holds that*

$$\Gamma([\mathscr{D}_i]) = \Gamma([\mathscr{D}_{\sigma(i)}]).$$

None of the rules of Example 4 satisfy this property. However, if we modify Γ_2 to take the intersection of all the coherent sets of the individuals, it will satisfy anonymity.

Theorem 5 (Democracy). *Assume that* $|\mathscr{X}| \geq 3$. *For any social rule* Γ *that satisfies:*

- *unlimited domain or unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*
- *anonymity,*

there exists a unique oligarchy, which is the whole society.

Thus, the social rule cannot contradict any of the choices of the individuals, and must incorporate those options where all of them agree. Notice moreover that, by Proposition 3 and Theorem 5, the only social rule satisfying these hypotheses is

$$(\forall[\mathscr{D}_i] \in \mathscr{A})\,\Gamma([\mathscr{D}_i]) = \cap_{i\in\mathscr{V}}\mathscr{D}_i.$$

This social rule is recognized in the literature and is referred to as the *unanimity rule* in Walley, 1982.

It is also possible to prove that, *provided the other axioms in Theorem 4 are satisfied*, democracy is equivalent to anonymity.

Proposition 4. *Assume that* $|\mathscr{X}| \geq 3$. *For any social rule* Γ *that satisfies:*

- *unlimited domain or unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*

if the whole society is an oligarchy, then Γ *satisfies anonymity.*

2.2 Coherent social rules

After detailing whether and how standard results in social choice transform with desirability, we study the problem of aggregating opinions modeled by coherent sets of gambles at a more fundamental level by focusing on coherence.

To this end, we start by representing profiles in the joint space $\mathscr{V} \times \mathscr{X}$ of individuals and their opinions: to do so we extend the generic gamble $f_i \in \mathscr{D}_i$ into a gamble f on $\mathscr{V} \times \mathscr{X}$ that is equal to f_i when we consider individual i and is equal to zero otherwise. In other words, f is a gamble conditional on considering individual i, see Definition 7. We then call $\mathscr{D}|i$ the representation of $\mathscr{D}_i$ in the space $\mathscr{V} \times \mathscr{X}$ obtained by taking the (conditional) extension of the gambles in $\mathscr{D}_i$. At this point we can take the union of $\mathscr{D}|i$ for all individuals i; this allows us to talk of all individuals and their opinions together, without having introduced any new information. The natural extension of this union then, represents all the 'logical' implications of the individuals' opinions, see Section 1.1.3.

We have shown in particular that this natural extension is always coherent and its $\mathscr{X}$-marginal, see Definition 6, is given by $\cap_{i\in\mathscr{V}}\mathscr{D}_i$.

Lemma 6. *Consider a set of individuals $\mathscr{V}$ and a profile $[\mathscr{D}_i]$. For each $i \in \mathscr{V}$, let us define $\mathscr{D}|i$ on the product space $\mathscr{V} \times \mathscr{X}$ by:*

$$\mathscr{D}|i := \{f \in \mathscr{L}(\mathscr{V} \times \mathscr{X}) : f = \mathbb{I}_i \otimes f_i, f_i \in \mathscr{D}_i\}, \tag{2.1}$$

where $\mathbb{I}_i \otimes f_i$ is the gamble given by

$$(\forall (j,z) \in \mathscr{V} \times \mathscr{X})\ \mathbb{I}_i \otimes f_i(j,z) := \begin{cases} f_i(z) & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\cap_{i\in\mathscr{V}}\mathscr{D}_i = \mathrm{Marg}_{\mathscr{X}}(\mathscr{E})$, where

$$\mathscr{E} := \mathrm{posi}(\cup_{i\in\mathscr{V}}(\mathscr{D}|i) \cup \mathscr{L}^+(\mathscr{V} \times \mathscr{X})) \tag{2.2}$$

is a coherent set.

This result implies that we can always merge the individuals' opinions into a compromise that is consistent with those opinions. In addition, the fact that the individuals' shared viewpoints, represented by the set $\cap_{i\in\mathscr{V}}\mathscr{D}_i$, coincide with the $\mathscr{X}$-marginal of $\mathscr{E}$, implies that at the very least the compromise will have to be implemented via weak Pareto.

Under this light, weak Pareto turns out to be the last frontier before incoherence. This suggests the following definition.

Definition 33 (Coherent social rule). *A* coherent social rule *is a social rule Γ such that for every admissible profile $[\mathscr{D}_i]$:*

$$\Gamma([\mathscr{D}_i]) \supseteq \mathrm{Marg}_{\mathscr{X}}(\mathscr{E}),$$

where $\mathscr{E}$ is defined as in Eq. (2.2).

Weak Pareto results not only to be necessary for keeping together the individuals' opinions coherently, but also sufficient.

Theorem 6. *Let Γ be a social rule. Consider a set of individuals $\mathscr{V}$ with profile $[\mathscr{D}_i]$, and let $\mathscr{D}|i$ be given by Eq. (2.1) for every $i \in \mathscr{V}$. Let us define also*

$$\mathscr{E} := \left\{ \sum_{i\in\mathscr{V}} \mathbb{I}_i \otimes f_i : (\forall i \in \mathscr{V})\ f_i \in \mathscr{D}_i \cup \{0\} \right\} \setminus \{0\}.$$

Then

1. *$\mathscr{E}$ is equal to the set in Eq. (2.2).*
2. *$\Gamma([\mathscr{D}_i]) \supseteq \mathrm{Marg}_{\mathscr{Z}}(\mathscr{E})$ is equivalent to the existence of a coherent set $\mathscr{E}' \supseteq \mathscr{E}$ such that $\Gamma([\mathscr{D}_i]) = \mathrm{Marg}_{\mathscr{Z}}(\mathscr{E}')$.*
3. *The smallest such set is*
$$\mathscr{E}' := \{\mathbb{I}_{\mathscr{V}} \otimes f_0 + \sum_{i \in \mathscr{V}} \mathbb{I}_i \otimes f_i : f_0 \in \Gamma([\mathscr{D}_i]) \cup \{0\}, (\forall i \in \mathscr{V}) f_i \in \mathscr{D}_i \cup \{0\}\} \setminus \{0\}, \tag{2.3}$$
where $\mathbb{I}_{\mathscr{V}} \otimes f_0(j,z) := f_0(z)$ for every $(j,z) \in \mathscr{V} \times \mathscr{Z}$.
4. *For all $i \in \mathscr{V}$, $\mathscr{E}'|i = \mathscr{D}|i$.*

Point 2 in the theorem, in fact, shows that weak Pareto is equivalent to the existence of a social rule that is consistent with the individuals' opinions. Weak pareto therefore is the 'guardian' of the coherence of a social rule with the individuals' opinions: all and only the coherent social rules are those that contain $\bigcap_{i \in \mathscr{V}} \mathscr{D}_i$.

Points 3–4 detail instead some characteristics of the least-committal set $\mathscr{E}'$, in the space $\mathscr{V} \times \mathscr{Z}$, associated with a coherent social rule: point 3 details its form; point 4 reassures us that $\mathscr{E}'$ yields the same conditionals from which we started.

Next we analyse in more detail a special case of strict complete coherent social rule that associates to a set of linear previsions another linear prevision.

To talk about social rules defined in terms of coherent lower previsions or sets of finitely additive probabilities we shall make use of the one-to-one correspondences existing between these models and coherent sets of strictly desirable gambles, see Section 1. Properties of these rules derive from the analogous ones established for coherent sets. For instance, we shall say that a social rule Γ defined on coherent lower previsions satisfies weak Pareto if and only if the social rule Γ' that we can determine on the associated sets of strictly desirable gambles by
$$\Gamma'([\mathscr{D}_i^+]) = \{f : \Gamma([\underline{P}_i])(f) > 0\} \cup \mathscr{L}^+, \text{ where } \underline{P}_i = \sigma(\mathscr{D}_i^+) \text{ for every } i, \tag{2.4}$$
satisfies weak Pareto. Similar considerations hold for the other properties, and for social rules defined on credal sets. For more details on how the different axioms and government systems are represented in these contexts, see Theorem 19 in Appendix B.

Corollary 2. *Let Γ be a social rule that assigns, to any profile $[P_i]$ of linear previsions, another linear prevision P. If Γ satisfies weak Pareto, then for any profile $[P_i]$ there exists a probability mass function π over $\mathscr{V}$ such that $P = \sum_{i \in \mathscr{V}} \pi(i) P_i$.*

This result shows that in the strict complete case coherent social rules turn out to be just mixtures of the original linear previsions. In other words, this places coherent social rules, in the precise case, very close to *linear opinion pooling* [Stewart and Quintana, 2018]; the main difference being that in linear pooling the experts' weights are traditionally fixed while in coherent social rules they can change with the profile. Both, however, are based on the idea of pooling opinions by taking convex combinations, i.e., by a linear approach to pooling.

In the more general incomplete case, a similar idea is at the basis of the *convex opinion pooling* [Stewart and Quintana, 2018, Section 4, Section 5], a special case of coherent social rule forming the group opinion as the convex hull of the individual probabilistic opinions. Being a coherent social rule, it is compatible with our framework and, in particular, it corresponds to the special case that expresses total ignorance a priori about the relative importance of the individuals. As such it can yield quite an imprecise group opinion. It is wort noting that taking the convex hull of the probability measures associated with a class of maximal sets of gambles is equivalent to taking the intersection of these sets [Walley, 1982, Section 7.1]. Therefore, one can view convex opinion pooling as a specialized instance of Walley [1982]'s unanimity rule, reaffirming our previous conclusion that it represents a *rational* approach to opinion aggregation.[2]

Convex opinion pooling also solves the problem of *probabilistic independence preservation*, i.e., it forms a group opinion that regard two events as independent if all the agents agree that they are independent [Stewart and Quintana, 2018]. As natural it may seems, the latter property is not respected by linear opinion pooling [Stewart and Quintana, 2018]. Moreover, it satisfies the so-called *strong setwise function property*, which can be seen as a probabilistic analogue of the independence of irrelevant alternatives of social choice. It requires that the probabilistic opinion of a group of individuals for an event A remains the same if the individual opinions for A do not change, regardless the opinions for other events. The assumption of independence of irrelevant alternatives given in the Arrowian context in fact, appears to be particularly suited to model voting problems but not so much to model problems involving aggregating (imprecise) beliefs: it neglects individuals' opinions other than their bare votes. It states in fact that if every voter's preference between an alternative x and an alternative y remains unchanged, then the group's preference between x and y will also remain unchanged, independently from any other preference among different alternatives

[2] We express our gratitude to a thesis reviewer for this keen observation.

[Arrow, 1951, Chapter 3].[3] In the context of desirability, this translates asking that whether a gamble f belongs to the aggregated decision set depends only on which individuals are endorsing f. Nevertheless, this is still associated with a voting scenario. Indeed, consider, for instance, $n = 2$ experts expressing their opinions by means of two coherent lower previsions $\underline{P}_1, \underline{P}_2$. The requirement of independence of irrelevant alternatives here established would imply that if a gamble f satisfies $\underline{P}_1(f) = -0.1$, $\underline{P}_2(f) = 100$ in one profile and $\underline{P}_1(f) = -100$, $\underline{P}_2(f) = 0.1$ in another, it should be included in both the aggregated coherent sets of strictly desirable gambles or in neither of them, regardless the strength of the opinions involved. The same criterion, however, may be a sensible assumption when we want to avoid considering the experts' opinions other than their acceptability assessments, so as to avoid, e.g., that stronger opinions are favoured.

Investigating whether some axioms and results from social choice are sensible in opinion pooling, is surely something related to our work. In the future therefore, it would be interesting to deepen this aspect. For the moment, a list of essential and desirable properties for the aggregation of imprecise beliefs was already presented by Walley [1982]. While we agree that his *coherence*, which requires that the group opinion is still represented by a coherent set of gambles, and unanimity (weak Pareto) should be normative prescriptions, we think other suggested properties may be too restrictive.

As an application of Corollary 2 in the more traditional case in which weights are constant with respect to the profile, we also showed that even without requiring independence of irrelevant alternatives, we can find a dictatorship. To this end, we consider the usual situation in applications: the one where it is possible to distinguish individuals' probabilities from utilities.

Definition 34 (State independence). *A probability measure P on $\Omega \times \mathcal{X}$ is* state independent *if and only if possibilities and prizes are stochastically independent. A strict complete social rule Γ* satisfies state independence *if and only if for each profile $[P_i]$, the linear prevision $\Gamma([P_i])$ corresponds to a probability measure satisfying state independence. It is said to have* state independent domain *if and only if*

[3] The expression 'Independence of irrelevant alternatives' nevertheless, adopts various interpretations in different contexts. When focusing solely on the winners of an election process, it is often used to indicate the famous *Sen's α* property [Sen, 1970, p.17]. This property asserts that if an element in subset S_1 of S_2 of alternatives is the best option in S_2, it remains the best in S_1. In other words, the removal of some suboptimal alternatives should not alter the status of the best choice. This property, however, should not be confused with the condition of the same name proposed by Arrow, 1951, which instead concerns the functional relationship between social preference and individual preferences [Sen, 1970, p.17].

for each profile $[P_i]$ in its domain, P_i corresponds to a state independent probability measure for every i.

Theorem 7. *Let Γ be a social rule defined on a set of profiles $\mathscr{A}$ composed by linear previsions $[P_i]$. If Γ satisfies:*

- *state independent domain,*
- *weak Pareto,*
- *strict completeness,*
- *state independence,*

and if then there is a probability mass function π on $\mathscr{V}$ such that for all $[P_i] \in \mathscr{A}$ the linear prevision $\Gamma([P_i])$ can be written as $\sum_{i \in \mathscr{V}} \pi(i)P_i$, then there exists $j \in \mathscr{V}$ such that $\pi(j) = 1$.

A similar result is already proven in Seidenfeld et al., 1989. Here, however, we deduce something different: that π (independent of the profile) is degenerate, meaning that it is always the same individual that determines the collective choices, differently from Seidenfeld et al., 1989's result. In other words, we obtain a dictatorship with precise degenerates.

Further discussion and future lines of research on this topic can be found in Section 5.

2.3 Summary

In this chapter, we formulate a general problem of opinions aggregation where both individual and group's opinions are modeled through coherent sets of gambles. Coherent sets of gambles are capable of representing individual and group viewpoints, which can take various forms and may encompass both belief and value considerations. Furthermore, they represent a very easy formalism to work with.

In this context, we initially explore whether *Arrowian*-style constraints can still be applied to this alternative problem formulation. To address this query, we first reinterpret the core concepts of *social choice*, the field where Arrow's constraints were originally formulated, using elements of desirability. Subsequently, we redefine its primary findings, revealing that a version of Arrow's impossibility theorem can indeed be established within this framework. When aggregating opinions that adhere to the adaptation of Arrow's theorem therefore, it results in a form of *dictatorship*. However, desirability's capability to model incomplete information offers a way to evade this outcome. Indeed, by relaxing Arrow's *completeness* axiom, we can attain an *oligarchy* or even a *democracy*, wherein not just one individual makes decisions for the group. Nevertheless, it's important to note that achieving this outcome while respecting the other Arrowian axioms requires us to consider desirable only those gambles that are deemed desirable by all members of the oligarchy or democracy. Consequently, the only aggregating function that aligns with this relaxed Arrowian framework and leads to a democracy is Walley's *unanimity rule*.

Next, we move to analyse the aggregation problem through the lens of *coherence*, the foundational concept that underlies desirability. In particular, we concentrate on identifying the minimal criteria necessary to arrive at a group opinion *consistent* with the individuals ones. Our investigation reveals that the adaptation of Arrow's *weak Pareto axiom* can serve this purpose.

Walley's unanimity rule also proves to be 'rational' within this alternative framework, reinforcing our earlier conclusion that it represents a 'rational' approach to opinion aggregation. Particularly, when individual opinions are conveyed through single probability distributions, it simplifies to *convex opinion pooling*—a widely recognized method for preserving crucial properties during the aggregation process. When also the group is required to specify a unique probability measure, it simplifies to a form of linear pooling with profile-dependent weights, thus justifying the use of a linear approach for pooling.

We conclude the chapter by unveiling conditions that may result in dictatorship in practical, commonly encountered scenarios.

The diagram below offers a graphical representation of the key findings discovered in this chapter.

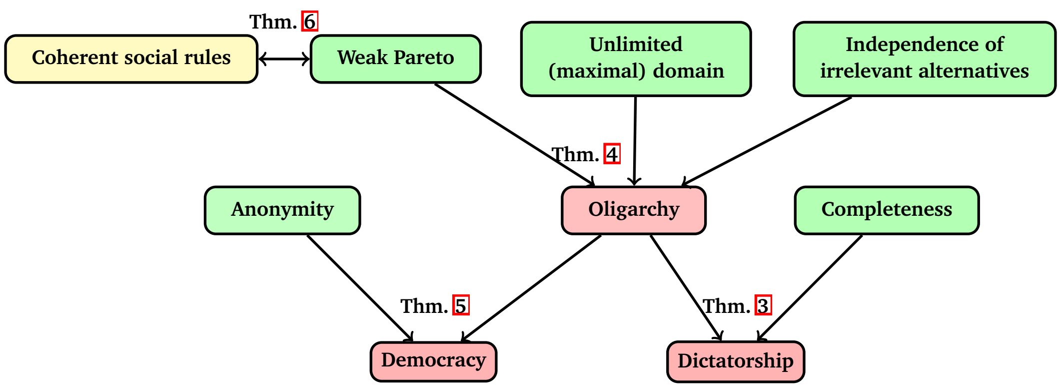

Figure 2.3. A graphical summary of the results presented.

Chapter 3

Information algebras and desirability

Information algebras, together with their *non-idempotent* variants called *valuation algebras*, are general algebraic structures to manage knowledge or information. They abstract away the most important features that appear in nearly every representation of information and, at such an abstract level, they provide operations and architectures for inference.

Information and valuation algebras subsume several formalisms in computer science both inspired by probability theory, like discrete probability potentials as used for example in Bayesian networks, and non-probabilistic systems, like relational databases, multiple systems of formal logic, and so on [Wilson and Mengin, 1999; Kohlas, 2003; Kohlas and Schmid, 2020; Kohlas, 2017].

The first idea of a generic structure to manage information can be found in Shenoy and Shafer [1990]. In the context of Bayesian networks indeed, Lauritzen and Spiegelhalter [1988] have shown that if a joint distribution factorizes, computation of marginals can be made feasible by arranging the operations in such a way that they take place 'locally' in the smaller domains of the factors. Such schemes of computations are called *local computations*. Shenoy and Shafer [1990] introduced for the first time an abstract, axiomatic system capturing the essence of local computation as introduced in Lauritzen and Spiegelhalter [1988].

After realizing that many of the formalisms in computer science are essentially instances of this system, Kohlas [2003] explicitly formulated it as an algebraic structure, denoted as *valuation algebra*, and studied it in detail as a general framework for inference.

The basic idea behind valuation algebras is the following. In many different formalisms to manage information, information essentially comes in pieces that refer to different *domains* or *questions* of interest. These blocks are somehow combined or aggregated to represent the whole of the information. Inference then usually means to extract from the whole of the knowledge the part relevant to a given question.

This leads to an algebraic structure composed by a set of 'pieces of information' $\mathbf{\Phi}$ and a set of questions $\boldsymbol{Q}$, manipulable by two basic operations: *combination*, to aggregate the pieces of information, and *extraction*, to extract from a piece of information the part related to a specific question. In Kohlas [2003], questions $\boldsymbol{Q}$ are assumed to regard *logically independent variables* with unknown values,[1] the so-called *multivariate model* for questions. In Kohlas [2017], this assumption is relaxed considering more general structures.

A valuation algebra satisfying the additional property of *Idempotency*, which requires that a piece of information combined with a part of it still gives the same information, is known as an *information algebra*. Idempotency is a very intuitive property of information. It adds structure to a valuation algebra allowing for simpler architectures for local computations and connections with information systems [Kohlas, 2003]. Moreover, it allows to introduce a partial order between pieces of information representing their relative informative content.

Since desirability induces instances of valuation algebras that are in particular information algebras, in what follows we concentrate only on these special structures.

The view of information as pieces regarding different questions of interest leads to two equivalent formulations of information algebras: the *labeled* and the *domain-free* one. The main difference is that in labeled information algebras, pieces of information are explicitly linked to questions they refer to while in domain-free ones they are treated as abstract entities, unrelated to particular questions. The labeled form is in general more convenient for computational considerations while the domain-free one is more suitable for theoretical issues Kohlas [2003].

Information algebras can also be placed in the realm of logic. Indeed, they further arise from particular *information systems* $(\boldsymbol{L}, \boldsymbol{Cl}, \mathfrak{S})$, see Section 1.1.3 and [Kohlas, 2003, Section 6.4]. Vice versa, it is possible to show that every information algebra induces this kind of system. Thus, information systems can be considered as an alternative way to represent information algebras [Kohlas, 2003, Section 6.4].

These considerations offer a very natural link with desirability, see Section 1, which is precisely the starting point of our works. Specifically, in Casanova et al.

[1]We say that variables in a set $\{X_i\}_{i \in I}$ where I is a non-empty set are *logically independent* if any value in the Cartesian product of their sets of possible values is assumed to be possible. If needed, we assume the axiom of choice.

[2022a], we show that, assuming a multivariate model for questions, coherent sets of gambles as well as coherent lower previsions induce instances of information algebras. In Casanova, Kohlas and Zaffalon [2021]; Kohlas et al. [2021]; Casanova et al. [2022b], analogous results are obtained considering also questions of more general forms.

Information algebras permit to encapsulate desirability in a more general structure to manage information. This allows us to further generalise results obtained with desirability, such as the ones related to the treatment of the *marginal problem*, see Section 3.1.3 and Section 3.2.2. In this context, desirability provides a general formulation of the problem comprehensive of most of the frameworks within which it is generally studied. Information algebras allows us to generalise it even more. Desirability moreover, forms a very general instance of information algebras able to manipulate very different forms of information, see Section 1. In the future, we plan to further explore the capability of desirability of modeling frameworks handled by information algebras. More details about future lines of research could be found in Section 5.

In the following Section 3.1 and Section 3.2, we systematically present our research findings in this domain. Our approach is structured as follows: each section begins with an in-depth review of the relevant literature, followed by a comprehensive exposition of our research results.

In Section 3.1, our focus is on deriving instances of information algebras from desirability, considering a multivariate model for questions. We commence this section by revisiting the definition of information algebras as provided in the literature, as outlined in Kohlas [2003], both in the domain-free and labeled versions. Subsequently, Sections 3.1.1, 3.1.2, and 3.1.3 offer a detailed presentation of the results obtained in our work Casanova et al., 2022a.

In contrast, Section 3.2 explores more comprehensive models for questions. Here, we introduce our definition of a 'generalized' information algebra, building upon our prior works Casanova, Kohlas and Zaffalon, 2021; Kohlas et al., 2021; Casanova et al., 2022b and aligning it with the established literature definition in Kohlas, 2017. Following this, in Section 3.2.1 and Section 3.2.2, we provide a summary of the outcomes and findings regarding the link between 'generalized' information algebras and desirability, as detailed in our previous works Casanova, Kohlas and Zaffalon, 2021; Kohlas et al., 2021; Casanova et al., 2022b.

The proofs of all the results presented in Section 3.1 can be found in Appendix C.1,the proofs of the ones presented in Section 3.2 can be found instead in Appendix C.2.

3.1 Information algebras: multivariate model

Here we recall the basic definitions given in Kohlas [2003] and related works.

As previously said, in this section we consider reasoning and inference to be concerned with a set of logically independent variables with unknown values. So, we consider a set $\mathbf{\Phi}$ of pieces of information representing knowledge about the possible values of a set of logically independent variables $\{X_i\}_{i\in I}$, where I is a non-empty index set. Pieces of information express knowledge on the whole or on a subset of the variables involved. Questions of interest are thus modeled as subsets of $\{X_i\}_{i\in I}$ or, alternatively, as subsets of I. This model for questions is called *multivariate model*, see [Kohlas, 2017, Section 2.2].

We are now ready to give the definition of a *domain-free* and a *labeled* information algebra in this context [Kohlas, 2003, Section 2, Section 3, Section 6;Kohlas and Schmid, 2014, Section 2, Section 3].

Definition 35 (**Domain-free information algebra**)**.** *A* domain-free information algebra *is a two-sorted algebra* $(\mathbf{\Phi}, \mathrm{P}(I); \cap, \cup, \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$*, where:*

- $(\mathbf{\Phi}; \cdot, \mathbf{0}, \mathbf{1})$ *is a commutative semigroup with* $\cdot : \mathbf{\Phi} \times \mathbf{\Phi} \to \mathbf{\Phi}$ *defined by* $\boldsymbol{\phi}, \boldsymbol{\psi} \mapsto \boldsymbol{\phi} \cdot \boldsymbol{\psi}$*, and with* $\mathbf{0}$ *and* $\mathbf{1}$ *as its null and unit elements respectively,*

- $(\mathrm{P}(I); \cap, \cup)$ *is the lattice constructed from the power-set of a non-empty index set of variables* I *ordered by inclusion,*

- $\boldsymbol{\epsilon} : \mathbf{\Phi} \times \mathrm{P}(I) \to \mathbf{\Phi}$*, defined by* $\boldsymbol{\phi}, S \mapsto \boldsymbol{\epsilon}_S(\boldsymbol{\phi})$*,*

satisfying moreover the following properties:

- Nullity*: for any* $S \subseteq I$*,*
$$\boldsymbol{\epsilon}_S(\mathbf{0}) = \mathbf{0};$$

- Idempotency*: for any* $\boldsymbol{\phi} \in \mathbf{\Phi}$ *and* $S \subseteq I$*,*
$$\boldsymbol{\epsilon}_S(\boldsymbol{\phi}) \cdot \boldsymbol{\phi} = \boldsymbol{\phi};$$

- Combination*: for any* $\boldsymbol{\phi}, \boldsymbol{\psi} \in \mathbf{\Phi}$ *and* $S \subseteq I$*,*
$$\boldsymbol{\epsilon}_S(\boldsymbol{\epsilon}_S(\boldsymbol{\phi}) \cdot \boldsymbol{\psi}) = \boldsymbol{\epsilon}_S(\boldsymbol{\phi}) \cdot \boldsymbol{\epsilon}_S(\boldsymbol{\psi});$$

- Transitivity*: for any* $\boldsymbol{\phi} \in \mathbf{\Phi}$ *and* $S, T \subseteq I$*,*
$$\boldsymbol{\epsilon}_S(\boldsymbol{\epsilon}_T(\boldsymbol{\phi})) = \boldsymbol{\epsilon}_T(\boldsymbol{\epsilon}_S(\boldsymbol{\phi})) = \boldsymbol{\epsilon}_{S\cap T}(\boldsymbol{\phi});$$

- Support*: for any* $\boldsymbol{\phi} \in \mathbf{\Phi}$*,*
$$\boldsymbol{\epsilon}_I(\boldsymbol{\phi}) = \boldsymbol{\phi}.$$

Since from an index set I we can always construct the lattice $(\mathrm{P}(I);\cap,\cup)$*, we henceforth write* $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$ *as a short notation for* $(\boldsymbol{\Phi}, \mathrm{P}(I); \cap, \cup, \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$.

Notice in particular that, by Idempotency, we also have $\boldsymbol{\epsilon}_S(\mathbf{1}) = \boldsymbol{\epsilon}_S(\mathbf{1}) \cdot \mathbf{1} = \mathbf{1}$.

The first binary operation $\cdot$ defined on a domain-free information algebra $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$ is called *combination*. For any pair of pieces of information $\boldsymbol{\phi}, \boldsymbol{\psi} \in \boldsymbol{\Phi}$, $\boldsymbol{\phi} \cdot \boldsymbol{\psi}$ represents the information obtained aggregating $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$. $(\boldsymbol{\Phi}; \cdot, \mathbf{0}, \mathbf{1})$ is then required to be a commutative semigroup in order to mimic the intuitive properties of the ‘aggregation’ of information. The null element represents contradiction, hence combined with any piece of information it generates again contradiction, and the unit element represents vacuous information that combined with any piece of information does not add anything new.

The second binary operation $\boldsymbol{\epsilon}$ is called *extraction*. Given a piece of information $\boldsymbol{\phi}$ and a question $S \subseteq I$, $\boldsymbol{\epsilon}_S(\boldsymbol{\phi})$ represents the information regarding the question S extracted from the piece of information $\boldsymbol{\phi}$. If $\boldsymbol{\epsilon}_S(\boldsymbol{\phi}) = \boldsymbol{\phi}$, S is called a *support* of $\boldsymbol{\phi}$. In this case $\boldsymbol{\phi}$ is just an information on S.

The intuition behind the other axioms is the following. Nullity says that extraction from contradiction still gives contradiction. Idempotency says that combining a piece of information with part of it gives nothing new. The property of Combination is for a domain-free information algebra the most important requirement: if we combine the part of a piece of information relating to the question $S \subseteq I$ with any other piece of information and extract the part relating to S from the aggregated information, then we may as well *first* extract the part regrading S from the second piece of information and then combine. This is most important for local computation. Transitivity assures instead that the order of successive extractions does not matter. This will no longer be true when questions of interest are modeled with more general structures, see Section 3.2. Finally, Support is useful to construct equivalent labeled versions of domain-free information algebras, as shown later on.

Definition 36 (**Labeled information algebra**)**.** *A* labeled information algebra *is a two-sorted (partial) algebra* $(\tilde{\boldsymbol{\Phi}}, \mathrm{P}(I); \cap, \cup, \boldsymbol{d}, \cdot, \{\tilde{\mathbf{0}}_S\}_{S \in \mathrm{P}(I)}, \{\tilde{\mathbf{1}}_S\}_{S \in \mathrm{P}(I)}, \boldsymbol{\pi})$*, where:*

- $\boldsymbol{d} : \tilde{\boldsymbol{\Phi}} \to \mathrm{P}(I)$*,*
- $(\tilde{\boldsymbol{\Phi}}; \cdot)$ *is a commutative semigroup with* $\cdot : \tilde{\boldsymbol{\Phi}} \times \tilde{\boldsymbol{\Phi}} \to \tilde{\boldsymbol{\Phi}}$ *defined by* $\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\psi}} \mapsto \tilde{\boldsymbol{\phi}} \cdot \tilde{\boldsymbol{\psi}}$. *For all* $S \subseteq I$*, there exist an element* $\tilde{\mathbf{0}}_S$ *and an element* $\tilde{\mathbf{1}}_S$ *with* $\boldsymbol{d}(\tilde{\mathbf{0}}_S) = S$ *and* $\boldsymbol{d}(\tilde{\mathbf{1}}_S) = S$ *such that for all* $\tilde{\boldsymbol{\phi}} \in \tilde{\boldsymbol{\Phi}}$ *with* $\boldsymbol{d}(\tilde{\boldsymbol{\phi}}) = S$*,* $\tilde{\mathbf{0}}_S \cdot \tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\phi}} \cdot \tilde{\mathbf{0}}_S = \tilde{\mathbf{0}}_S$ *and* $\tilde{\mathbf{1}}_S \cdot \tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\phi}} \cdot \tilde{\mathbf{1}}_S = \tilde{\boldsymbol{\phi}}$*,*
- $(\mathrm{P}(I);\cap,\cup)$ *is the lattice constructed from the power-set of a non-empty index set of variables* I *ordered by inclusion,*

- $\pi : dom(\pi) \subseteq \tilde{\mathbf{\Phi}} \times \mathrm{P}(I) \to \tilde{\mathbf{\Phi}}$, *defined by* $\tilde{\boldsymbol{\phi}}, S \mapsto \pi_S(\tilde{\boldsymbol{\phi}})$ *for* $S \subseteq \boldsymbol{d}(\tilde{\boldsymbol{\phi}})$*,*

satisfying moreover the following properties:

1. Labeling*: for* $\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\psi}} \in \tilde{\mathbf{\Phi}}$*,*
$$\boldsymbol{d}(\tilde{\boldsymbol{\phi}} \cdot \tilde{\boldsymbol{\psi}}) = \boldsymbol{d}(\tilde{\boldsymbol{\phi}}) \cup \boldsymbol{d}(\tilde{\boldsymbol{\psi}});$$
2. Marginalisation*: for* $\tilde{\boldsymbol{\phi}} \in \tilde{\mathbf{\Phi}}$ *and* $S \subseteq \boldsymbol{d}(\tilde{\boldsymbol{\phi}})$*,*
$$\boldsymbol{d}(\pi_S(\tilde{\boldsymbol{\phi}})) = S;$$
3. Nullity *and* Neutrality*: for* $S, T \subseteq I$*,*
$$\tilde{\mathbf{0}}_S \cdot \tilde{\mathbf{0}}_T = \tilde{\mathbf{0}}_{S \cup T};$$
$$\tilde{\mathbf{1}}_S \cdot \tilde{\mathbf{1}}_T = \tilde{\mathbf{1}}_{S \cup T};$$
4. Stability*: for* $S, T \subseteq I$ *with* $T \subseteq S$*,*
$$\pi_T(\tilde{\mathbf{1}}_S) = \tilde{\mathbf{1}}_T.$$
5. Idempotency*: for* $\tilde{\boldsymbol{\phi}} \in \tilde{\mathbf{\Phi}}$ *and* $S \subseteq \boldsymbol{d}(\tilde{\boldsymbol{\phi}})$*,*
$$\pi_S(\tilde{\boldsymbol{\phi}}) \cdot \tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{\phi}};$$
6. Combination*: for* $\tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{\psi}} \in \tilde{\mathbf{\Phi}}$ *with* $\boldsymbol{d}(\tilde{\boldsymbol{\phi}}) = S$ *and* $\boldsymbol{d}(\tilde{\boldsymbol{\psi}}) = T$*,*
$$\pi_S(\tilde{\boldsymbol{\phi}} \cdot \tilde{\boldsymbol{\psi}}) = \tilde{\boldsymbol{\phi}} \cdot \pi_{S \cap T}(\tilde{\boldsymbol{\psi}});$$
7. Transitivity*: for* $\tilde{\boldsymbol{\phi}} \in \tilde{\mathbf{\Phi}}$ *and* $T \subseteq S \subseteq \boldsymbol{d}(\tilde{\boldsymbol{\phi}})$*,*
$$\pi_T(\pi_S(\tilde{\boldsymbol{\phi}})) = \pi_T(\tilde{\boldsymbol{\phi}});$$

Since from an index set I *we can always construct the lattice* $(\mathrm{P}(I); \cap, \cup)$*, we henceforth write* $(\tilde{\mathbf{\Phi}}, I; \boldsymbol{d}, \cdot, \{\tilde{\mathbf{0}}_S\}_{S \subseteq I}, \{\tilde{\mathbf{1}}_S\}_{S \subseteq I}, \pi)$ *as a short notation for* $(\tilde{\mathbf{\Phi}}, \mathrm{P}(I); \cap, \cup, \boldsymbol{d}, \cdot,$ $\{\tilde{\mathbf{0}}_S\}_{S \in \mathrm{P}(I)}, \{\tilde{\mathbf{1}}_S\}_{S \in \mathrm{P}(I)}, \pi)$*.*

Unlike before, here we explicitly link any piece of information to the subset of variables to which it refers. This is highlighted by a new operation $\boldsymbol{d}$ called *labeling* that associates to each piece of information $\tilde{\boldsymbol{\phi}}$ its *domain* $\boldsymbol{d}(\tilde{\boldsymbol{\phi}})$. This link is what, in general, makes labeled information algebras more suitable for computational purposes: it allows to limit the memory requirement only to what is needed.

The other two operations $\cdot$ and π, called *combination* and *marginalisation* or *projection* respectively, are instead the counterpart of combination and extraction operations defined on a domain-free information algebra. Here however, from any piece of information $\tilde{\boldsymbol{\phi}}$ it is possible to extract information only regarding domains smaller than (or equal to) $\boldsymbol{d}(\tilde{\boldsymbol{\phi}})$. Axioms required to a labeled information algebra are also very similar to the ones required to a domain-free one.

Classical notions of *homomorphism* and *subalgebra* can be examined with respect to information algebras [Kohlas, 2003, Section 3.3]. We recall here their definitions only for domain-free information algebras having the same set of questions. Analogous definitions can be established for the labeled case.

Definition 37 (**Homomorphism - Domain-free version**). *Let* $(\boldsymbol{\Phi}^1, I; \cdot^1, \mathbf{0}^1, \mathbf{1}^1, \epsilon^1)$ *and* $(\boldsymbol{\Phi}^2, I; \cdot^2, \mathbf{0}^2, \mathbf{1}^2, \epsilon^2)$ *be two domain-free information algebras. A mapping* $h : \boldsymbol{\Phi}^1 \to \boldsymbol{\Phi}^2$ *is called a* homomorphism *if, for any* $\boldsymbol{\phi}^1, \boldsymbol{\psi}^1 \in \boldsymbol{\Phi}^1$ *and* $S \subseteq I$*, we have:*

1. $h(\boldsymbol{\phi}^1 \cdot^1 \boldsymbol{\psi}^1) = h(\boldsymbol{\phi}^1) \cdot^2 h(\boldsymbol{\psi}^1)$,

2. $h(\mathbf{0}^1) = \mathbf{0}^2$ *and* $h(\mathbf{1}^1) = \mathbf{1}^2$,

3. $h(\epsilon_S^1(\boldsymbol{\phi}^1)) = \epsilon_S^2(h(\boldsymbol{\phi}^1))$.

If there is a homomorphism h from $(\boldsymbol{\Phi}^1, I; \cdot^1, \mathbf{0}^1, \mathbf{1}^1, \epsilon^1)$ to $(\boldsymbol{\Phi}^2, I; \cdot^2, \mathbf{0}^2, \mathbf{1}^2, \epsilon^2)$, the first information algebra is said to be *homomorphic* to the second one. If h is injective it is called *embedding* and $(\boldsymbol{\Phi}^1, I; \cdot^1, \mathbf{0}^1, \mathbf{1}^1, \epsilon^1)$ is said to be *embedded* into $(\boldsymbol{\Phi}^2, I; \cdot^2, \mathbf{0}^2, \mathbf{1}^2, \epsilon^2)$. If h is also bijective, it is an *isomorphism* and $(\boldsymbol{\Phi}^1, I; \cdot^1, \mathbf{0}^1, \mathbf{1}^1, \epsilon^1)$ is *isomorphic* to $(\boldsymbol{\Phi}^2, I; \cdot^2, \mathbf{0}^2, \mathbf{1}^2, \epsilon^2)$.

Definition 38 (**Subalgebra - Domain-free version**). *Let* $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$ *be a domain-free information algebra and consider* $\boldsymbol{\Phi}' \subseteq \boldsymbol{\Phi}$*. If:*

1. $(\forall \boldsymbol{\phi}', \boldsymbol{\psi}' \in \boldsymbol{\Phi}')\ \boldsymbol{\phi}' \cdot \boldsymbol{\psi}' \in \boldsymbol{\Phi}'$,

2. $\mathbf{0}, \mathbf{1} \in \boldsymbol{\Phi}'$,

3. $(\forall \boldsymbol{\phi}' \in \boldsymbol{\Phi}', S \subseteq I)\ \epsilon_S(\boldsymbol{\phi}') \in \boldsymbol{\Phi}'$,

$(\boldsymbol{\Phi}', I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$ *qualifies as a domain-free information algebra and is precisely denoted as a* subalgebra *of* $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$.

It is possible to show moreover that labeled and domain-free versions of information algebras are equivalent [Kohlas, 2003, Section 3.2]. Let us consider a domain-free information algebra $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$. Consider then the set of pairs

$$\tilde{\boldsymbol{\Phi}} := \{(\boldsymbol{\phi}, S), \boldsymbol{\phi} \in \boldsymbol{\Phi}, S \subseteq I : \epsilon_S(\boldsymbol{\phi}) = \boldsymbol{\phi}\}. \tag{3.1}$$

These pairs can be considered as pieces of information $\boldsymbol{\phi}$, labeled by their domains. Now, let us define on $\tilde{\boldsymbol{\Phi}}$ and $\mathsf{P}(I)$ the following operations expressed in terms of the ones defined on $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$.

- Labeling: $\boldsymbol{d} : \tilde{\boldsymbol{\Phi}} \to \mathsf{P}(I)$, defined by $(\boldsymbol{\phi}, S) \mapsto \boldsymbol{d}(\boldsymbol{\phi}, S) := S$.

- Combination: $\cdot : \tilde{\boldsymbol{\Phi}} \times \tilde{\boldsymbol{\Phi}} \to \tilde{\boldsymbol{\Phi}}$, defined by
$$(\boldsymbol{\phi}, S), (\boldsymbol{\psi}, T) \mapsto (\boldsymbol{\phi}, S) \cdot (\boldsymbol{\psi}, T) := (\boldsymbol{\phi} \cdot \boldsymbol{\psi}, S \cup T).$$
- Marginalisation: $\pi : dom(\pi) \subseteq \tilde{\boldsymbol{\Phi}} \times \mathrm{P}(I) \to \tilde{\boldsymbol{\Phi}}$, defined by
$$(\boldsymbol{\phi}, S), T \mapsto \pi_T(\boldsymbol{\phi}, S) := (\boldsymbol{\epsilon}_T(\boldsymbol{\phi}), T),$$
for every $T \subseteq S \subseteq I$.

It can be shown that $(\tilde{\boldsymbol{\Phi}}, I; \boldsymbol{d}, \cdot, \{(\mathbf{0}, S)\}_{S \subseteq I}, \{(\mathbf{1}, S)\}_{S \subseteq I}, \boldsymbol{\pi})$, where $\boldsymbol{d}, \cdot$ and $\boldsymbol{\pi}$ are the operations just defined on $\tilde{\boldsymbol{\Phi}}$ and $\mathrm{P}(I)$, is a labeled information algebra [Kohlas, 2003, p. 54]. From this labeled information algebra it is then possible to reconstruct a domain-free information algebra that is isomorphic to the original one. Similarly, it is possible to start with a labeled algebra, construct the associated domain-free information algebra and then reconstruct, by the procedure just introduced, a labeled information algebra that is again essentially (up to isomorphism) the same as the original one [Kohlas, 2003, p. 54].

As introduced before, axioms defining a domain-free and a labeled version of an information algebra lead to the definition of a partial order on pieces of information that is called *information order* [Kohlas, 2003, Section 6.2].

Definition 39 (Information order - Domain-free version). *Consider a domain -free information algebra $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$. Given $\boldsymbol{\phi}, \boldsymbol{\psi} \in \boldsymbol{\Phi}$ we say that $\boldsymbol{\phi} \leq \boldsymbol{\psi}$ if and only if $\boldsymbol{\phi} \cdot \boldsymbol{\psi} = \boldsymbol{\psi}$.*

This means that $\boldsymbol{\psi}$ is more informative than $\boldsymbol{\phi}$, if adding $\boldsymbol{\phi}$ to $\boldsymbol{\psi}$ gives nothing new. Notice that, in particular, $\mathbf{1} \leq \boldsymbol{\phi} \leq \mathbf{0}$ for every $\boldsymbol{\phi} \in \boldsymbol{\Phi}$.

The same definition can be given for the labeled case.

In certain information algebras there are maximally informative elements called *atoms* [Kohlas, 2003, Definition 6.13].

Definition 40 (Atoms - Domain-free version). *Given a domain-free information algebra $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$, an element $\boldsymbol{\alpha} \in \boldsymbol{\Phi}$ is called an* atom *if and only if*

- $\boldsymbol{\alpha} \neq \mathbf{0}$,
- $(\forall \boldsymbol{\phi} \in \boldsymbol{\Phi})\ \boldsymbol{\alpha} \leq \boldsymbol{\phi} \Rightarrow \boldsymbol{\phi} = \boldsymbol{\alpha}$ *or* $\boldsymbol{\phi} = \mathbf{0}$.

This says that no information, except the null information, can be more informative than an atom. A similar definition can be given also in the labeled case, where atoms are maximally informative elements with respect to their domains.

Definition 41 (Atoms - Labeled version). *Given a labeled information algebra $(\tilde{\boldsymbol{\Phi}}, I; \boldsymbol{d}, \cdot, \{\tilde{\mathbf{0}}_S\}_{S \subseteq I}, \{\tilde{\mathbf{1}}_S\}_{S \subseteq I}, \pi)$, an element $\tilde{\boldsymbol{\alpha}} \in \tilde{\boldsymbol{\Phi}}$ with $\boldsymbol{d}(\tilde{\boldsymbol{\alpha}}) = S$ is called an* atom relative to *S, if and only if*

- $\tilde{\alpha} \neq \tilde{\mathbf{0}}_S$,
- $(\forall \tilde{\phi} \in \tilde{\Phi}$ *with* $d(\tilde{\phi}) = S)\ \tilde{\alpha} \leq \tilde{\phi} \Rightarrow \tilde{\phi} = \tilde{\alpha}$ *or* $\tilde{\phi} = \tilde{\mathbf{0}}_S$.

Let us denote with $At(\Phi)$ the set of all atoms of a domain-free information algebra $(\Phi, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$. Let us define also, for every element $\phi \in \Phi$,

$$At(\phi) := \{\alpha \in At(\Phi) : \phi \leq \alpha\}. \tag{3.2}$$

We can identify the following types of domain-free information algebras [Kohlas, 2003, Definition 6.15].

Definition 42 (**Atomic information algebras - Domain free version**). *A domain-free information algebra* $(\Phi, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$ *is called:*

1. atomic, *if and only if for all* $\phi \in \Phi$, $\phi \neq \mathbf{0}$, $At(\phi)$ *is not empty,*
2. atomistic *or* atomic composed, *if and only if it is atomic and for all* $\phi \in \Phi$, $\phi \neq \mathbf{0}$,
$$\phi = \bigwedge At(\phi).$$

Analogous definitions are valid for the labeled case.

3.1.1 Information algebras of coherent sets of gambles

Domain-free information algebras also arise from particular *information systems* $(L, Cl, \mathfrak{S})$,[2] where Cl and $\mathfrak{S}$ respect some properties [Kohlas, 2003, Chapter 6.4]. In this case, pieces of information correspond to the closed subsets of L:

$$\Phi_{Cl} := \{E \subseteq L : E = Cl(E)\}$$

and questions are modeled through the sublanguages $\mathfrak{S}$,[3]. Operations of combination and extraction can then be simply defined as:

1. *Combination*: $E_1 \cdot E_2 := Cl(E_1 \cup E_2)$, $E_1, E_2 \in \Phi_{Cl}$, i.e., the combination of two pieces of information E_1, E_2 is the least closed set containing both of them;
2. *Extraction*: $\epsilon_S(E) := Cl(E \cap S)$, $E \in \Phi_{Cl}, S \in \mathfrak{S}$, i.e., the information contained in a piece E regarding the question modeled by the sublanguage S is the least closed set that contains E focused on the elements of S.

[2]See Section 1.1.3.

[3]Sublanguages here considered are assumed to form lattices (with respect to set inclusion) isomorphic to lattices of subsets of variables. For this reason, in what follows, we can indifferently represent questions both as sublanguages and as sets of variables.

Starting from a similar observation for desirability, we have induced a *domain-free information algebra of coherent sets of gambles*.

As observed in Section 1.1.3, indeed, in this context we can consider $\mathscr{L}$ as a language and $\mathscr{C}$ as a closure operator. Moreover, we can consider $\Phi(\Omega) = \mathbb{D}(\Omega) \cup \{\mathscr{L}(\Omega)\}$ as a set of pieces of information, since it corresponds to the set of all and only the closed subsets of $\mathscr{L}(\Omega)$ with respect to the closure operator $\mathscr{C}$.

Modeling questions is instead a bit more involved. As before, we assume questions to regard a set of logically independent variables $\{X_i\}_{i\in I}$. For this reason, we assume also a special form for the possibility space Ω, that of a product space: $\Omega := \times_{i\in I} \Omega_i$, where Ω_i is the set of possible values of the variable X_i for every i. Elements ω in Ω can therefore be seen as functions $\omega : I \to \Omega$, so that $\omega_i := \omega(i) \in \Omega_i$, for any $i \in I$. For any subset S of I, we also let $\Omega_S := \times_{i\in S} \Omega_i$.

In this way, sets $\mathscr{D} \in \mathbb{D}(\Omega)$ represent an information about the whole or a subset of the index set of variables $\{X_i\}_{i\in I}$. Questions can then be identified as subsets $S \subseteq I$ or, alternatively, as sublanguages formed by sets of gambles that only depend on variables $\{X_i\}_{i\in S}$. These correspond to the sets $\mathscr{L}_{\mathscr{P}_S}(\Omega)$ of $\mathscr{P}_S$-measurable gambles (see Definition 5) where, given $\omega, \omega' \in \Omega$, ω, ω' are in the same block of $\mathscr{P}_S$ if and only if $\omega|S = \omega'|S$ (here $\omega|S$ is the restriction of the map ω to S). In what follows we refer to $\mathscr{P}_S$-measurable gambles also as S-measurable gambles for simplicity. For the same reason, for every $S \subseteq I$, we indicate also $\mathscr{L}_{\mathscr{P}_S}(\Omega)$ as $\mathscr{L}_S(\Omega)$, or $\mathscr{L}_S$. In particular, if $I = \emptyset$, then $\mathscr{L}(\Omega_I) = \mathbb{R}$, the set of constant gambles [de Cooman et al., 2011, Section 2.3], moreover $\mathscr{L}_\emptyset := \mathbb{R}$, and $\mathscr{L}_I = \mathscr{L}(\Omega)$.

Following the previous reasoning, we can define the following operations:

1. Combination. $\cdot : \Phi \times \Phi \to \Phi$, defined by
$$\mathscr{D}_1, \mathscr{D}_2 \mapsto \mathscr{D}_1 \cdot \mathscr{D}_2 := \mathscr{D}_1 \vee \mathscr{D}_2 := \mathscr{C}(\mathscr{D}_1 \cup \mathscr{D}_2);$$
2. Extraction. $\epsilon : \Phi \times \mathrm{P}(I) \to \Phi$, defined by
$$\mathscr{D}, S \mapsto \epsilon_S(\mathscr{D}) := \mathscr{C}(\mathscr{D} \cap \mathscr{L}_S).$$

Combination corresponds to the join operation on Φ. Extraction instead essentially corresponds to the marginalisation operation defined on coherent sets of gambles. Given a coherent sets of gambles $\mathscr{D}$ indeed, $\mathscr{C}(\mathscr{D} \cap \mathscr{L}_S)$ contains the same information of the $\mathscr{P}_S$—marginal set of $\mathscr{D}$, $\mathscr{D} \cap \mathscr{L}_S$, see Definition 6. This consideration will be particularly important for the construction of a labeled version of the information algebra of coherent sets of gambles.

It is then possible to formally demonstrate that Φ and $\mathrm{P}(I)$, enriched with the operations defined above, induce a domain-free information algebra $(\Phi, I;$

$\cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$, called *domain-free information algebra of coherent sets of gambles*.

Theorem 8. 1. $(\Phi; \cdot, 0, 1)$ *is a commutative semigroup with a null element* $0 = \mathscr{L}$ *and a unit element* $1 = \mathscr{L}^+$.

2. *For any* $S \subseteq I$, $\epsilon_S(0) = 0$.

3. *For any* $\mathscr{D} \in \Phi$ *and* $S \subseteq I$, $\epsilon_S(\mathscr{D}) \cdot \mathscr{D} = \mathscr{D}$.

4. *For any* $\mathscr{D}_1, \mathscr{D}_2 \in \Phi$, *and* $S \subseteq I$, $\epsilon_S(\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2) = \epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2)$.

5. *For any* $\mathscr{D} \in \Phi$ *and* $S, T \subseteq I$, $\epsilon_S(\epsilon_T(\mathscr{D})) = \epsilon_T(\epsilon_S(\mathscr{D})) = \epsilon_{S \cap T}(\mathscr{D})$.

6. *For any* $\mathscr{D} \in \Phi$, $\epsilon_I(\mathscr{D}) = \mathscr{D}$.

It is possible to notice moreover that $(\Phi^+, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ forms a subalgebra of $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$. In what follows, we refer to the former as the *domain-free information algebra of coherent sets of strictly desirable gambles*.

Example 5. *Let us consider again the framework of Example 1, Example 2, and Example 3. Let us suppose now that other two detectives, Bob and Carol, are sent to investigate with Alice. The three detectives start to work together to discover the identity of the murderer. They consider other aspects besides their height, such as their possible motive for the crime, their hair colour and their hair length. For simplicity, let us model these characteristics by using four binary variables:*

- X_1: *representing the height of the murderer.* $X_1 = 0$ *stands for a short murderer (height within the range* $[1.5, 1.7)m$*) and* $X_1 = 1$ *stands for a tall murderer (height in the range* $[1.7, 2]m$*);*

- X_2: *representing the possible motive of the murderer.* $X_2 = 0$ *stands for an economic motive,* $X_2 = 1$ *stands for a motive of passion;*

- X_3: *representing the hair colour of the murderer.* $X_3 = 0$ *stands for a bright one,* $X_3 = 1$ *stands for a dark one;*

- X_4: *representing the haircut of the murderer.* $X_4 = 0$ *stands for a short cut,* $X_4 = 1$ *stands for a long one.*

The three carry out separately their investigations. At the end, they collect together the information obtained.

To model their conclusions about all the characteristics of the murderer, we construct a possibility space collecting together the variables involved. Since we do not have other information, we consider all the combined values of the variables to *be possible, i.e., we assume* $X_1, ..., X_4$ *to be logically independent. Thus, we shall consider* $\mathscr{X} := \times_{i=1}^{4} \mathscr{X}_i$, *where* $\mathscr{X}_i := \{0,1\}$ *is the set of the possible values of the variable* X_i *for every* $i \in \{1,...,4\}$, *and denote by* x *its generic element with components* x_1, x_2, x_3, x_4. *The information that each detective has collected about* $\mathscr{X}$ *can be modeled as a coherent set of gambles:*

- *Alice thinks the murderer is more likely to be tall than short. Therefore, as seen in the previous examples, her beliefs can be represented through the set:*
$$\mathscr{D}_1 := \mathrm{posi}(\{\mathbb{I}_{\{X_1=1\}} - \mathbb{I}_{\{X_1=0\}}\} \cup \mathscr{L}^+).$$

- *Bob analyses the modalities of the crime and he concludes the motive has to be a motive of passion. His beliefs therefore can be expressed through the coherent set of strictly desirable gambles:*
$$\mathscr{D}_2^+ := \{f \in \mathscr{L} : \min_{\{x_1,x_3,x_4\} \in \{0,1\}} f(x_1, 1, x_3, x_4) > 0\} \cup \mathscr{L}^+.$$
This set indeed corresponds to an agent certain that $X_2 = 1$ *and nothing more.*

- *Carol finds a long black hair on the crime scene that cannot have been left there before or after the crime and that does not belong to the victim. She is then sure that it must belong to the murderer. She is also more scrupulous than Alice about the dynamics of the crime, discovering in particular that it must be sure that the murderer is tall. Her beliefs can then be represented through the coherent set of strictly desirable gambles:*
$$\mathscr{D}_3^+ := \{f \in \mathscr{L} : \min_{x_2 \in \{0,1\}} f(1, x_2, 1, 1) > 0\} \cup \mathscr{L}^+.$$
This set indeed corresponds to an agent certain that $X_1 = 1, X_3 = 1, X_4 = 1$ *and nothing more.*

Information algebras give us a tool to combine all these pieces of information together. In this case we obtain:
$$\mathscr{D} := \mathscr{C}(\mathscr{D}_1 \cup \mathscr{D}_2^+ \cup \mathscr{D}_3^+) = \mathscr{C}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+) = \{f \in \mathscr{L} : f(1,1,1,1) > 0\} \cup \mathscr{L}^+. \quad (3.3)$$

Indeed, $\mathscr{D}_1 \subseteq \mathscr{D}_3^+$. *Moreover,* $\mathscr{E}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+) \subseteq \{f \in \mathscr{L} : f(1,1,1,1) > 0\} \cup \mathscr{L}^+$, *hence* $\mathscr{C}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+) = \mathscr{E}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+) \subseteq \{f \in \mathscr{L} : f(1,1,1,1) > 0\} \cup \mathscr{L}^+$.

Vice versa, let us consider $g \in \{f \in \mathscr{L} : f(1,1,1,1) > 0\}$ *and let us define the two gambles:*
$$(\forall x \in \mathscr{X})\ g_1(x) := \begin{cases} \lambda & \text{if } x_2 = 1, x \neq (1,1,1,1), \\ g(x) - \mu & \text{if } x = (1,0,1,1), \\ g(x)/2 & \text{otherwise}, \end{cases}$$

$$(\forall x \in \mathscr{X})\ g_2(x) := \begin{cases} g(x)-\lambda & \textit{if } x_2=1, x \neq (1,1,1,1), \\ \mu & \textit{if } x=(1,0,1,1), \\ g(x)/2 & \textit{otherwise}, \end{cases}$$

with $\lambda, \mu > 0$. *Then,* $g = g_1 + g_2$ *with* $g_1 \in \mathscr{D}_2^+$ *and* $g_2 \in \mathscr{D}_3^+$*, hence* $g \in \mathscr{E}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+) = \mathscr{C}(\mathscr{D}_2^+ \cup \mathscr{D}_3^+)$. *Thus, we obtain Eq.* (3.3).

$\mathscr{D}$, *as expected, corresponds to an agent certain that* $X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 1$. *Thus, the three detectives together can conclude that the murderer has to be tall, with a motive of passion and with long and dark hair.*

Information algebras give us also a tool to extract information about a specific question of interest. Questions of interest, in this case, can be modeled as subsets $S \subseteq \{1,2,3,4\}$. *Let us say we are interested in the information about the motive of the crime that the three detectives have collected together. We obtain:*

$$\epsilon_{\{2\}}(\mathscr{D}) = \epsilon_{\{2\}}(\mathscr{D}_2^+ \cdot \mathscr{D}_3^+) = \mathscr{D}_2^+ \cdot \epsilon_{\{2\}}(\mathscr{D}_3^+) = \mathscr{D}_2^+ \cdot \mathscr{L}^+ = \mathscr{D}_2^+,$$

using $\epsilon_{\{2\}}(\mathscr{D}_2^+) = \mathscr{D}_2^+$*and the Combination axiom of domain-free information algebras. Intuitively, since no one, except for Bob, has information about the crime motive, the information about the motive the group has must correspond to the one of Bob: the motive for the crime has to be a motive of passion.*

Following the standard procedure illustrated above to derive a labeled information algebra from a domain-free one, it is possible to derive a *labeled information algebra of coherent sets of gambles* $(\hat{\Phi}, I; d, \cdot, \{(\mathscr{L}, S)\}_{S \subseteq I}, \{(\mathscr{L}^+, S)\}_{S \subseteq I}, \pi)$ from the domain-free one, where:

- $\hat{\Phi} := \bigcup_{S \subseteq I} \{(\mathscr{D}, S) : \mathscr{D} \in \Phi,\ \epsilon_S(\mathscr{D}) = \mathscr{D}\}$;
- $d : \hat{\Phi} \to \mathrm{P}(I)$, defined by $(\mathscr{D}, S) \mapsto d(\mathscr{D}, S) := S$;
- $\cdot : \hat{\Phi} \times \hat{\Phi} \to \hat{\Phi}$, defined by
$$(\mathscr{D}_1, S), (\mathscr{D}_2, T) \mapsto (\mathscr{D}_1, S) \cdot (\mathscr{D}_2, T) := (\mathscr{D}_1 \cdot \mathscr{D}_2, S \cup T);$$
- $\pi : dom(\pi) \subseteq \hat{\Phi} \times \mathrm{P}(I) \to \hat{\Phi}$, defined by
$$(\mathscr{D}, S), T \mapsto \pi_T(\mathscr{D}, S) := (\epsilon_T(\mathscr{D}), T),$$
for every $T \subseteq S \subseteq I$.

As mentioned above, coherent sets $\mathscr{D} = \epsilon_S(\mathscr{D}) = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_S)$ contain the same information of their $\mathscr{P}_S$—marginals $\mathscr{D} \cap \mathscr{L}_S$. As shown in Section 1 moreover, the latter are in a one-to-one correspondence with sets $\tilde{\mathscr{D}}$ directly defined on blocks of $\mathscr{P}_S$, i.e., on Ω_S. Therefore, also the sets $\mathscr{D} = \epsilon_S(\mathscr{D}) = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_S)$ are inherently determined by gambles on $\mathscr{L}(\Omega_S)$. Representing them as couples $(\mathscr{D}, S)$, thus, is a waste of computational resources since sets $\mathscr{D}$ are still defined on the whole set

Ω. For this reason, we propose an alternative version of this labeled information algebra where sets are directly determined by gambles defined on Ω_S. We then show that these two labeled information algebras are isomorphic.

For every subset S of I, let us define

$$\tilde{\Phi}_S(\Omega) := \{(\tilde{\mathscr{D}}, S) : \ \tilde{\mathscr{D}} \in \Phi(\Omega_S) := \mathbb{D}(\Omega_S) \cup \{\mathscr{L}(\Omega_S)\}\}$$

and

$$\tilde{\Phi}(\Omega) := \bigcup_{S \subseteq I} \tilde{\Phi}_S(\Omega),$$

often shortened to $\tilde{\Phi}_S$ and $\tilde{\Phi}$ respectively.

As introduced in Section 1, there is a one-to-one correspondence between gambles in $\mathscr{L}_S(\Omega_R)$, with $S \subseteq R \subseteq I$, and gambles in $\mathscr{L}(\Omega_S)$. In particular, given $f \in \mathscr{L}_S(\Omega_R)$, we indicate with $f^{\downarrow S}$ the corresponding gamble in $\mathscr{L}(\Omega_S)$ defined as $f^{\downarrow S}(\omega_S) := f(\omega_R)$ for every $\omega_S \in \Omega_S$ and $\omega_R \in \Omega_R$ such that $\omega_R|S = \omega_S$. Analogously, given $f \in \mathscr{L}(\Omega_S)$ we indicate with $f^{\uparrow R}$ the corresponding gamble in $\mathscr{L}_S(\Omega_R)$ defined as $f^{\uparrow R}(\omega_R) := f(\omega_R|S)$ for every $\omega_R \in \Omega_R$. Similar maps can be defined at the level of sets of gambles, operating gamble by gamble. This allows us to define on $\tilde{\Phi}$ and $\mathrm{P}(I)$ the following operations.

1. Labeling. $d : \tilde{\Phi} \to \mathrm{P}(I)$, defined by $(\tilde{\mathscr{D}}, S) \mapsto d(\tilde{\mathscr{D}}, S) := S$.

2. Combination. $\cdot : \tilde{\Phi} \times \tilde{\Phi} \to \tilde{\Phi}$, defined by
$$(\tilde{\mathscr{D}}_1, S), (\tilde{\mathscr{D}}_2, T) \mapsto (\tilde{\mathscr{D}}_1, S) \cdot (\tilde{\mathscr{D}}_2, T) := (\mathscr{C}(\tilde{\mathscr{D}}_1^{\uparrow S \cup T}) \cdot \mathscr{C}(\tilde{\mathscr{D}}_2^{\uparrow S \cup T}), S \cup T),$$
where $\mathscr{C}(\tilde{\mathscr{D}}_1^{\uparrow S \cup T}) \cdot \mathscr{C}(\tilde{\mathscr{D}}_2^{\uparrow S \cup T})$ is the combination defined for sets in $\Phi(\Omega_{S \cup T})$.

3. Marginalisation. $\pi : dom(\pi) \subseteq \tilde{\Phi} \times \mathrm{P}(I) \to \tilde{\Phi}$, defined by
$$(\tilde{\mathscr{D}}, S), T \mapsto \pi_T(\tilde{\mathscr{D}}, S) := ((\epsilon_T(\tilde{\mathscr{D}}) \cap \mathscr{L}_T(\Omega_S))^{\downarrow T}, T),$$
for every $T \subseteq S \subseteq I$, where $\epsilon_T(\tilde{\mathscr{D}})$ is the extraction defined for sets in $\Phi(\Omega_S)$ and $\mathrm{P}(I)$.

Now, consider the map $h : \hat{\Phi} \to \tilde{\Phi}$, defined by

$$\mathscr{D}, S \mapsto h(\mathscr{D}, S) := ((\epsilon_S(\mathscr{D}) \cap \mathscr{L}_S(\Omega))^{\downarrow S}, S) = ((\mathscr{D} \cap \mathscr{L}_S(\Omega))^{\downarrow S}, S). \tag{3.4}$$

The following theorem shows that h is bijiective and mantains operations defined on $\hat{\Phi}$ and $\mathrm{P}(I)$. Therefore, it follows that: $(\tilde{\Phi}, I; d, \cdot, \{(\mathscr{L}(\Omega_S), S)\}_{S \subseteq I}, \{(\mathscr{L}^+(\Omega_S), S)\}_{S \subseteq I}, \pi)$ satisfies the axioms defining a labeled information algebra and it is isomorphic to $(\hat{\Phi}, I; d, \cdot, \{(\mathscr{L}(\Omega), S)\}_{S \subseteq I}, \{(\mathscr{L}^+(\Omega), S)\}_{S \subseteq I}, \pi)$. From now on, we refer only to the former as the labeled information algebra of coherent sets of gambles.

Theorem 9. *The map h defined in Eq. (3.4) satisfies the following properties.*

1. *For any* $(\mathscr{D}_1, S), (\mathscr{D}_2, T) \in \hat{\Phi}$,
$$h((\mathscr{D}_1, S) \cdot (\mathscr{D}_2, T)) = h(\mathscr{D}_1, S) \cdot h(\mathscr{D}_2, T);$$
2. *for any* $S \subseteq I$,
$$h(\mathscr{L}(\Omega), S) = (\mathscr{L}(\Omega_S), S);$$
3. *for any* $S \subseteq I$,
$$h(\mathscr{L}^+(\Omega), S) = (\mathscr{L}^+(\Omega_S), S);$$
4. *for any* $(\mathscr{D}, S) \in \hat{\Phi}$, $T \subseteq S$,
$$h(\pi_T(\mathscr{D}, S)) = \pi_T(h(\mathscr{D}, S));$$
5. *h is bijective.*

Example 6. *Let us consider again the framework of Example 5. It is possible to notice that:*

- $\epsilon_{S_1}(\mathscr{D}_1) = \mathscr{D}_1$; *where* $S_1 := \{1\}$;
- $\epsilon_{S_2}(\mathscr{D}_2^+) = \mathscr{D}_2^+$, *where* $S_2 := \{2\}$;
- $\epsilon_{S_3}(\mathscr{D}_3^+) = \mathscr{D}_3^+$, *where* $S_3 := \{1,3,4\}$.

Therefore, $(\mathscr{D}_1, S_1), (\mathscr{D}_2^+, S_2), (\mathscr{D}_3^+, S_3) \in \hat{\Phi}$. *Let us consider now* $\mathscr{X}_S := \times_{i \in S} \mathscr{X}_i$ *for every* $S \subseteq \{1,2,3,4\}$, *and the map h defined in Eq. (3.4). We have:*

- $h(\mathscr{D}_1, S_1) = (\tilde{\mathscr{D}}_1, S_1)$, *where* $\tilde{\mathscr{D}}_1 := \operatorname{posi}(\{\mathbb{I}_{\{1\}} - \mathbb{I}_{\{0\}}\} \cup \mathscr{L}^+(\mathscr{X}_{S_1}))$;
- $h(\mathscr{D}_2^+, S_2) = (\tilde{\mathscr{D}}_2^+, S_2)$, *where* $\tilde{\mathscr{D}}_2^+ := \{f \in \mathscr{L}(\mathscr{X}_{S_2}) : f(1) > 0\} \cup \mathscr{L}^+(\mathscr{X}_{S_2})$;
- $h(\mathscr{D}_3^+, S_3) = (\tilde{\mathscr{D}}_3^+, S_3)$, *where* $\tilde{\mathscr{D}}_3^+ := \{f \in \mathscr{L}(\mathscr{X}_{S_3}) : f(1,1,1) > 0\} \cup \mathscr{L}^+(\mathscr{X}_{S_3})$.

Theorem 9 guarantees that h maintains combination and extraction, therefore
$$(\tilde{\mathscr{D}}_1, S_1) \cdot (\tilde{\mathscr{D}}_2^+, S_2) \cdot (\tilde{\mathscr{D}}_3^+, S_3) = h(\mathscr{D}_1, S_1) \cdot h(\mathscr{D}_2^+, S_2) \cdot h(\mathscr{D}_3^+, S_3) =$$
$$= h((\mathscr{D}_1, S_1) \cdot (\mathscr{D}_2^+, S_2) \cdot (\mathscr{D}_3^+, S_3)) := h(\mathscr{D}_1 \cdot \mathscr{D}_2^+ \cdot \mathscr{D}_3^+, I) = h(\mathscr{D}, I) = (\mathscr{D}, I),$$
and
$$\pi_{S_2}((\tilde{\mathscr{D}}_1, S_1) \cdot (\tilde{\mathscr{D}}_2^+, S_2) \cdot (\tilde{\mathscr{D}}_3^+, S_3)) = \pi_{S_2}(h(\mathscr{D}_1, S_1) \cdot h(\mathscr{D}_2^+, S_2) \cdot h(\mathscr{D}_3^+, S_3)) =$$
$$= \pi_{S_2}(h(\mathscr{D}, I)) = h(\pi_{S_2}(\mathscr{D}, I)) := h(\epsilon_{S_2}(\mathscr{D}), S_2) = h(\mathscr{D}_2^+, S_2) = (\tilde{\mathscr{D}}_2^+, S_2),$$
where
$$\mathscr{D} := \{f \in \mathscr{L}(\mathscr{X}) : f(1,1,1,1) > 0\} \cup \mathscr{L}^+(\mathscr{X}),$$
as in Example 5.

We then showed that both the domain-free and the labeled version of the information algebra of coherent sets of gambles are atomistic.

Maximal coherent sets of gambles $M \in \mathbb{M}$, see Section 1, are atoms in $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$. Indeed, they differ from $\mathscr{L}$ and they have the property that, in information order,

$$M \leq \mathscr{D} \text{ for } \mathscr{D} \in \Phi \Rightarrow \mathscr{D} = M \text{ or } \mathscr{D} = \mathscr{L}.$$

It is easy to prove in fact that, in this case, the information order coincides with set inclusion, i.e., $\mathscr{D}_1 \leq \mathscr{D}_2 \iff \mathscr{D}_1 \subseteq \mathscr{D}_2$.

Let $At(\Phi)$ denote the set of atoms of $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$, i.e., the set of maximal sets of gambles. For any set $\mathscr{D} \in \Phi$ then, let $At(\mathscr{D})$ denote the set of maximal sets of gambles containing $\mathscr{D}$,

$$At(\mathscr{D}) := \{M \in At(\Phi) : \mathscr{D} \leq M\}.$$

The following properties are satisfied.

1. For any set $\mathscr{D} \in \mathbb{D}$, there is a set $M \in At(\Phi)$ so that in information order $\mathscr{D} \leq M$ (i.e., $\mathscr{D} \subseteq M$). Hence, $At(\mathscr{D})$, for $\mathscr{D}$ coherent, is never empty.

2. For any set $\mathscr{D} \in \mathbb{D}$, we have
$$\mathscr{D} = \bigcap At(\mathscr{D}).$$

These properties are recalled in Section 1 and proved in de Cooman and Quaeghebeur, 2012, Theorem 3, Corollary 4. Thus, $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ is atomistic.

Regarding its labeled version instead, we have that $(\tilde{M}, S)$ where $\tilde{M} \in \mathbb{M}(\Omega_S)$ are atoms relative to S, for every $S \subseteq I$. Let us indicate with $At_S(\tilde{\Phi})$ the set of all its atoms relative to S and with $At_S(\tilde{\mathscr{D}}, S)$ the set of elements of $At_S(\tilde{\Phi})$ dominating (in information order) $(\tilde{\mathscr{D}}, S)$, for every $(\tilde{\mathscr{D}}, S) \in \tilde{\Phi}$.

The properties of the domain-free information algebra $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ of being atomic and atomistic carry over to its labeled version.

1. *Atomic:* For any element $(\tilde{\mathscr{D}}, S) \in \tilde{\Phi}$, $S \subseteq I$ with $\tilde{\mathscr{D}} \in \mathbb{D}(\Omega_S)$, there is an atom relative to S, $(\tilde{M}, S) \in At_S(\tilde{\Phi})$, so that $(\tilde{\mathscr{D}}, S) \leq (\tilde{M}, S)$.

2. *Atomistic:* For any element $(\tilde{\mathscr{D}}, S) \in \tilde{\Phi}$, $S \subseteq I$, with $\tilde{\mathscr{D}} \in \mathbb{D}(\Omega_S)$, $(\tilde{\mathscr{D}}, S) = \bigwedge\{(\tilde{M}, S) : (\tilde{M}, S) \in At_S(\tilde{\mathscr{D}}, S)\}$.

3.1.2 Information algebras of coherent lower previsions

From the domain-free information algebra of coherent sets of strictly desirable gambles, it is possible to derive a domain-free information algebra of coherent lower previsions.

Let us consider, as before, a possibility space $\Omega := \times_{i\in I} \Omega_i$ where I is an index set and Ω_i is the set of the possible values of a variable X_i, for every $i \in I$. In this context, let us consider the domain-free information algebra of coherent sets of strictly desirable gambles defined on Ω, $(\Phi^+(\Omega), I; \cdot, \mathscr{L}(\Omega), \mathscr{L}^+(\Omega), \epsilon)$, and $\underline{\Phi}(\Omega) := \underline{\mathbb{P}}(\Omega) \cup \{\sigma(\mathscr{L}(\Omega))\}$, as defined in Section 1.

On $\underline{\Phi}$ and $\mathrm{P}(I)$ it is possible to define an operation of combination and one of extraction analogous to the ones of $(\Phi^+, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$.

1. Combination. $\cdot : \underline{\Phi} \times \underline{\Phi} \to \underline{\Phi}$, defined by
$$\underline{P}_1, \underline{P}_2 \mapsto \underline{P}_1 \cdot \underline{P}_2 := \underline{E}^*(\max\{\underline{P}_1, \underline{P}_2\}), \tag{3.5}$$
where $\underline{E}^*$ is defined in Definition 17. [4]

2. Extraction. $\underline{e} : \underline{\Phi} \times \mathrm{P}(I) \to \underline{\Phi}$, defined by
$$\underline{P}, S \mapsto \underline{e}_S(\underline{P}) := \underline{E}^*(\underline{P}_S),$$
where $\underline{P}_S$ is the $\mathscr{P}_S$-marginal of $\underline{P}$, see Definition 16.

With the following result it is possible to show that $\underline{\Phi}$ and $\mathrm{P}(I)$, with the two operations defined above, induce a domain-free information algebra $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$, which is in particular isomorphic to $(\Phi^+, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$, subalgebra of $(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$. In what follows, we refer to $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$ as the *domain-free information algebra of coherent lower previsions*.

Theorem 10. *Let $\mathscr{D}_1^+$, $\mathscr{D}_2^+, \mathscr{D}^+ \subseteq \mathscr{L}$ be coherent sets of strictly desirable gambles and $S \subseteq I$. Then*

1. $(\forall f \in \mathscr{L})\ \sigma(\mathscr{L})(f) = \infty$, $\sigma(\mathscr{L}^+)(f) = \inf f$,

2. $\sigma(\mathscr{D}_1^+ \cdot \mathscr{D}_2^+) = \sigma(\mathscr{D}_1^+) \cdot \sigma(\mathscr{D}_2^+)$,

3. $\sigma(\epsilon_S(\mathscr{D}^+)) = \underline{e}_S(\sigma(\mathscr{D}^+))$.

$(\Phi, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ is instead only *weakly homomorphic* to $(\Phi^+, I; \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ and $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$, in the sense specified by the following theorem and corollary.

[4] It can be regarded as the analogous of the combination operation defined on elements of Φ^+. Indeed, let us consider $\mathscr{D}_1^+, \mathscr{D}_2^+ \in \Phi^+$ such that $\mathscr{D}_1^+ \cdot \mathscr{D}_2^+ \neq \mathscr{L}$ and define $\underline{P}_1 := \sigma(\mathscr{D}_1^+)$ and $\underline{P}_2 := \sigma(\mathscr{D}_2^+)$. We have
$$\sigma(\mathscr{D}_1^+ \cup \mathscr{D}_2^+)(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathscr{D}_1^+ \cup \mathscr{D}_2^+\} = \max\{\underline{P}_1(f), \underline{P}_2(f)\},$$
for every $f \in dom(\sigma(\mathscr{D}_1^+ \cup \mathscr{D}_2^+))$. See also Theorem 10.

Theorem 11. *Let $\mathscr{D}_1$, $\mathscr{D}_2$ and $\mathscr{D}$ be coherent sets of gambles and $S \subseteq I$. Consider also the map from $\mathbb{D}$ to $\mathbb{D}^+$: $\mathscr{D} \to \mathscr{D}^+ := \tau^+(\sigma(\mathscr{D}))$.*

1. $\mathscr{D}_1 \cdot \mathscr{D}_2 \neq \mathscr{L} \Rightarrow \mathscr{D}_1 \cdot \mathscr{D}_2 \mapsto (\mathscr{D}_1 \cdot \mathscr{D}_2)^+ = \mathscr{D}_1^+ \cdot \mathscr{D}_2^+$,
2. $\epsilon_S(\mathscr{D}) \mapsto (\epsilon_S(\mathscr{D}))^+ = \epsilon_S(\mathscr{D}^+)$.

Corollary 3. *Let $\mathscr{D}_1$, $\mathscr{D}_2$ and $\mathscr{D}$ be coherent sets of gambles and $S \subseteq I$.*

1. $\mathscr{D}_1 \cdot \mathscr{D}_2 \neq \mathscr{L} \Rightarrow \sigma(\mathscr{D}_1 \cdot \mathscr{D}_2) = \sigma(\mathscr{D}_1) \cdot \sigma(\mathscr{D}_2)$,
2. $\sigma(\epsilon_S(\mathscr{D})) = \underline{e}_S(\sigma(\mathscr{D}))$.

The weak homomorphism cannot be extended to a homomorphism as the following example shows.

Example 7. *Let us consider the framework of the previous examples in this section. Consider moreover the following slightly different versions of $\mathscr{D}_2^+$.*

$$\mathscr{D}_2' := \mathscr{D}_2^+ \cup \{f \in \mathscr{L} : \min_{\{x_1,x_3,x_4\} \in \{0,1\}} f(x_1,1,x_3,x_4) = 0 < \min_{\{x_3,x_4\} \in \{0,1\}} f(0,0,x_3,x_4)\},$$

$$\mathscr{D}_2'' := \mathscr{D}_2^+ \cup \{f \in \mathscr{L} : \min_{\{x_1,x_3,x_4\} \in \{0,1\}} f(x_1,1,x_3,x_4) = 0 < \min_{\{x_3,x_4\} \in \{0,1\}} f(1,0,x_3,x_4)\}.$$

They are coherent. Moreover, by Eq. (1.10), we know that $\tau^+(\sigma(\mathscr{D}_2')) = \tau^+(\sigma(\mathscr{D}_2'')) = \mathscr{D}_2^+$.

We have also $\mathscr{D}_2' \cdot \mathscr{D}_2'' = \mathscr{L}$. Indeed, $0 = f + g$, where $f \in \mathscr{D}_2'$, $g \in \mathscr{D}_2''$ are defined as:

$$(\forall x \in \mathscr{X})\, f(x) := \begin{cases} 1 & \text{if } x = (0,0,0,0) \\ 1 & \text{if } x = (0,0,0,1) \\ 1 & \text{if } x = (0,0,1,0) \\ 1 & \text{if } x = (0,0,1,1) \\ -1 & \text{if } x = (1,0,0,0), \\ -1 & \text{if } x = (1,0,0,1) \\ -1 & \text{if } x = (1,0,1,0) \\ -1 & \text{if } x = (1,0,1,1) \\ 0 & \text{otherwise} \end{cases} \quad g(x) := \begin{cases} -1 & \text{if } x = (0,0,0,0) \\ -1 & \text{if } x = (0,0,0,1) \\ -1 & \text{if } x = (0,0,1,0) \\ -1 & \text{if } x = (0,0,1,1) \\ 1 & \text{if } x = (1,0,0,0) \\ 1 & \text{if } x = (1,0,0,1) \\ 1 & \text{if } x = (1,0,1,0) \\ 1 & \text{if } x = (1,0,1,1) \\ 0 & \text{otherwise.} \end{cases}$$

So that, Theorem 11 is not applicable. Indeed, we have $\mathscr{L} = (\mathscr{D}_2' \cdot \mathscr{D}_2'')^+ \neq (\mathscr{D}_2')^+ \cdot (\mathscr{D}_2'')^+ := \tau^+(\sigma(\mathscr{D}_2')) \cdot \tau^+(\sigma(\mathscr{D}_2'')) = \mathscr{D}_2^+ \cdot \mathscr{D}_2^+ = \mathscr{D}_2^+$. Analogous reasoning can be made regarding results of Corollary 3. Indeed, we have $\sigma(\mathscr{D}_2^+) = \sigma(\mathscr{D}_2^+) \cdot \sigma(\mathscr{D}_2^+) = \sigma(\mathscr{D}_2') \cdot \sigma(\mathscr{D}_2'') \neq \sigma(\mathscr{D}_2' \cdot \mathscr{D}_2'') = \sigma(\mathscr{L})$.

Let us consider instead $\mathscr{D}_1, \mathscr{D}_2^+, \mathscr{D}_3^+$ as defined in Example 5. In this case we have $\mathscr{D} := \mathscr{D}_1 \cdot \mathscr{D}_2^+ \cdot \mathscr{D}_3^+ \neq \mathscr{L}$. Therefore, $\mathscr{D}_1^+ \cdot \mathscr{D}_2^+ \cdot \mathscr{D}_3^+ = \mathscr{D}^+ = \mathscr{D}$.

Moving to lower previsions, we have:

$$(\forall f \in \mathscr{L})\ \underline{P}_1(f) := \sigma(\mathscr{D}_1)(f) = \inf\{P(f) : P \in \mathbb{P}, P(X_1 = 1) \geq P(X_1 = 0)\},$$

where the equality is proven in Example 3,

$$(\forall f \in \mathscr{L})\ \underline{P}_2(f) := \sigma(\mathscr{D}_2^+)(f) = \sigma(\mathscr{D}_2')(f) = \sigma(\mathscr{D}_2'')(f) = \min_{\{x_1,x_3,x_4\}\in\{0,1\}} f(x_1, 1, x_3, x_4),$$

$$(\forall f \in \mathscr{L})\ \underline{P}_3(f) := \sigma(\mathscr{D}_3^+)(f) = \min_{x_2\in\{0,1\}} f(1, x_2, 1, 1).$$

Since $\mathscr{D} := \mathscr{D}_1 \cdot \mathscr{D}_2^+ \cdot \mathscr{D}_3^+ \neq \mathscr{L}$, also Corollary 3 is applicable. Hence, we have

$$\sigma(\mathscr{D}_1) \cdot \sigma(\mathscr{D}_2^+) \cdot \sigma(\mathscr{D}_3^+) = \sigma(\mathscr{D}_1 \cdot \mathscr{D}_2^+ \cdot \mathscr{D}_3^+) = \sigma(\mathscr{D}),$$

where, in particular, $\sigma(\mathscr{D})(f) = f(1, 1, 1, 1)$ for every $f \in \mathscr{L}$. Hence, it is a linear *prevision. Moreover, we have:*

$$\underline{e}_{S_2}(\sigma(\mathscr{D}_1) \cdot \sigma(\mathscr{D}_2^+) \cdot \sigma(\mathscr{D}_3^+)) = \underline{e}_{S_2}(\sigma(\mathscr{D})) = \sigma(\epsilon_{S_2}(\mathscr{D})) = \sigma(\mathscr{D}_2^+).$$

From the domain-free information algebra of coherent lower previsions, we can deduce two labeled information algebras, following the same reasoning used for coherent sets of gambles. The first one is the following:

$$(\underline{\hat{\Phi}}, I; d, \cdot, \{(\sigma(\mathscr{L}), S)\}_{S\subseteq I}, \{(\sigma(\mathscr{L}^+), S)\}_{S\subseteq I}, \underline{\pi}),$$

where:

- $\underline{\hat{\Phi}} := \bigcup_{S\subseteq I}\{(\underline{P}, S) :\ \underline{P} \in \underline{\Phi},\ \underline{e}_S(\underline{P}) = \underline{P}\}$;
- $d : \underline{\hat{\Phi}} \to \mathrm{P}(I)$, defined by $(\underline{P}, S) \mapsto d(\underline{P}, S) := S$;
- $\cdot : \underline{\hat{\Phi}} \times \underline{\hat{\Phi}} \to \underline{\hat{\Phi}}$, defined by
$$(\underline{P}_1, S), (\underline{P}_2, T) \mapsto (\underline{P}_1, S) \cdot (\underline{P}_2, T) := (\underline{P}_1 \cdot \underline{P}_2, S \cup T);$$
- $\underline{\pi} : dom(\underline{\pi}) \subseteq \underline{\hat{\Phi}} \times \mathrm{P}(I) \to \underline{\hat{\Phi}}$, defined by
$$(\underline{P}, S), T \mapsto \underline{\pi}_T(\underline{P}, S) := (\underline{e}_T(\underline{P}), T),$$
for every $T \subseteq S \subseteq I$.

The second one is introduced, as usual, to save computational resources. As before, in what follows we refer only to this one as the labeled version of the information algebra of coherent lower previsions. For any set $S \subseteq I$, we define

$$\underline{\tilde{\Phi}}_S(\Omega) := \{(\underline{\tilde{P}}, S) : \underline{\tilde{P}} \in \underline{\Phi}(\Omega_S)\},$$

where $\underline{\Phi}(\Omega_S) := \underline{\mathbb{P}}(\Omega_S) \cup \{\sigma(\mathscr{L}(\Omega_S))\}$. Furthermore, we consider

$$\underline{\tilde{\Phi}}(\Omega) := \bigcup_{S\subseteq I} \underline{\tilde{\Phi}}_S(\Omega).$$

As usual, we omit the possibility space whenever possible.

On $\tilde{\underline{\Phi}}$ and $\mathrm{P}(I)$ we define the following operations.

1. Labeling: $d : \tilde{\underline{\Phi}} \to \mathrm{P}(I)$, defined by $(\tilde{\underline{P}}, S) \mapsto d(\tilde{\underline{P}}, S) := S$.
2. Combination. $\cdot : \tilde{\underline{\Phi}} \times \tilde{\underline{\Phi}} \to \tilde{\underline{\Phi}}$, defined by
$$(\tilde{\underline{P}}_1, S), (\tilde{\underline{P}}_2, T) \mapsto (\tilde{\underline{P}}_1, S) \cdot (\tilde{\underline{P}}_2, T) := (\underline{E}^*(\tilde{\underline{P}}_1^{\uparrow S\cup T}) \cdot \underline{E}^*(\tilde{\underline{P}}_2^{\uparrow S\cup T}), S \cup T),$$
where given a lower prevision $\underline{P}$ with $dom(\underline{P}) \subseteq \mathscr{L}(\Omega_Z)$ with $Z \subseteq S \cup T$, $\underline{P}^{\uparrow S\cup T}(f) := \underline{P}(f^{\downarrow Z})$, for every $f \in \mathscr{L}_Z(\Omega_{S\cup T})$ such that $f^{\downarrow Z} \in dom(\underline{P})$.
3. Marginalisation. $\underline{\pi} : dom(\underline{\pi}) \subseteq \tilde{\underline{\Phi}} \times \mathrm{P}(I) \to \tilde{\underline{\Phi}}$, defined by
$$(\tilde{\underline{P}}, S), T \mapsto \underline{\pi}_T(\tilde{\underline{P}}, S) := (\underline{e}_T(\tilde{\underline{P}})_T^{\downarrow T}, T),$$
for every $T \subseteq S \subseteq I$, where given a lower prevision $\underline{P}$ with $dom(\underline{P}) \subseteq \mathscr{L}(\Omega_Z)$ with $T \subseteq Z$, $\underline{P}_T^{\downarrow T}(f) := \underline{P}_T(f^{\uparrow Z}) = \underline{P}(f^{\uparrow Z})$ for every $f \in \mathscr{L}(\Omega_T)$ such that $f^{\uparrow Z} \in dom(\underline{P})$.

$\tilde{\underline{\Phi}}(\Omega)$ and $\mathrm{P}(I)$ equipped with these operations form a labeled information algebra $(\tilde{\underline{\Phi}}(\Omega), I; d, \cdot, \{(\sigma(\mathscr{L}(\Omega_S)), S)\}_{S\subseteq I}, \{(\sigma(\mathscr{L}^+(\Omega_S)), S)\}_{S\subseteq I}, \underline{\pi})$. Indeed, again analogously to what we have done for coherent set of gambles, we can introduce the map $\underline{h} : \hat{\underline{\Phi}} \to \tilde{\underline{\Phi}}$, defined by
$$(\underline{P}, S) \mapsto \underline{h}(\underline{P}, S) := (\underline{e}_S(\underline{P})_S^{\downarrow S}, S) = (\underline{P}_S^{\downarrow S}, S) = (\sigma((\tau^+(\underline{P}) \cap \mathscr{L}_S)^{\downarrow S}), S).^5 \tag{3.6}$$
Thus, introducing the function $\tilde{\sigma} : \tilde{\Phi} \to \tilde{\underline{\Phi}}$, defined by
$$(\tilde{\mathscr{D}}, S) \mapsto \tilde{\sigma}(\tilde{\mathscr{D}}, S) := (\sigma(\tilde{\mathscr{D}}), S),$$
we have
$$\underline{h}(\underline{P}, S) = \tilde{\sigma}(h(\tau^+(\underline{P}), S)).$$

We then show that $\underline{h}$ is bijiective and mantains operations.

Theorem 12. *The map $\underline{h}$ defined in Eq. (3.6) satisfies the following properties.*

1. *For any $(\underline{P}_1, S)$, $(\underline{P}_2, T) \in \hat{\underline{\Phi}}(\Omega)$,*
$$\underline{h}((\underline{P}_1, S) \cdot (\underline{P}_2, T)) = \underline{h}(\underline{P}_1, S) \cdot \underline{h}(\underline{P}_2, T);$$
2. *for any $S \subseteq I$,*
$$\underline{h}(\sigma(\mathscr{L}(\Omega)), S) = (\sigma(\mathscr{L}(\Omega_S)), S);$$

[5] The last equivalence can be proven as follows.
$$(\forall f \in \mathscr{L}(\Omega_S))\ \sigma((\tau^+(\underline{P}) \cap \mathscr{L}_S)^{\downarrow S})(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in (\tau^+(\underline{P}) \cap \mathscr{L}_S)^{\downarrow S}\}$$
$$= \sup\{\mu \in \mathbb{R} : f^{\uparrow I} - \mu \in (\tau^+(\underline{P}) \cap \mathscr{L}_S)\} =: \underline{P}_S^{\downarrow S}(f).$$

3. *for any* $S \subseteq I$,
$$\underline{h}(\sigma(\mathscr{L}^+(\Omega)), S) = (\sigma(\mathscr{L}^+(\Omega_S)), S);$$

4. *for any* $(\underline{P}, S) \in \underline{\hat{\Phi}}(\Omega)$, $T \subseteq S$,
$$\underline{h}(\underline{\pi}_T(\underline{P}, S)) = \underline{\pi}_T(\underline{h}(\underline{P}, S));$$

5. $\underline{h}$ *is bijective.*

Example 8. *Let us consider again the framework of the previous examples in this section.*

It is possible to notice that $(\underline{P}_1, S_1), (\underline{P}_2, S_2), (\underline{P}_3, S_3) \in \underline{\hat{\Phi}}$. *A formal proof of this fact is provided by Corollary 15 in Appendix C.1. Thus, we can define:*

- $(\underline{\tilde{P}}_1, S_1) := \underline{h}(\underline{P}_1, S_1) = \tilde{\sigma}(h(\tau^+(\underline{P}_1), S_1)) = (\sigma(\tilde{\mathscr{D}}_1), S_1)$, *where*
$$(\forall f \in \mathscr{L}(\mathscr{X}_{S_1}))\ \sigma(\tilde{\mathscr{D}}_1)(f) = \inf\{P(f) \in \mathbb{P} : P(1) \geq P(0)\}.$$
In particular, to prove the equality $\tilde{\sigma}(h(\tau^+(\underline{P}_1), S_1)) = (\sigma(\tilde{\mathscr{D}}_1), S_1)$, *we can observe that*
$$\tilde{\sigma}(h(\tau^+(\underline{P}_1), S_1)) = \tilde{\sigma}(h(\mathscr{D}_1^+, S_1)) := (\sigma((\mathscr{D}_1^+ \cap \mathscr{L}_{S_1})^{\downarrow S_1}), S_1),$$
where $\mathscr{D}_1^+ := \tau^+(\sigma(\mathscr{D}_1)) = \tau^+(\underline{P}_1)$, *as defined above in this subsection.*[6] *Now, since by construction and Eq. (1.10), we have*
$$(\mathscr{D}_1^+ \cap \mathscr{L}_{S_1})^{\downarrow S_1} = ((\mathscr{D}_1 \cap \mathscr{L}_{S_1})^{\downarrow S_1})^+,$$
it follows that
$$\begin{aligned}\tilde{\sigma}(h(\tau^+(\underline{P}_1), S_1)) &= (\sigma(((\mathscr{D}_1 \cap \mathscr{L}_{S_1})^{\downarrow S_1})^+), S_1) = \\ &= (\sigma((\mathscr{D}_1 \cap \mathscr{L}_{S_1})^{\downarrow S_1}), S_1) =: (\sigma(\tilde{\mathscr{D}}_1), S_1).\end{aligned}$$

- $(\underline{\tilde{P}}_2, S_2) := \underline{h}(\underline{P}_2, S_2) = \tilde{\sigma}(h(\tau^+(\underline{P}_2), S_2)) = \tilde{\sigma}(h(\mathscr{D}_2^+, S_2)) = (\sigma(\tilde{\mathscr{D}}_2^+), S_2)$, *where*
$$(\forall f \in \mathscr{L}(\mathscr{X}_{S_2}))\ \sigma(\tilde{\mathscr{D}}_2^+)(f) := f(1),$$

- $(\underline{\tilde{P}}_3, S_3) := \underline{h}(\underline{P}_3, S_3) = \tilde{\sigma}(h(\tau^+(\underline{P}_3), S_3)) = \tilde{\sigma}(h(\mathscr{D}_3^+, S_3)) = (\sigma(\tilde{\mathscr{D}}_3^+), S_3)$, *where*
$$(\forall f \in \mathscr{L}(\mathscr{X}_{S_3}))\ \sigma(\tilde{\mathscr{D}}_3^+)(f) := f(1,1,1).$$

Additional results analogous to the ones provided for sets of gambles in Example 6 can be shown also for lower previsions.

As for coherent sets of gambles, both the domain-free and the labeled information algebras of coherent lower previsions, are atomistic. On $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}),$

[6] $(\mathscr{D}_1^+, S_1) \in \hat{\Phi}$. Indeed, $\epsilon_{S_1}(\mathscr{D}_1^+) = (\epsilon_{S_1}(\mathscr{D}_1))^+ = \mathscr{D}_1^+$ by Theorem 11 and $\epsilon_S(\mathscr{D}_1) = \mathscr{D}_1$.

$\sigma(\mathscr{L}^+), \underline{e})$ indeed, we can establish a partial order $\leq$ on its pieces of information that coincides with dominance on lower previsions in $\underline{\Phi}$. Linear previsions are then the maximal elements in $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$ with respect to this order.

Lemma 7. *Let $\underline{P}$ be an element of $\underline{\Phi}$ and P a linear prevision. Then $P \leq \underline{P}$ implies either $\underline{P} = P$ or $\underline{P}(f) = +\infty$ for all $f \in \mathscr{L}$.*

In particular, $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$ is atomistic. Indeed, let us define $At(\underline{\Phi})$ as the set of atoms of $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$, i.e., the set of linear previsions. For any lower prevision $\underline{P} \in \underline{\Phi}$, let then $At(\underline{P})$ denote the subset of linear previsions dominating $\underline{P}$,

$$At(\underline{P}) := \{P \in At(\underline{\Phi}) : \underline{P} \leq P\}.$$

The following properties are true.

Theorem 13. *Consider the set of lower previsions $\underline{\Phi}$. If $\underline{P} \in \underline{\mathbb{P}}$, then $At(\underline{P}) \neq \emptyset$ and*

$$\underline{P} = \min At(\underline{P}).$$

For the proof of this theorem, see Theorem 2.6.3 and Theorem 3.3.3 in Walley [1991].

The elements $(\tilde{P}, S)$, where $\tilde{P} \in \mathbb{P}(\Omega_S)$ and $S \subseteq I$, are instead atoms relative to S of the labeled information algebra $(\underline{\tilde{\Phi}}(\Omega), I; d, \cdot, \{(\sigma(\mathscr{L}(\Omega_S)), S)\}_{S\subseteq I}, \{(\sigma(\mathscr{L}^+(\Omega_S)), S)\}_{S\subseteq I}, \underline{\pi})$, that is, if $(\tilde{P}, S) \leq (\underline{\tilde{P}}, S)$, then either $(\underline{\tilde{P}}, S) = (\tilde{P}, S)$ or $(\underline{\tilde{P}}, S) = (\sigma(\mathscr{L}(\Omega_S)), S)$, which is the null element for label S. This still follows from Lemma 7.

Also in this case, the properties of the domain-free information algebra $(\underline{\Phi}, I; \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$ of being atomic and atomistic carry over to this labeled version. Let $At_S(\underline{\tilde{\Phi}})$ be the set of atoms $(\tilde{P}, S)$ relative to S of the labeled information algebra $(\underline{\tilde{\Phi}}(\Omega), I; d, \cdot, \{(\sigma(\mathscr{L}(\Omega_S)), S)\}_{S\subseteq I}, \{(\sigma(\mathscr{L}^+(\Omega_S)), S)\}_{S\subseteq I}, \underline{\pi})$, and let $At_S(\underline{\tilde{P}}, S)$ be the subset of $At_S(\underline{\tilde{\Phi}})$ dominating $(\underline{\tilde{P}}, S) \in \underline{\tilde{\Phi}}$.

- *Atomic:* For any element $(\underline{\tilde{P}}, S) \in \underline{\tilde{\Phi}}$ with $\underline{\tilde{P}} \in \underline{\mathbb{P}}(\Omega_S)$, $At_S(\underline{\tilde{P}}, S)$ is not empty.
- *Atomistic*: For any element $(\underline{\tilde{P}}, S) \in \underline{\tilde{\Phi}}$, with $\underline{\tilde{P}} \in \underline{\mathbb{P}}(\Omega_S)$, we have $(\underline{\tilde{P}}, S) = \bigwedge At_S(\underline{\tilde{P}}, S)$.

3.1.3 Application: marginal problem

Information algebras provide also a general framework to treat the *marginal problem*.

The *marginal problem*, which received a long-standing interest in the literature, is the problem of checking whether some given marginal probabilistic assessments have a common joint probabilistic model [Boole, 1854; Fréchet, 1951; Vorob'ev, 1962].

To see an example, suppose we are given a few marginal probability functions over some variables: e.g., $P_1(X_1), P_2(X_2), P_3(X_1, X_3, X_4)$. We can ask ourselves whether there is a joint probability $P(X_1, X_2, X_3, X_4)$ from which we can reproduce P_1, P_2, P_3 by marginalisation.

A necessary condition for the compatibility of a number of marginal assessments is their *pairwise compatibility*, that is, the equality of the marginals over common variables. In general, however, this is not sufficient to guarantee global compatibility. Beeri et al. [1983] established a necessary and sufficient condition for pairwise compatibility to imply the global one: *the running intersection property (RIP)*, which essentially requires the existence of a total order on the marginals such that if any two marginals have variables in common, then all the marginals between them in the order contain those variables too.

An analogous problem is also studied within several other calculi of AI, such as: the theory of relational databases, possibility theory, Dempster-Shafer's theory of belief functions, etcetera [Studenỳ, 1995]. In all of these contexts, again, the running intersection property is the key factor for pairwise compatibility of less-dimensional knowledge representations being equivalent to the existence of a global one.

In [Miranda and Zaffalon, 2020] the marginal problem is treated with the instruments of desirability. Its modeling capability indeed permits a very general formulation of the problem comprehensive of most of the formalisms within which it is generally studied, see Chapter 1.1. Desirability, moreover, is not constrained by measurability issues.

In our work Casanova et al. [2022a], we reconsider the marginal problem at an even more general level. Specifically, we show that results analogous to the ones reached in Miranda and Zaffalon [2020] can be obtained with simpler proofs by using only properties of information algebras. The latter, moreover, cover other frameworks within which the marginal problem is studied that cannot be directly modeled through coherent sets of gambles, such as relational databases.

In the rest of the subsection we recall the main results found in Casanova et al. [2022a]. To easily connect the reformulation of the marginal problem given here with the one given in Miranda and Zaffalon [2020], we express them by using coherent sets of gambles. Since, however, we only use definitions and properties of domain-free information algebras, our results live at this level of generality.

We consider the same framework examined in the rest f the section, that of a multivariate model for questions. We assume therefore a possibility space Ω of the form: $\Omega := \times_{i\in I} \Omega_i$, where Ω_i is the set of possible values of a variable X_i for every $i \in I$. Using the language of information algebras, we represent partial assessments about variables $\{X_i\}_{i\in I}$ as coherent sets of gambles having support $S_i \subseteq I$. Given a family of partial assessments, our goal is to find a global model, again modeled as a coherent set of gambles, from which we can reproduce the initial models by extraction.

First of all, let us start by defining a weaker notion of compatibility for coherent sets of gambles called *consistency*.

Definition 43 (Consistency for coherent sets of gambles). *A finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ is* consistent*, or $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are* consistent*, if and only if $\mathscr{L} \neq \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$.*

As we will see later on, this property is required to extend the results of Miranda and Zaffalon [2020] to more general information algebras.

Now, we translate the definitions of pairwise compatibility and global compatibility of coherent sets of gambles given in Definition 9 and 10 of Miranda and Zaffalon, 2020 in the language of information algebras.

Definition 44 (Pairwise compatibility for coherent sets of gambles). *Two coherent sets of gambles $\mathscr{D}_i$ and $\mathscr{D}_j$, where $\mathscr{D}_i$ has support S_i and $\mathscr{D}_j$ support S_j, are called* pairwise compatible *if and only if*

$$\epsilon_{S_i\cap S_j}(\mathscr{D}_i) = \epsilon_{S_i\cap S_j}(\mathscr{D}_j). \tag{3.7}$$

Analogously, a finite family of coherent sets of gambles $\mathscr{D}_1, ..., \mathscr{D}_n$, where $\mathscr{D}_i$ has support S_i for every $i \in \{1, ..., n\}$ respectively, is pairwise compatible, or again $\mathscr{D}_1, ..., \mathscr{D}_n$ are pairwise compatible, if and only if pairs $\mathscr{D}_i, \mathscr{D}_j$ are pairwise compatible for every $i, j \in \{1, ...n\}$.

Definition 45 (Compatibility for coherent sets of gambles). *A finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$, where $\mathscr{D}_i$ has support S_i for every $i \in \{1, \ldots, n\}$ respectively, is called* compatible*, or $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are called* compatible*, if and only if there is a coherent set of gambles $\mathscr{D}$ such that $\epsilon_{S_i}(\mathscr{D}) = \mathscr{D}_i$ for $i = 1, \ldots, n$.*

Consistency and pairwise compatibility are necessary conditions for compatibility.

Lemma 8. *Consider a finite family of coherent sets of gambles $\mathscr{D}_1, ..., \mathscr{D}_n$ having supports $S_1, .., S_n$ respectively. If they are compatible, they are consistent.*

Lemma 9. *Consider a finite family of coherent sets of gambles $\mathscr{D}_1, ..., \mathscr{D}_n$ having supports $S_1, .., S_n$ respectively. If they are compatible, they are pairwise compatible.*

In general, however, these conditions are not sufficient for compatibility. In [Miranda and Zaffalon, 2020, Theorem 2], it is shown that the *running intersection property* could provide a solution for this issue.

Running intersection property (RIP) For $i = 1$ to $n-1$ there is an index $p(i)$, $i+1 \leq p(i) \leq n$ such that

$$S_i \cap S_{p(i)} = S_i \cap (\cup_{j=i+1}^{n} S_j).$$

In the context of information algebras the situation is very similar.

Theorem 14. *Consider a finite family of consistent coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ with $n > 1$ where $\mathscr{D}_i$ has support S_i for every $i \in \{1, ..., n\}$ respectively. If $S_1, \ldots, S_n$ satisfy RIP and $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are pairwise compatible, then they are compatible and $\epsilon_{S_i}(\mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n) = \mathscr{D}_i$ for $i = 1, \ldots, n$.*

We remark that from a point of view of information, compatibility of pieces of information is not always desirable. It is indeed a kind of (conditional) independence condition. Consider a family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ with supports $S_1, ..., S_n$ respectively. If $\mathscr{D}_i = \epsilon_{S_i}(\mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n)$ means that the pieces of information $\mathscr{D}_j$ for $j \neq i$ give no new information relative to variables in S_i. If, instead, the family $\mathscr{D}_1, \ldots, \mathscr{D}_n$ is not compatible, but consistent in the sense that $\mathscr{D} := \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n \neq \mathscr{L}$, then $\mathscr{D}_j$ may provide additional information on the variables in S_i for $i \neq j$ [Kohlas, 2003; Casanova et al., 2022a].

Example 9. *Consider again the framework of the previous examples. The sets of variables $S_1 := \{1\}, S_2 := \{2\}, S_3 := \{1, 3, 4\}$ satisfy the running intersection property: $S_1 \cap S_3 = S_1 \cap (S_2 \cup S_3)$. We verified above that $\mathscr{D}_1, \mathscr{D}_2^+, \mathscr{D}_3^+$ are consistent, hence they, with their support S_1, S_2, S_3, are compatible if and only if they are pairwise compatible. Let us check therefore, their pairwise compatibility.*

$\mathscr{D}_1$ and $\mathscr{D}_2^+$, with their supports S_1, S_2, are trivially pairwise compatible because they represent information on disjoint sets of variables:

$$\epsilon_{S_1 \cap S_2}(\mathscr{D}_1) = \mathscr{L}^+ = \epsilon_{S_1 \cap S_2}(\mathscr{D}_2^+).$$

The same is true for $\mathscr{D}_2^+$ and $\mathscr{D}_3^+$:

$$\epsilon_{S_2 \cap S_3}(\mathscr{D}_2^+) = \mathscr{L}^+ = \epsilon_{S_2 \cap S_3}(\mathscr{D}_3^+).$$

It remains only to check the compatibility of $\mathscr{D}_1$ and $\mathscr{D}_3^+$.

$$\epsilon_{S_1 \cap S_3}(\mathscr{D}_1) = \mathscr{D}_1,$$

$$\epsilon_{S_1 \cap S_3}(\mathscr{D}_3^+) = \{f \in \mathscr{L} : \min_{\{x_2, x_3, x_4 \in \{0,1\}} f(1, x_2, x_3, x_4) > 0\} \cup \mathscr{L}^+ \neq \epsilon_{S_1 \cap S_3}(\mathscr{D}_1).$$

They are not pairwise compatible, hence they cannot be compatible. The intuitive explanation is that $\mathscr{D}_3^+$ *gives us more information about variable* X_1 *than* $\mathscr{D}_1$*. Carol, in fact, believes the murderer* is *tall, while Alice thinks only that it is more probable that the murderer is tall than low.*

This is a desirable property in this context: Carol gives the group additional information about the height of the murderer with respect to the ones given by Alice. However, if in place of Alice we consider Dave who already thinks that the murderer is tall, i.e., whose beliefs can be modeled with the coherent set of strictly desirable gambles

$$\dot{\mathscr{D}}_1^+ := \{f \in \mathscr{L} : \min_{x_2,x_3,x_4 \in \{0,1\}} f(1,x_2,x_3,x_4) > 0\} \cup \mathscr{L}^+,$$

then $\dot{\mathscr{D}}_1^+, \mathscr{D}_2^+, \mathscr{D}_3^+$ *are consistent, pairwise compatible and hence compatible:*

$$\epsilon_{S_1 \cap S_2}(\dot{\mathscr{D}}_1^+) = \mathscr{L}^+ = \epsilon_{S_1 \cap S_2}(\mathscr{D}_2^+),$$
$$\epsilon_{S_1 \cap S_3}(\dot{\mathscr{D}}_1^+) = \dot{\mathscr{D}}_1^+ = \epsilon_{S_1 \cap S_3}(\mathscr{D}_3^+).$$

3.2 Generalised information algebras

In Kohlas [2003], sets of questions are represented as lattices of subsets of variables. In the draft work of Kohlas [2017], the structure of questions is generalised considering *quasi-separoid*. The latter are join-semilattices together with a three-place relation abstracting the notion of conditional independence, essential for local computations.[7]

An important example of quasi-separoids arise from *semilattices of partitions*, the basic model for questions assumed in our works Kohlas et al. [2021]; Casanova, Kohlas and Zaffalon [2021]; Casanova et al. [2022b]. Modeling questions as partitions allows us to be as general as Walley [1991], while avoiding to work with too abstract structures. Partitions indeed permit to work with any possibility space, not necessarily with a Cartesian product structure. This opens new possibilities. For instance, diagnostic trees and, more generally, dependent variables constrained by given relations among them can be modeled by partitions, see Shafer et al., 1987. It allows us also to differentiate between impossible events and events with zero probability, since the former can now be directly excluded from the possibility space. Moreover, it permits to avoid a proliferation of notation (see Example 10).

[7]In the literature of Bayesian networks, the underlying structures of the domains of information to be used in local computation are called *join* or *junction trees*, which determine certain structures of conditional independence. Similar structures can be established also with respect to q-separoids and allow for local computation as well [Kohlas, 2017].

Partitions lead in addition to a natural connection with *set algebras*, prototypical form of information algebras where pieces of information are represented in the simplest way as sets of answers to questions of interest, manipulable by usual set operations [Kohlas, 2003].

In what follows, we first present the definition of *generalised* domain-free information algebra as used in Kohlas et al. [2021]; Casanova, Kohlas and Zaffalon [2021]; Casanova et al. [2022b], which is equivalent to the one known in the literature from Kohlas [2017]. We formally demonstrate the equivalence of these definitions in Theorem 21 within Appendix C.2. Next, we delve into related concepts derived from the literature, with a specific focus on questions modeled as partitions. Within this context, we introduce a 'set algebra' that operates on subsets of a possibility space Ω. Finally, in Section 3.2.1 and Section 3.2.2, we focus on the application of generalised information algebras to desirability theory, showing the results we obtained on the topic in Kohlas et al. [2021]; Casanova, Kohlas and Zaffalon [2021]; Casanova et al. [2022b].

As before, we denote with $\boldsymbol{\Phi}$ the set whose elements are considered to represent pieces of information. As model for questions, we consider instead a *quasi-separoid* $(\boldsymbol{Q};\vee,\perp)$ [Kohlas, 2017, Section 2.1]. *Quasi-separoids* $(\boldsymbol{Q};\vee,\perp)$, also denoted with the term *q-separoids*, are tuples composed by a join-semilattice $(\boldsymbol{Q};\vee)$, whose elements are indicated with lower case letters $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}...$, and a three-place relation $\perp$ satisfying the following conditions:

C1 $(\forall \boldsymbol{x},\boldsymbol{y} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{y}$;

C2 $(\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z} \Rightarrow \boldsymbol{y} \perp \boldsymbol{x} | \boldsymbol{z}$;

C3 $(\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{w},\boldsymbol{z} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z}$ and $\boldsymbol{w} \leq \boldsymbol{y} \Rightarrow \boldsymbol{x} \perp \boldsymbol{w} | \boldsymbol{z}$;

C4 $(\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z}$ implies $\boldsymbol{x} \perp \boldsymbol{y} \vee \boldsymbol{z} | \boldsymbol{z}$.

Usually, in literature, two additional conditions are assumed [Dawid, 2001]:

C5 $(\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{w},\boldsymbol{z} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z}$ and $\boldsymbol{w} \leq \boldsymbol{y} \Rightarrow \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z} \vee \boldsymbol{w}$;

C6 $(\forall \boldsymbol{x},\boldsymbol{y},\boldsymbol{w},\boldsymbol{z} \in \boldsymbol{Q})\ \boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z}$ and $\boldsymbol{x} \perp \boldsymbol{w} | \boldsymbol{y} \vee \boldsymbol{z} \Rightarrow \boldsymbol{x} \perp \boldsymbol{y} \vee \boldsymbol{w} | \boldsymbol{z}$.

In this case we obtain a *separoid* $(\boldsymbol{Q};\vee,\perp)$. Separoids are often identified as the fundamental mathematical structure underlying different concepts of conditional independence developed both in probabilistic contexts and logical ones [Dawid, 1979; Spohn, 1980; Dawid, 2001]. They express the idea that: if $\boldsymbol{x} \perp \boldsymbol{y} | \boldsymbol{z}$, only the part relative to $\boldsymbol{z}$ of an information relative to $\boldsymbol{x}$ is relevant as an information relative to $\boldsymbol{y}$, and vice versa.

In Kohlas [2017], systems of questions are assumed to only form q-separoids, with the same interpretation. The weaker notion of conditional independence they subsume, indeed, is sufficient to allow for local computations in this more general context [Kohlas, 2017]. Join semilattices are instead required to be able to compare questions by their granularity and, given two questions $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{Q}$, to work with their combined question $\boldsymbol{x} \vee \boldsymbol{y}$.

Now, we are ready to give our formal definition of *generalised* domain-free information algebras, as employed in Kohlas et al. [2021]; Casanova, Kohlas and Zaffalon [2021]; Casanova et al. [2022b].

Definition 46 (**Generalised domain-free information algebra**)**.** *A* generalised domain-free information algebra *is a two-sorted structure* $(\boldsymbol{\Phi}, \boldsymbol{Q}; \vee, \perp, \cdot, \boldsymbol{0}, \boldsymbol{1}, \epsilon)$,[8] *where:*

- $(\boldsymbol{\Phi}; \cdot, \boldsymbol{0}, \boldsymbol{1})$ *is a commutative semigroup with* $\cdot : \boldsymbol{\Phi} \times \boldsymbol{\Phi} \to \boldsymbol{\Phi}$ *defined by* $\boldsymbol{\phi}, \boldsymbol{\psi} \mapsto \boldsymbol{\phi} \cdot \boldsymbol{\psi}$*, and with* $\boldsymbol{0}$ *and* $\boldsymbol{1}$ *as its null and unit elements respectively,*

- $(\boldsymbol{Q}; \vee, \perp)$ *is a q-separoid,*

- $\epsilon : \boldsymbol{\Phi} \times \boldsymbol{Q} \to \boldsymbol{\Phi}$ *defined by* $\boldsymbol{\phi}, \boldsymbol{x} \mapsto \epsilon_{\boldsymbol{x}}(\boldsymbol{\phi})$*,*

satisfying moreover the following properties:

1. *Nullity: for any* $\boldsymbol{x} \in \boldsymbol{Q}$*,*
$$\epsilon_{\boldsymbol{x}}(\boldsymbol{0}) = \boldsymbol{0};$$

2. *Idempotency: for any* $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ *and* $\boldsymbol{x} \in \boldsymbol{Q}$*,*
$$\epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \boldsymbol{\phi} = \boldsymbol{\phi};$$

3. *Combination: for any* $\boldsymbol{\phi}, \boldsymbol{\psi}$ *and* $\boldsymbol{x} \in \boldsymbol{Q}$*,*
$$\epsilon_{\boldsymbol{x}}(\epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \boldsymbol{\psi}) = \epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \epsilon_{\boldsymbol{x}}(\boldsymbol{\psi});$$

4. *Extraction: for any* $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ *and* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{Q}$*, such that* $\epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) = \boldsymbol{\phi}$ *and* $\boldsymbol{x} \vee \boldsymbol{z} \perp \boldsymbol{y} \vee \boldsymbol{z} | \boldsymbol{z}$*,*
$$\epsilon_{\boldsymbol{y} \vee \boldsymbol{z}}(\boldsymbol{\phi}) = \epsilon_{\boldsymbol{y} \vee \boldsymbol{z}}(\epsilon_{\boldsymbol{z}}(\boldsymbol{\phi}));$$

5. *Support: For any* $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ *there is an* $\boldsymbol{x} \in \boldsymbol{Q}$ *so that* $\epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) = \boldsymbol{\phi}$*, i.e., a* support *of* $\boldsymbol{\phi}$*, and whenever* $\epsilon_{\boldsymbol{x}}(\boldsymbol{\phi}) = \boldsymbol{\phi}$*, then* $\epsilon_{\boldsymbol{y}}(\boldsymbol{\phi}) = \boldsymbol{\phi}$ *for every* $\boldsymbol{y} \geq \boldsymbol{x}$*,* $\boldsymbol{y} \in \boldsymbol{Q}$*.*

[8] See [Hodges, 1993, Section 1.1].

Note that, by Idempotency, we also have $\boldsymbol{\epsilon_x(1) = \epsilon_x(1) \cdot 1 = 1}$.

The intuition behind this definition is similar to the one given in Section 3.1 for a domain-free information algebra and the operations $\cdot$, ϵ are called with the same names. Dissimilarities are due to the different model for questions assumed.

This definition is in particular equivalent to the one given in Section 5.2 of Kohlas, 2017, see Theorem 21 in Appendix C.2. We use this alternative formulation to be closer to the one given in Section 3.1.

Definitions of *homomorphism*, *sub-structure*, *information order*, *atom*, *atomic* and *atomistic* (generalised) domain-free information algebras can be given also in this context, analogous to the ones of Section 3.1, see Kohlas, 2017, Section 5, Section 6.

A useful example of q-separoids is given by a semilattice of partitions with a suitable conditional independence relation [Kohlas, 2017, Section 2.2].

Consider a possibility space Ω. Questions about Ω can directly be modeled by listing their possible answers, i.e., they can be modeled as partitions $\mathscr{P}_x$ of Ω where x is in some index set Q, whose blocks represent worlds $\omega \in \Omega$ where $\mathscr{P}_x$ has the same answer. Partitions can also be equivalently expressed with equivalence relations $\equiv_x$ on Ω: given $\omega, \omega' \in \Omega$, $\omega \equiv_x \omega'$ if and only if ω, ω' are in the same block of partition $\mathscr{P}_x$. For simplicity, in what follows, we identify partitions and equivalence relations with their indexes $x \in Q$. In the special case where the questions of interest cover all the possible partitions of Ω, we denote Q as U. Q and U depend of course on Ω. But we do not emphasize this dependency, because it will always be clear from the context to which domain Ω they refer.

Questions modeled in this way can be ordered with respect to granularity: $x \leq y$ if and only if y is finer than x, i.e., for every $\omega, \omega' \in \Omega$, $\omega \equiv_y \omega'$ implies $\omega \equiv_x \omega'$ or, equivalently, block $[\omega]_y$ of partition $\mathscr{P}_y$ is contained in some block $[\omega]_x$ of partition $\mathscr{P}_x$.[9]

The set of all the partitions of Ω, $\{\mathscr{P}_x\}_{x \in U}$, with this order relation induces a lattice [Grätzer, 2003, Section 4]. In this lattice, the join of two partitions $\mathscr{P}_x, \mathscr{P}_y$ corresponds to the partition obtained as the non-empty intersections of blocks of $\mathscr{P}_x$ with blocks of $\mathscr{P}_y$. We indicate it with $x \vee y$. Hence, $(U; \vee)$, where $\vee$ is the join operation defined above, is a join-semilattice. The definition of meet instead is somewhat more involved.

We can also define a conditional independence relation on partitions that, for $n = 2$, satisfies properties C1–C4. For a proof of this fact see Kohlas, 2017,

[9] In the literature usually the inverse order between partitions is considered. However, this order better corresponds to our natural order of questions by granularity.

Theorem 2.6.

Definition 47 (**Conditionally Independent Partitions**). *Consider a finite set of partitions of Ω, $\mathscr{P}_1, \ldots \mathscr{P}_n$, and a block B of another partition $\mathscr{P}$ of Ω (contained or not in the list $\mathscr{P}_1, \ldots, \mathscr{P}_n$), then define for $n \geq 1$,*

$$R_B(\mathscr{P}_1, \ldots, \mathscr{P}_n) := \{(B_1, \ldots, B_n) : B_i \in \mathscr{P}_i, \cap_{i=1}^n B_i \cap B \neq \emptyset\}.$$

We call $\mathscr{P}_1, \ldots, \mathscr{P}_n$ conditionally independent given *$\mathscr{P}$, if and only if for all blocks B of $\mathscr{P}$, $R_B(\mathscr{P}_1, \ldots, \mathscr{P}_n) = R_B(\mathscr{P}_1) \times \cdots \times R_B(\mathscr{P}_n)$.*

$\mathscr{P}_x, \mathscr{P}_y$ are conditionally independent given $\mathscr{P}_z$ if and only if, for every answer B_z to the question $\mathscr{P}_z$, knowing also an answer to $\mathscr{P}_x$ (or $\mathscr{P}_y$) compatible with B_z, does not give us additional information regarding the answer to $\mathscr{P}_y$ (respectively $\mathscr{P}_x$), except that it must be again compatible with B_z. This is true if and only if for every $B_x \in \mathscr{P}_x, B_y \in \mathscr{P}_y, B_z \in \mathscr{P}_z$ such that $B_x \cap B_z \neq \emptyset$ and $B_y \cap B_z \neq \emptyset$, we have $B_x \cap B_y \cap B_z \neq \emptyset$. In this case we write $x \bot y | z$. We may also say that $x \bot y | z$ if and only if $\omega \equiv_z \omega'$ implies the existence of an element $\omega'' \in \Omega$ such that $\omega \equiv_{x \vee z} \omega''$ and $\omega' \equiv_{y \vee z} \omega''$. Analogous considerations can be made for more than two partitions conditionally independent given a third one.

Given the above considerations, we can conclude that $(U; \vee)$ with the conditional independence relation $\bot$ defined above forms a q-separoid $(U; \vee, \bot)$. It is possible to show that also every join-subsemilattice $(Q; \vee)$ of $(U; \vee)$ induces a q-separoid $(Q; \vee, \bot)$, where $\bot$ corresponds again to the conditional independence relation among partitions [Kohlas, 2017, Theorem 2.6]. In this case, for simplicity, we call $(Q; \vee, \bot)$ a *sub-q-separoid* of $(U; \vee, \bot)$.

Multivariate models form a particular instance of q-separoids of partitions.

Let us suppose that $\Omega = \bigtimes_{i \in I} \Omega_i$ for some non-empty index set I. Let us consider now partitions $\mathscr{P}_S$ of Ω with $S \subseteq I$ such that ω, ω' are in the same block (resp. $\omega \equiv_S \omega'$) if and only if $\omega | S = \omega' | S$. They induce a q-separoid. Identifying partitions with their indexes indeed, we have that $(\mathrm{P}(I); \vee, \bot)$ is a q-separoid, where $\vee$ and $\bot$ are respectively the join and the conditional independence relation among partitions defined above, see [Kohlas, 2017, Chapter 2.2]. In particular, $(\mathrm{P}(I); \vee)$ induces a lattice $(\mathrm{P}(I); \vee, \wedge)$, where $S \vee T = S \cup T$ and $S \wedge T := S \cap T$, for every $S, T \subseteq I$, see again [Kohlas, 2017, Chapter 2.2]. Moreover, we have also that $S \bot T | R$ if and only if $(S \cup R) \cap (T \cup R) = R$ or, equivalently, $S \cap T \subseteq R$ [Kohlas, 2017, Chapter 2.2]. These properties imply that a generalised domain-free information algebra having a multivariate model for questions satisfies also the following axiom [Kohlas, 2017, Lemma 5.3]:

Commutative Extraction: For any $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ and $S, T \subseteq I$,

$$\boldsymbol{\epsilon}_T(\boldsymbol{\epsilon}_S(\boldsymbol{\phi})) = \boldsymbol{\epsilon}_{S \cap T}(\boldsymbol{\phi}) = \boldsymbol{\epsilon}_S(\boldsymbol{\epsilon}_T(\boldsymbol{\phi})). \tag{3.8}$$

The resulting axiomatic framework obtained incorporating this latter axiom to the other specified above corresponds to the one introduced in Kohlas [2003], employed in our work Casanova et al. [2022a] and recalled in Section 3.1.

Archetypes of information algebras are so-called *set algebras*. They are generalised information algebras where pieces of information are modeled in the simplest way as subsets of some universe, combination is set intersection, and extraction is related to the so-called *saturation operators* [Kohlas, 2017, Section 2.2].

In our works Casanova, Kohlas and Zaffalon [2021]; Casanova et al. [2022b] we showed, in particular, how a set algebra where pieces of information are subsets of a possibility space Ω and questions of interest are represented as partitions can be embedded into the generalised domain-free information algebras induced by desirability. Set algebras can be regarded as the algebraic counterparts of classical propositional logic, thus this embedding can be interpreted as another way to show that latter is formally part of the theory of imprecise probabilities, see Section 1.1.3.

In this section, we present the set algebra in question. The relationship between this algebra and the generalized information algebras deriving from desirability will be revisited in the following subsection.

As above, let us consider a generic possibility space Ω. For any partition $\mathscr{P}_x$ of Ω, or equivalently for every $x \in U$, it is possible to construct a *saturation operator* σ_x defined as:

$$(\forall S \subseteq \Omega)\, \sigma_x(S) := \{\omega \in \Omega : (\exists \omega' \in S)\, \omega \equiv_x \omega'\}, \tag{3.9}$$

which can be rewritten also as

$$(\forall S \subseteq \Omega)\, \sigma_x(S) = \cup\{B_x : B_x \text{ block of } \mathscr{P}_x,\ B_x \cap S \neq \emptyset\}.$$

Saturation operators satisfy the properties summarised in the next result, similar to Kohlas, 2017, Lemma 2.1.[10]

Lemma 10. *For any $S, T \subseteq \Omega$ and any partition $\mathscr{P}_x$ with $x \in U$, we have:*

1. $\sigma_x(\emptyset) = \emptyset$,
2. $S \subseteq \sigma_x(S)$,
3. $\sigma_x(\sigma_x(S) \cap T) = \sigma_x(S) \cap \sigma_x(T)$,

[10] I express gratitude to a reviewer of this thesis for having pointed out that a saturation operator can be interpreted as a closure operator associated with the topology induced by its related partition [Steen et al., 1978, p.43]. Consequently, it naturally inherits all the properties of a topological closure operator. Notably, these properties also encompass those of a closure operator as defined in Section 1.1.3, which correspond to items 2, 4, and 5 of Lemma 10.

4. $\sigma_x(\sigma_x(S)) = \sigma_x(S)$,

5. $S \subseteq T \Rightarrow \sigma_x(S) \subseteq \sigma_x(T)$,

6. $\sigma_x(\sigma_x(S) \cap \sigma_x(T)) = \sigma_x(S) \cap \sigma_x(T)$.

Now, let us consider the following tuple:

$$(P_Q(\Omega), Q; \vee, \perp, \cap, \emptyset, \Omega, \sigma),$$

where:

- $(Q; \vee, \perp)$ is a sub-q-separoid of $(U; \vee, \perp)$;
- $P_Q(\Omega) := \{S \subseteq \Omega : \exists\, x \in Q,\ \sigma_x(S) = S\}$ is the set of subsets of Ω *saturated with respect to* Q;
- $\sigma : P_Q(\Omega) \times Q \to P_Q(\Omega)$ is defined by $S, x \mapsto \sigma_x(S)$, where σ_x is defined in turn in Eq. (3.9).

We claim it is a set algebra where $P_Q(\Omega)$ is the set of pieces of information, $(Q; \vee, \perp)$ is the q-separoid of questions of interest, $\cap$ is the combination operation (with $\emptyset$ and Ω as the null and the unit elements respectively) and σ, defined above, is the extraction operation.

We can prove the previous claim step by step as follows:

- $(Q; \vee, \perp)$ is a q-separoid by hypothesis.
- The Support axiom is satisfied. Indeed, by definition of $P_Q(\Omega)$, for every $S \in P_Q(\Omega)$ there exists $x \in Q$, such that $\sigma_x(S) = S$. Moreover, if $x \leq y$ and $y \in Q$, then $\omega \equiv_y \omega'$ implies $\omega \equiv_x \omega'$, so that $\sigma_y(S) \subseteq \sigma_x(S)$. Therefore, if $\sigma_x(S) = S$, then $S \subseteq \sigma_y(S) \subseteq \sigma_x(S) = S$.
- $(P_Q(\Omega); \cap, \emptyset, \Omega)$ is a commutative semigroup with $\emptyset$ as the null element and Ω as the unit one. Clearly, $\emptyset, \Omega \in P_Q(\Omega)$ and they are the null and the unit element respectively, with respect to set intersection. Therefore, the only property left to prove is that $P_Q(\Omega)$ is closed with respect to intersection. This is true. Consider indeed two sets $S, T \in P_Q(\Omega)$, such that $\sigma_x(S) = S$ and $\sigma_y(T) = T$ for some $x, y \in Q$, respectively. Then, thanks to the Support axiom, we know that $\sigma_{x \vee y}(S) = S$ and $\sigma_{x \vee y}(T) = T$ with $x \vee y \in Q$, since $(Q; \vee)$ is a join-semilattice. Therefore, thanks to Lemma 10, we have
$$\sigma_{x \vee y}(S \cap T) = \sigma_{x \vee y}(\sigma_{x \vee y}(S) \cap \sigma_{x \vee y}(T)) = \sigma_{x \vee y}(S) \cap \sigma_{x \vee y}(T) = S \cap T.$$

- $\sigma_x(S) \in P_Q(\Omega)$ for every $S \in P_Q(\Omega)$, $x \in Q$, by item 4 of Lemma 10.
- Nullity, Idempotency and Combination axioms follow from items 1–3 of Lemma 10.

It remains only to prove the Extraction axiom but it follows from the theorem below.

Theorem 15. *Given a sub-q-separoid $(Q; \vee, \perp)$ of $(U; \vee, \perp)$, consider $x, y, z \in Q$, such that $x \vee z \perp y \vee z | z$. Then, for any $S \subseteq \Omega$,*

$$\sigma_{y\vee z}(\sigma_x(S)) = \sigma_{y\vee z}(\sigma_z(\sigma_x(S))).$$

3.2.1 Generalised information algebras of coherent sets of gambles and coherent lower previsions

Let us consider again, as in Section 3.1.1, a possibility space Ω and the set $\Phi(\Omega) := \mathbb{D}(\Omega) \cup \{\mathcal{L}(\Omega)\}$. Here, however, we do not put any restriction on the form of the possibility space. Let us consider also a sub-q-separoid $(Q; \vee, \perp)$ of the set of all partitions of Ω, $(U; \vee, \perp)$, such that for ever $\mathcal{D}$ there exists $x \in Q$ so that $\epsilon_x(\mathcal{D}) = \mathcal{D}$, where ϵ_x is defined in Eq.(3.10).[11] Let us then consider these two operations on Φ and Q, analogous to the ones defined in Section 3.1.1:

- Combination: $\cdot : \Phi \times \Phi \to \Phi$, defined by
$$\mathcal{D}_1, \mathcal{D}_2 \mapsto \mathcal{D}_1 \cdot \mathcal{D}_2 := \mathcal{C}(\mathcal{D}_1 \cup \mathcal{D}_2),$$
where $\mathcal{C}$ operator is defined in Eq. (1.4),
- Extraction. $\epsilon : \Phi \times Q \to \Phi$, defined by
$$\mathcal{D}, x \mapsto \epsilon_x(\mathcal{D}) := \mathcal{C}(\mathcal{D} \cap \mathcal{L}_x), \tag{3.10}$$
where $\mathcal{L}_x$ is the set of all $\mathcal{P}_x$-measurable gambles.

The following theorem shows that $(\Phi(\Omega), Q; \vee, \perp, \cdot, \mathcal{L}(\Omega), \mathcal{L}^+(\Omega), \epsilon)$ or, for short, $(\Phi, Q; \vee, \perp, \cdot, \mathcal{L}, \mathcal{L}^+, \epsilon)$, where $\cdot$, ϵ are defined as above on Φ and Q is a generalised domain-free information algebra.

Theorem 16. *1. $(\Phi; \cdot, 0, 1)$ is a commutative semigroup with a null element $0 = \mathcal{L}$ and a unit element $1 = \mathcal{L}^+$.*

2. For any $x \in Q$, $\epsilon_x(0) = 0$.

[11] Notice that the finest top partition of Ω (all blocks consisting of exactly one element $\omega \in \Omega$) is a support of all the sets $\mathcal{D} \in \Phi$. Therefore, any sub-q-separoid $(Q; \vee, \perp)$ of $(U; \vee, \perp)$ containing this partition satisfies the requirement.

3. *For any* $\mathscr{D} \in \Phi$ *and any* $x \in Q$, $\epsilon_x(\mathscr{D}) \cdot \mathscr{D} = \mathscr{D}$.

4. *For any* $\mathscr{D}_1, \mathscr{D}_2 \in \Phi$ *and any* $x \in Q$, $\epsilon_x(\epsilon_x(\mathscr{D}_1) \cdot \mathscr{D}_2) = \epsilon_x(\mathscr{D}_1) \cdot \epsilon_x(\mathscr{D}_2)$.

5. *For any* $\mathscr{D} \in \Phi$ *and* $x, y, z \in Q$, *if* $\epsilon_x(\mathscr{D}) = \mathscr{D}$ *and* $x \vee z \perp y \vee z | z$, *then:*
$$\epsilon_{y \vee z}(\mathscr{D}) = \epsilon_{y \vee z}(\epsilon_z(\mathscr{D})).$$

6. *Given* $\mathscr{D} \in \Phi$ *and* $x \in Q$. *If* x *is a support of* $\mathscr{D}$, *then any* $y \geq x$, $y \in Q$ *is also a support of* $\mathscr{D}$.

In what follows, we refer to $(\Phi, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ as the *generalised domain-free information algebra of coherent sets of gambles*. Since the definitions of combination, information order and atoms are analogous to the ones given for its corresponding domain-free information algebra, it is possible to prove that $(\Phi, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ is also atomistic with maximal coherent sets as atoms.

Similarly to Section 3.1.2, see Theorem 22 in Appendix C.2, we can show that $(\underline{\Phi}, Q; \vee, \perp, \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$, with $\underline{\Phi}(\Omega) := \underline{\mathbb{P}}(\Omega) \cup \{\sigma(\mathscr{L}(\Omega))\}$, $(Q; \vee, \perp)$ a q-separoid satisfying the same requirement established for sets of gambles, and combination and extraction defined analogously to the ones introduced in Section 3.1.2:

1. Combination. $\cdot : \underline{\Phi} \times \underline{\Phi} \to \underline{\Phi}$, defined by
$$\underline{P}_1, \underline{P}_2 \mapsto \underline{P}_1 \cdot \underline{P}_2 := \underline{E}^*(\max\{\underline{P}_1, \underline{P}_2\}), \tag{3.11}$$
where $\underline{E}^*$ is defined in Definition 17.

2. Extraction. $\underline{e} : \underline{\Phi} \times Q \to \underline{\Phi}$, defined by
$$\underline{P}, x \mapsto \underline{e}_x(\underline{P}) := \underline{E}^*(\underline{P}_x),$$
where $\underline{P}_x$ is the $\mathscr{P}_x$-marginal of $\underline{P}$,

is a generalised domain-free information algebra. Similar to before it is then isomorphic to $(\Phi^+, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$, sub-structure of $(\Phi, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ that can be similarly constructed with Φ^+ in place of Φ, and it is atomistic with linear previsions as atoms.

We now link these generalised domain-free information algebras to set algebras. In particular, let us consider $(P_Q(\Omega), Q; \vee, \perp, \cap, \emptyset, \Omega, \sigma)$ defined as in Section 3.2 with the same q-separoid of questions $(Q; \vee, \perp)$ of $(\Phi^+, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$. We show that the former can formally be embedded in the latter, hence it can also be embedded in $(\Phi, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$ and $(\underline{\Phi}, Q; \vee, \perp, \cdot, \sigma(\mathscr{L}), \sigma(\mathscr{L}^+), \underline{e})$.

For any set $S \in P_Q(\Omega)$ indeed, we can define
$$\mathscr{D}_S^+ := \{f \in \mathscr{L}(\Omega) : \inf_{\omega \in S} f(\omega) > 0\} \cup \mathscr{L}^+(\Omega). \tag{3.12}$$

If $S \neq \emptyset$, it is a coherent set of strictly desirable gambles,[12] otherwise it corresponds to $\mathscr{L}(\Omega)$. The map $h_{\mathscr{D}} : S \mapsto \mathscr{D}_S^+$ is then a homomorphism between $(P_Q(\Omega), Q; \vee, \perp, \cap, \emptyset, \Omega, \sigma)$ and $(\Phi^+, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$.

Theorem 17. *Let $S, T \in P_Q(\Omega)$ and $x \in Q$. Then*

1. $\mathscr{D}_S^+ \cdot \mathscr{D}_T^+ = \mathscr{D}_{S\cap T}^+$,
2. $\mathscr{D}_\emptyset^+ = \mathscr{L}(\Omega)$, $\mathscr{D}_\Omega^+ = \mathscr{L}^+(\Omega)$,
3. $\epsilon_x(\mathscr{D}_S^+) = \mathscr{D}_{\sigma_x(S)}^+$.

$h_{\mathscr{D}}$, moreover, is in particular injective, indeed if $\mathscr{D}_S^+ = \mathscr{D}_T^+$, then $S = T$. So $(P_Q(\Omega), Q; \vee, \perp, \cap, \emptyset, \Omega, \sigma)$ is also embedded into $(\Phi^+, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$.

Example 10. *Let us consider again the framework of Example 5. Suppose now that the list of the suspected people boils down to the following individuals:*

- *subject a: low, blond short hair, motive of passion;*
- *subject b: tall, black long hair, motive of passion;*
- *subject c: low, blond long hair, economic motive;*
- *subject d: tall, black short hair, economic motive;*
- *subject e: tall, black short hair, motive of passion.*

Let us construct then a new possibility space composed by the suspected people: $\Omega := \{a, b, c, d, e\}$. Beliefs of the detectives Alice, Bob and Carol can be translated on these suspected people and can be modeled respectively with the following coherent sets of gambles on:

- $$\mathscr{D}_1' := \text{posi}(\{\mathbb{I}_{\{b,d,e\}} - \mathbb{I}_{\{a,c\}}\} \cup \mathscr{L}^+),$$
  which is associated to the coherent lower prevision:
  $$(\forall f \in \mathscr{L})\ \underline{P}_1'(f) := \inf\{P(f) : P \in \mathbb{P},\ P(\{b,d,e\}) \geq P(\{a,c\})\}.$$
- $$(\mathscr{D}_2')^+ := \{f \in \mathscr{L} : \min_{\omega \in \{a,b,e\}} f(\omega) > 0\} \cup \mathscr{L}^+,$$
  which is equivalent to the coherent lower prevision:
  $$(\forall f \in \mathscr{L})\ \underline{P}_2'(f) := \min_{\omega \in \{a,b,e\}} f(\omega).$$

[12] Indeed, $\underline{P}(f) := \inf_S(f)$ for every $f \in \mathscr{L}$ with $S \neq \emptyset$, is a coherent lower prevision [Troffaes and de Cooman, 2014, Section 5.4].

- $$(\mathscr{D}_3')^+ := \{f \in \mathscr{L} : f(b) > 0\} \cup \mathscr{L}^+,$$

  which is equivalent to the linear prevision

  $$(\forall f \in \mathscr{L})\, P_3'(f) := f(b).$$

Notice that we could have represented subjects using the same framework of Example 5, identifying every suspect with his characteristics (height, motive, hair colour and haircut). However, this would have caused a proliferation of notation since some of the possible combination of the characteristics considered do not have a correspondence with the suspects (e.g., there is no short suspect with black short hair).

In this context therefore, we can proceed in a natural way to represent characteristics of interest of the suspects, i.e., the questions of interest, by using partitions:

$$\mathscr{P}_{height} := \{\{a, c\}, \{b, d, e\}\},$$
$$\mathscr{P}_{motive} := \{\{a, b, e\}, \{c, d\}\},$$
$$\mathscr{P}_{hair\ colour} := \{\{a, c\}, \{b, d, e\}\},$$
$$\mathscr{P}_{hair\ cut} := \{\{a, d, e\}, \{b, c\}\}.$$

Similarly to before, we can then use instruments given by generalised (domain-free) information algebras to manipulate the information collected by detectives.

In this case, for example, merging together the detectives' beliefs we find Carols' ones, now represented directly through a maximal coherent set of strictly desirable gambles

$$\mathscr{D}' := \mathscr{D}_1' \cdot (\mathscr{D}_2')^+ \cdot (\mathscr{D}_3')^+ := \mathscr{C}(\mathscr{D}_1' \cup (\mathscr{D}_2')^+ \cup (\mathscr{D}_3')^+) = (\mathscr{D}_3')^+.$$

We can then extract the information about the murder's motive. Also in this case, we find Bob's information about the murderer.

$$\epsilon_{motive}(\mathscr{D}') = \epsilon_{motive}((\mathscr{D}_3')^+) := \mathscr{C}((\mathscr{D}_3')^+ \cap \mathscr{L}_{motive}) = (\mathscr{D}_2')^+.$$

Analogous operations can be performed expressing beliefs with coherent lower previsions.

3.2.2 Application: marginal problem

Here, we generalize the discussion about the marginal problem started in Section 3.1.3 to more general questions. To do so, we use tools provided by the generalised domain-free information algebra of coherent sets of gambles $(\Phi, Q; \vee, \perp, \cdot, \mathscr{L}, \mathscr{L}^+, \epsilon)$. Our results, however, are valid for every generalised domain-free information algebra, similarly to Section 3.1.3.

In this context, variables of interest are modeled through the questions they represent. Assessments about different questions are instead modeled through coherent sets of gambles having different supports.

We start by re-defining the concepts of consistency, pairwise compatibility and compatibility in this context.

Definition 48 (**Consistency for coherent sets of gambles**). *A finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ is* consistent*, or $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are* consistent*, if and only if $\mathscr{L} \neq \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$.*

Definition 49 (**Pairwise compatibility for coherent sets of gambles**). *Two coherent sets of gambles $\mathscr{D}_i$ and $\mathscr{D}_j$, where $\mathscr{D}_i$ has support x_i and $\mathscr{D}_j$ support x_j, are called* pairwise compatible *if and only if*

$$\epsilon_{x_i}(\mathscr{D}_i \cdot \mathscr{D}_j) = \mathscr{D}_i,$$
$$\epsilon_{x_j}(\mathscr{D}_i \cdot \mathscr{D}_j) = \mathscr{D}_j.$$

Analogously, a finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$, where $\mathscr{D}_i$ has support x_i for every $i \in \{1, \ldots, n\}$ respectively, is pairwise compatible, or again $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are pairwise compatible, if and only if pairs $\mathscr{D}_i, \mathscr{D}_j$ are pairwise compatible for every $i, j \in \{1, \ldots n\}$.

Definition 50 (**Compatibility for coherent sets of gambles**). *A finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$, where $\mathscr{D}_i$ has support x_i for every $i \in \{1, \ldots, n\}$ respectively, is called* compatible*, or $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are called* compatible*, if and only if there is a coherent set of gambles $\mathscr{D}$ such that $\epsilon_{x_i}(\mathscr{D}) = \mathscr{D}_i$ for $i = 1, \ldots, n$.*

The different definition of pairwise compatibility follows from the fact that Q does not necessarily induce a lattice. For multivariate models for questions, however, this definition collapses to the one given in Section 3.1.3 (Theorem 23 in Appendix C.2).

As before, we can prove that both consistency and pairwise compatibility are necessary conditions for compatibility.

Lemma 11. *Consider a finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ having supports $x_1, .., x_n$ respectively. If they are compatible, they are consistent.*

Lemma 12. *Consider a finite family of coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ having supports $x_1, .., x_n$ respectively. If they are compatible, they are pairwise compatible.*

Pairwise compatibility is not sufficient to guarantee compatibility. However, up to consistency, it is sufficient if supports of the sets involved form an *hypertree* [Kohlas, 2017, Definition 4.3], a generalisation of the RIP property defined in Section 3.1.3.

Definition 51 (Hypertree). *Let $(\mathbf{Q}; \vee, \perp)$ be a q-separoid. A n-elements subset $\boldsymbol{S}$ of $\mathbf{Q}$ is called a* hypertree *if there is a numbering of its elements $\boldsymbol{S} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$ such that for any $i \in \{1, ..., n-1\}$ there is an element $b(i) > i$ in the numbering so that*

$$\boldsymbol{x}_i \perp \bigvee_{j=i+1}^{n} \boldsymbol{x}_j | \boldsymbol{x}_{b(i)}.$$

In this case we say that elements of $\boldsymbol{S}$ *form a hypertree*.

Theorem 18. *Consider a finite family of consistent coherent sets of gambles $\mathscr{D}_1, \ldots, \mathscr{D}_n$ with $n > 1$ where $\mathscr{D}_i$ has support x_i for every $i \in \{1, ..., n\}$ respectively. If $x_1, \ldots, x_n$ form a hypertree and $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are pairwise compatible, then they are compatible and $\epsilon_{x_i}(\mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n) = \mathscr{D}_i$ for $i = 1, \ldots, n$.*

To conclude, it is possible to notice that, if we assume a multivariate model for questions, a set of supports is a hypertree if and only if it satisfy the running intersection property. Therefore, Theorem 14 can be seen as a special case of Theorem 18.

Example 11. *Let us consider the framework of Example 10. It is possible to notice that:*

- $\epsilon_{height}(\mathscr{D}'_1) = \mathscr{D}'_1$;
- $\epsilon_{motive}((\mathscr{D}')^+_2) = (\mathscr{D}')^+_2$;
- $\epsilon_{haircolour \vee haircut}((\mathscr{D}')^+_3) = (\mathscr{D}')^+_3$.

In particular,

$$height \perp motive \vee (haircolour \ \vee \ haircut) \,|(haircolour \ \vee \ haircut)$$

therefore, height, motive, haircolour $\vee$ haircut form an hypertree. So, since they are consistent, $\mathscr{D}'_1, (\mathscr{D}')^+_2, (\mathscr{D}')^+_2$ are compatible if and only if they are pairwise compatible. However, as for the multivariate case, they are not pairwise compatible, indeed the information given by $(\mathscr{D}')^+_3$ about the murder's height is still more informative than the one given by $\mathscr{D}'_1$:

$$\epsilon_{height}(\mathscr{D}'_1 \cdot (\mathscr{D}')^+_3) = \mathscr{D}'_1 \cdot \epsilon_{height}((\mathscr{D}')^+_3) = \{f \in \mathscr{L} : \min_{\omega \in \{b,d,e\}} f(\omega) > 0\} \cup \mathscr{L}^+ \neq \mathscr{D}'_1.$$

If we consider however, analogously to Example 9, Dave's beliefs:

$$(\dot{\mathscr{D}}'_1)^+ := \{f \in \mathscr{L} : \min_{\omega \in \{b,d,e\}} f(\omega) > 0\} \cup \mathscr{L}^+,$$

we reach compatibility.

3.3 Summary

In this chapter, we create a bridge between desirability and *information algebras*, general algebraic structures to manage information providing basic operations and architectures for inference.

Information algebras abstract away many features that can be found in nearly every formalism for representing information. Typically indeed, information comes in *pieces* that refer to different *questions* or problems of interest. These components can be combined together to obtain more information. Inference then usually means to extract from the whole of the knowledge the part relevant to a given problem or question. This leads to an algebraic structure composed by a set of 'pieces of information' and a set of questions, manipulable by two basic operations: *combination*, to aggregate the pieces of information, and *extraction*, to extract from a piece of information the part related to a specific question. Axioms required to these operations permit, in particular, a very general formulations of architectures to support efficient inference.

In Kohlas, 2003, questions are initially assumed to regard logically independent variables, the so-called *multivariate model* for questions. This model is then generalized in Kohlas [2017] considering different types of questions. Information algebras moreover, could be expressed in two different, but equivalent, forms: the *labeled* one, where pieces of information are explicitly linked to questions they refer and therefore they permit to limit memory requirements to what is needed, and the *domain-free* one, where instead they are treated as more abstract entities more suitable for theoretical considerations. Another characteristic of information algebras is that they can be also expressed through *information systems*, tuples formed by a *language*, a *closure operator* and a family of *sublanguages*. Starting from this observation, we establish a connection with desirability. More precisely, in Section 3.1.1 and Section 3.1.2, we demonstrate that when we confine our focus to questions that can be expressed using a multivariate model, it becomes possible to incorporate coherent sets of gambles and coherent lower previsions into the structure of both a domain-free and a labeled information algebra. Then, in Section 3.1.3, we apply these results to study the *marginal problem*, simplifying and even expanding the treatment of the same problem found in the existing desirability literature. In Section 3.2, we broaden the scope of our findings to encompass scenarios involving more general questions. Within this context, we also establish that a particular *set algebra* of subsets of the possibility space where gambles are defined can be embedded into the algebras of coherent sets of gambles and lower previsions. This is particular important since set algebras can be considered as prototypical information algebras, serving as the algebraic counterparts to classical propositional logic. Consequently, this integration can be viewed as an alternative approach to demonstrating that the latter is integrated into the framework of imprecise probabilities.

The results obtained in this chapter offer a novel algebraic analysis of desirability and imprecise probabilities in a broader sense. Furthermore, they enrich desirability by equipping it with the machinery for inference provided by information algebras, which permit to more easily deal with, and even generalise, some related problem such as the marginal one. Conversely, desirability serves as a general instance of information algebras, capable of handling very different forms of information.

Chapter 4

Non-linear desirability

In Section 1, we saw that rationality requirements of *Additivity* and *Positive Homogeneity* are justified by the crucial assumption that rewards of gambles are expressed in units of utility in a linear scale. This however can be a limitation in the interpretation of gambles, for example if we interpret gambles as payoff vectors of *monetary* lotteries.[1]

In frameworks close to Walley and Williams' desirability theory, some criticisms to these axioms can already be found in the works of Nau [1992] and Pelessoni and Vicig [2005, 2016], where convexity is proposed as a relaxation of these requirements. In particular, in Pelessoni and Vicig [2005, 2016] two different classes of lower previsions not necessarily coherent are proposed (*centered convex* and *centered 2-convex* lower previsions) with interesting applications in risk measurement.

In Section 1.1.2 moreover, we noted that rationality axioms D1, D2, D3, D4 are tightly linked to axioms on preferences required in the traditional decision-theoretic formalisation of incomplete preferences, as it was shown in Zaffalon and Miranda [2017, 2021]. In this context, relaxing axioms D3 and D4 is even more important since it permits to model preference relations not limited by the criticised axiom of *mixture independence*, see again Section 1.1.2.

For these main reasons, we present here some instances of *nonlinear* desirability, i.e., generalisations of standard desirability relaxing the rationality requirements of *Additivity* and *Positive Homogeneity*.

We limit our analysis to finite possibility spaces and, in this framework, we present an interpretation of both standard desirability and instances of nonlinear desirability proposed (*convex coherence* and *positive additive coherence*) as

[1]For example, for a large positive λ difficulties might be encountered at accepting λf for every acceptable gamble f, because of lack of market liquidity at some degree.

classification problems. Specifically, this means that we analyze different sets of rationality axioms and, for each one of them, we show that proving the existence of a set respecting these axioms and the will of an agent only providing a finite set of (almost) *acceptable* and a finite set of *rejectable*[2] gambles can be reformulated as a binary classification problem, usually nonlinear. To avoid to work with too complicated border structures, we assume as the standard formulation of desirability the one given in terms of D1'–D5' instead of the one given by D1–D4. Thus, we consider different axiomatizations of coherence that maintain axioms D1', D2',D5' and relax axioms D3', D4'.[3] Finally, by borrowing ideas from machine learning, we show the possibility of defining a *feature mapping*, which allows us to reformulate the above nonlinear classification problems as linear ones in higher-dimensional spaces.

A wider extension of the theory of desirability is given in the work of Miranda and Zaffalon [2023], more recently published. Here, they give a general definition of nonlinear desirability based on arbitrary closure operators comprehensive of the non-finite case. While their approach is more general, ours allows for the consideration of rejection statements and it is more closely tied to practical applications, see Chapter 5.

In the next sections we discuss in detail the results found in our works Casanova, Benavoli and Zaffalon [2021]; Casanova et al. [2023]. Specifically, in Section 4.1 and Section 4.2, we introduce our framework and we reformulate standard cases of desirability as classification problems. The latter will serve as the basis for the different definitions of coherence given in Section 4.3 and Section 4.4. In particular, in Section 4.2, we also introduce a procedure to reformulate the nonlinear classification problems emerging in our analysis as linear ones in higher-dimensional spaces. Finally, in Section 4.5, we derive lower previsions from sets of gambles satisfying each of the set of rationality axioms provided and in Section 4.6 we propose a possible generalisation of our procedure.

The proof of all the discussed material can be found in Appendix D.

[2]By rejecting a gamble, the agent expresses the idea that they consider the (almost) acceptability of that gamble unreasonable.

[3]Given our interpretation of gambles as (more generic) bets and the finiteness assumption we made on the possibility space, we have chosen to retain axioms D1' and D2', even though they may pose computational challenges in more general contexts [Benavoli et al., 2019]. As for axiom D5', we have opted to keep it because it strengthens the connection between desirability and probability theory, thus simplifying subsequent discussions on the relationship between sets and lower previsions. In the future, it may indeed be worthwhile to explore variants of desirability that relax also these assumptions.

4.1 Standard maximal coherence

Let us consider here and in the rest of the chapter, a finite possibility space $\Omega := \{\omega_1, \ldots, \omega_n\}$. In this context, gambles f can be regarded as column vectors, denoted as $\boldsymbol{f}$, in $\mathscr{L} = \mathbb{R}^n$. For simplicity, in what follows, we use the notation f_i to indicate $f(\omega_i)$ for every $i \in \{1, \ldots, n\}$. Hence, we have: (i) $\boldsymbol{f} \geq \boldsymbol{g}$ if and only if $f_i \geq g_i$ for every $i \in \{1, \ldots, n\}$; (ii) $\boldsymbol{f} \gneq \boldsymbol{g}$ if and only if $\boldsymbol{f} \geq \boldsymbol{g}$ and $f_k > g_k$ for some $k \in \{1, \ldots, n\}$; (iii) $\boldsymbol{f} > \boldsymbol{g}$ if and only if $f_i > g_i$ for every $i \in \{1, \ldots, n\}$. We also introduce the following sets that are extensively used later on:

$$T := \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f} \geq \boldsymbol{0}\},$$
$$F := \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f} < \boldsymbol{0}\}.$$

In this framework, suppose to have a modeler who considers an agent *rational* if their beliefs can be summarised in a maximal coherent set of almost desirable gambles $\overline{M}$, see Definition 14. Suppose, however, that the only information about the agent the modeller receives are two finite sets of gambles A and R such that, denoted with $\overline{\mathscr{K}}$ the generic set of gambles the agent finds almost desirable, $A \subseteq \overline{\mathscr{K}}$ and $\overline{\mathscr{K}} \cap R = \emptyset$ - in what follows, with a little abuse of nomenclature, we simply say that A and R are respectively *acceptable* and *rejectable* for the agent-.[4] In this situation, we can assume the modeller still regards the agent as rational if (and only if) there exists a maximal coherent set of almost desirable gambles $\overline{M}$ respecting their willingness to accept/reject gambles, i.e., such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$.

To establish the existence of such a set, it is possible to proceed by re-interpreting the equivalence between maximal coherent sets of almost desirable gambles and linear previsions/probability mass functions (see Section 1.1.1) as a binary linear classification problem. If we consider two finite sets of gambles A and R, in fact, there exists a maximal coherent set of almost desirable gambles $\overline{M} \subseteq \mathbb{R}^n$ such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$ if and only if there exists a linear prevision P such that $P(\boldsymbol{f}) \geq 0$ for every $\boldsymbol{f} \in A$ and $P(\boldsymbol{f}) < 0$ for every $\boldsymbol{f} \in R$, see Eq.(1.11). Every linear prevision, however, is an expected value operator, i.e., a linear operator $P(\boldsymbol{f}) = \boldsymbol{f}^\top \boldsymbol{\beta}$ for every $\boldsymbol{f} \in \mathbb{R}^n$, where $\boldsymbol{\beta} \in \mathbb{R}^n$ is a probability mass function: $\boldsymbol{\beta} \gneq 0$ and $\sum_{i=1}^n \beta_i = 1$. In particular, it is easy to verify that the agent is considered rational if and only if there exists a binary linear classifier LC defined

[4] This framework is closer to practical situations where an agent can only declare their opinions regarding a finite set of gambles. A similar framework for reasoning with uncertainty based on accept and reject statements about gambles has already been developed in Quaeghebeur et al. [2015]. In Quaeghebeur et al. [2015] however, also non-finite sets of acceptance/rejection statements are considered and the rationality axioms of *Additivity* and *Positive Homogeneity* are always assumed.

as:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ LC(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \boldsymbol{f}^\top \boldsymbol{\beta} \geq 0, \\ -1 & \text{otherwise,} \end{cases} \tag{4.1}$$

with $\boldsymbol{\beta} \in \mathbb{R}^n$,[5] classifying gambles in $A \cup T$ as 1 and gambles in $R \cup F$ as -1. More formally, we can introduce the following definitions and results.

Definition 52 (Linear separability). *A pair of sets of gambles (A, B) is* linearly separable *if and only if there exists a binary linear classifier LC of type* (4.1) *such that $LC(A) = 1$ and $LC(B) = -1$.*[6] *We indicate the set of these classifiers with $\mathrm{LC}(A, B)$.*[7]

As explained above, the following result is valid.

Proposition 5. *Given a pair of finite sets of gambles (A, R), there exists a maximal coherent set of almost desirable gambles $\overline{M} \subseteq \mathbb{R}^n$ such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$, if and only if $(A \cup T, R \cup F)$ is linearly separable.*

Notably, if there exists a maximal coherent set of almost desirable gambles $\overline{M}$ such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$, it is possible to construct a classifier $LC \in \mathrm{LC}(A \cup T, R \cup F)$ such that $\overline{M} = \{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) = 1\}$. Vice versa, if there exists a classifier $LC \in \mathrm{LC}(A \cup T, R \cup F)$, the region $\overline{M} = \{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) = 1\}$ is a maximal coherent set of almost desirable gambles such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$.

From the previous considerations, it is also possible to deduce the following proposition that simplifies the classification problem.

Proposition 6. *Consider a pair of finite sets of gambles (A, R). Every classifier $LC \in \mathrm{LC}(A \cup T, R \cup F)$ is a classifier $LC \in \mathrm{LC}(A, R)$ with parameter $\boldsymbol{\beta} \gneq 0$, and vice versa.*

Corollary 4. *Given a pair of finite sets of gambles (A, R), there exists a maximal coherent set of almost desirable gambles $\overline{M} \subseteq \mathbb{R}^n$, such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$, if and only if (A, R) is linearly separable and there exists a classifier $LC \in \mathrm{LC}(A, R)$ with parameter $\boldsymbol{\beta} \gneq 0$.*[8]

[5] We assume without loss of generality that the binary classifier has 1 and -1 as classes.

[6] With a little abuse of notation, with $LC(\mathscr{K}) = c$ for $\mathscr{K} \subseteq \mathbb{R}^n$, $c \in \{-1, 1\}$, we mean $LC(\boldsymbol{f}) = c$ for every $\boldsymbol{f} \in \mathscr{K}$. We will use the same notation also for the other types of binary classifiers considered later on.

[7] In this definition, we assume the classification constraints to hold only for non-empty sets. In particular, if $A = B = \emptyset$ we assume the pair (A, B) to be linearly separable and $\mathrm{LC}(A, B)$ to be the whole set of binary linear classifiers of type (4.1).

[8] We would like to express our gratitude to a reviewer of this thesis for suggesting that, given that $-\boldsymbol{f}^T \boldsymbol{\beta} = -(\boldsymbol{f}^T \boldsymbol{\beta})$, apart from the zeros, we could have alternatively asked the classifier to exclusively classify the set $A \cup (-R)$ as 1. However, we'd like to emphasize that this alternative representation isn't suitable for the other classifiers we considered in the following sections. This consideration further supports our choice to separately classify A and R.

If the parameter $\boldsymbol{\beta}$ of a binary linear classifier of type (4.1) is such that $\boldsymbol{\beta} \gneq 0$, i.e., the classifier satisfies the constraint provided by Proposition 6 and Corollary 4, it can be regarded as a probability mass function on Ω. This leads to the following reformulation of Corollary 4.

Corollary 5. *Given a pair of finite sets of gambles (A,R), there exists a maximal coherent set of almost desirable gambles $\overline{M} \subseteq \mathbb{R}^n$, such that $\overline{M} \supseteq A$ and $\overline{M} \cap R = \emptyset$, if and only if equivalently:*

- *there exists a binary linear classifier $LC \in LC(A,R)$ with (normalised) parameter $\boldsymbol{\beta} \gneq 0$;*
- *there exists a probability mass function on Ω, π, such that $E_\pi(\boldsymbol{f}) \geq 0$ for every $\boldsymbol{f} \in A$ and $E_\pi(\boldsymbol{f}) < 0$ for every $\boldsymbol{f} \in R$.*

Coonsiderations above lead to different representations of maximal sets of almost desirable gambles. We reported them in the following diagram pointing out the connections existing among them. We highlight in black the relations already known in literature and in green the ones established by us.

All the machinery here introduced will then be used in an analogous way in the following sections.

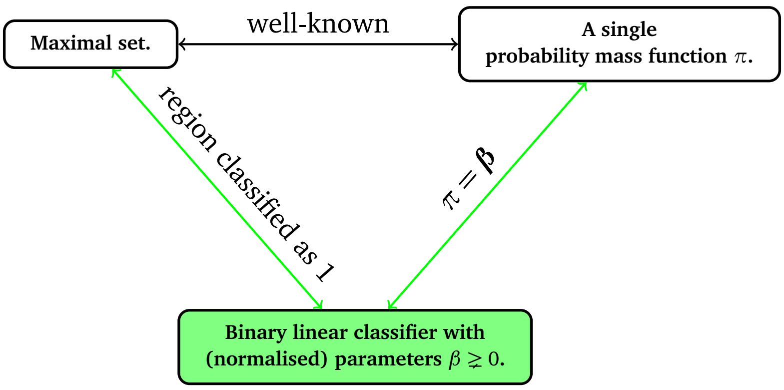

Figure 4.1. Diagram showing equivalent models for representing maximal sets of almost desirable gambles.

Given two finite sets A and R, there can be classifiers $LC \in \mathrm{LC}(A,R)$ with (normalised) parameter $\boldsymbol{\beta} \gneq 0$ identifying different regions $\{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) = 1\}$. How can we learn a unique one? An idea can be to learn a classifier leading to make the minimal assumptions on the agent's dispositions to accept gambles, i.e., a classifier that identifies the minimal acceptance region $\{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) =$

1}. However, all the regions $\{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) = 1\}$ are minimal since they are maximal sets. We need therefore other criteria. Another one could be to get, if it exists, the binary linear classifier leading to the probability distribution with the minimal *Kullback–Leibler divergence* from the uniform one:[9]

$$LC \in \mathrm{LC}(A, R) \text{ with (normalised) } \boldsymbol{\beta} \gneq 0, \text{ minimizing: } \sum_{i \in \{1,..,n\}} \beta_i \log_2 (n\beta_i),$$

but other criteria could be used.

The following example illustrates the results obtained in this section considering a specific numerical case and reports a numerical optimization procedure to solve the classification problem and find the parameter of the classifer. It also highlights some of the problems that our procedure can cause. It is possible indeed that the learning procedure illustrated above leads to a classifier whose parameter *does not* correspond to the real beliefs of an agent. On the other hand, it is also possible that the modeller judges an agent rational even if they are not. The idea however is that, in both the situations, augmenting the dimension of the finite sets of acceptable and/or rejectable gambles declared by the agent can solve the problem. Similar considerations can be made for the other classification tasks considered later on.

Example 12. *Let us consider again the framework of Example 1. In particular, let us consider $\Omega = \{$ Tall (t), Short (s)$\}$. A gamble $\boldsymbol{f}$ has therefore two components $\boldsymbol{f}(t) = f_1$ and $\boldsymbol{f}(s) = f_2$. In this context, we can think of outcomes of gambles directly as amounts in thousands of dollars.*

In this framework, let us consider Anne (an agent) who is disposed to accept/reject gambles on the basis of the linear prevision P_{Anne}, defined as $P_{Anne}(\boldsymbol{f}) := E_{\{2/3,1/3\}}(\boldsymbol{f})$,[10] for every $\boldsymbol{f} \in \mathbb{R}^n$ (this can be interpreted as she believes that the probability that the murderer is tall is 2/3). That is, she is characterised by the following maximal coherent set of almost desirable gambles:

$$\overline{M}_{Anne} := \{\boldsymbol{f} \in \mathbb{R}^n : P_{Anne}(\boldsymbol{f}) \geq 0\} := \{\boldsymbol{f} \in \mathbb{R}^n : E_{\{2/3,1/3\}}(\boldsymbol{f}) \geq 0\}.$$

Then, for example, she is disposed to accept gambles in $A_{Anne} := \{[-1,2]^\top, [2,-1]^\top, [1,-1]^\top, [4,-2]^\top, [1,-0.5]^\top\}$, and reject gambles in $R_{Anne} := \{[-3,2]^\top\}$.

In Figure 4.2 it is possible to find a graphical representation of the gambles contained in A_{Anne} and R_{Anne} in the two-dimensional Cartesian coordinate system

[9] This approach leads to a unique solution. Indeed, the discrete probability distribution that minimizes the Kullback-Leibler divergence with respect to the uniform one corresponds to the distribution that maximizes the *Shannon's entropy*. The latter then is unique if we consider linear constraints, see for example Oikonomou and Grünwald, 2015, Section 2.

[10] We indicate with $E_{\{p,1-p\}}(\boldsymbol{f})$, the expected value of a gamble $\boldsymbol{f}$ calculated with respect to the probability mass function that assigns probability p to Tall and $(1-p)$ to Short.

representing gambles' values respectively on the alternatives t and l. Gambles in A_{Anne} are denoted with (blue) points, gambles in R_{Anne} are instead denoted with (red) triangles.

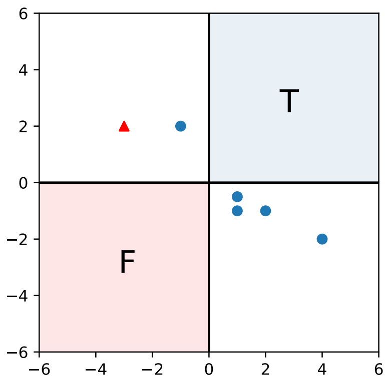

Figure 4.2. Gambles contained in A_{Anne} and R_{Anne}.

Suppose now that the only information a modeller, Bruce, who judges an agent rational if their beliefs are modeled by a maximal coherent set of almost desirable gambles, has on Anne is represented by the two sets A_{Anne}, R_{Anne} just defined. Bruce believes Anne is rational because, for example, the binary linear classifier LC of type (4.1) characterised by the parameter $\boldsymbol{\beta}^{(1)} := \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$ belongs to $LC(A_{Anne}, R_{Anne})$. Notice that this is the linear classifier in $LC(A_{Anne}, R_{Anne})$ with (normalised) parameter corresponding to the probability distribution closest to the uniform one with respect to the Kullback–Leibler divergence (the divergence between the two distributions is indeed 0).

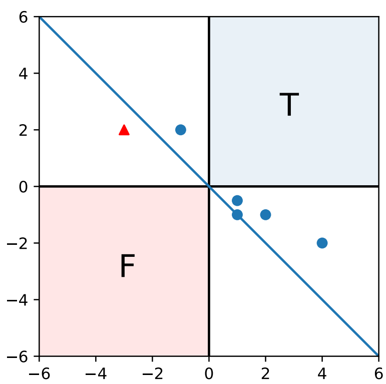

Figure 4.3. Graphical representation of the decision boundary of the linear classifier LC.

However, the region $\{\boldsymbol{f} \in \mathbb{R}^n : LC(\boldsymbol{f}) = 1\}$ does not *correspond to Anne's set of almost desirable gambles.*

The idea, however, is that by querying Anne on additional gambles, that is by augmenting the dimension of the sets A_{Anne} *and/or* R_{Anne}*, it shall be possible to find a binary linear classifier* $LC \in LC(A_{Anne} \cup T, R_{Anne} \cup F)$ *whose region classified as* 1 *corresponds to Anne's set of almost desirable gambles. For example, suppose that Anne adds the gamble* $[1, -2]^\top$ *to* A_{Anne} *forming a new set of acceptable gambles* $A'_{Anne} := A_{Anne} \cup \{[1, -2]^\top\}$*. Then, the binary linear classifier* $LC_A \in LC(A'_{Anne} \cup T, R_{Anne} \cup F)$ *corresponding to the probability distribution closest to the uniform one with respect to the Kullback-Leibler divergence can be found numerically by solving the convex optimization problem:*

$$\begin{cases} \max_{\{\boldsymbol{\beta} \in \mathbb{R}^2\}} (-\sum_{i=1}^2 \beta_i \log_2 \beta_i) \\ s.t. \\ -\boldsymbol{f}^\top \boldsymbol{\beta} \leq 0 \quad \forall \boldsymbol{f} \in A'_{Anne}, \\ \quad \boldsymbol{f}^\top \boldsymbol{\beta} \leq -\epsilon \quad \forall \boldsymbol{f} \in R_{Anne}, \\ \displaystyle\sum_{i=1}^2 \beta_i = 1, \\ \beta_1, \beta_2 \geq 0, \end{cases}$$

where ϵ *is a small positive number. The solution is* $\boldsymbol{\beta}^{(2)} := \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$*, which is also the parameter of the unique (with normalised parameter) binary linear classifier in* $LC(A'_{Anne} \cup T, R_{Anne} \cup F)$*. The decision boundary of this latter classifier is shown in Figure 4.4. In this case, the region* $\{\boldsymbol{f} \in \mathbb{R}^n : LC_A(\boldsymbol{f}) = 1\}$ corresponds *to Anne's set of almost desirable gambles and the components of* $\boldsymbol{\beta}^{(2)}$ *coincide with the values of her probability mass function on* Ω*.*

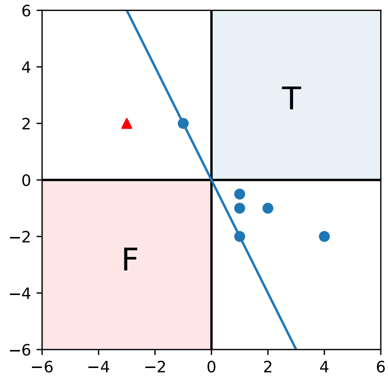

Figure 4.4. Graphical representation of the decision boundary of the linear classifier LC_A.

It is also possible that Bruce judges an agent to be rational when, actually, they are not.

Consider in this regard another agent Valerie, who is initially disposed to accept A_{Anne} and reject R_{Anne}, as well as Anne. Assume however that she, at a later time, declares to be disposed to reject also the gambles $[-2,4]^\top, [2,-2]^\top$ obtaining a new set of rejectable gambles $R_{Valerie} := R_{Anne} \cup \{[-2,4]^\top, [2,-2]^\top\}$. This leads to a nonlinearly separable *pair of finite sets of acceptable and rejectable gambles $(A_{Valerie}, R_{Valerie})$, where $A_{Valerie} = A_{Anne}$. The latter sets are represented in the following Figure 4.5. As usual, gambles in $A_{Valerie}$ are denoted with (blue) points and gambles in $R_{Valerie}$ are denoted with (red) triangles. Now Bruce is sure that she cannot be rational according to the criteria he has set.*

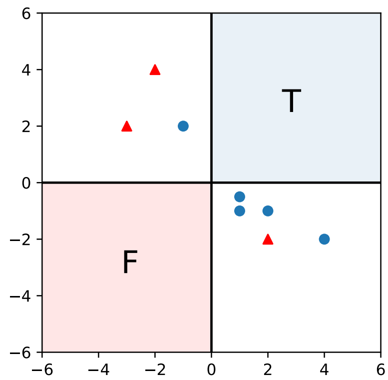

Figure 4.5. Example of a nonlinearly separable pair of finite sets of gambles.

4.2 Standard coherence

Let us consider now a modeller who takes D1'–D5' as basic rationality axioms. We can assume, similarly to before, they will judge an agent to be rational on the basis of a finite set of acceptable gambles A and a finite set of rejectable gambles R, if and only if there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$.

To establish the existence of such a set we can again re-interpret the duality between coherent sets of almost desirable gambles and convex and closed sets of linear previsions/probability mass functions, see Section 1.1.1, as a binary classification task. This time, however, the resulting classification problem is *nonlinear*.

In particular, in what follows, we concentrate on finitely generated coherent sets of almost desirable gambles $\overline{\mathscr{D}} = \overline{\mathscr{E}(A)}$ for some finite set $A \subseteq \mathbb{R}^n$, see Section 1.1.1.

Let us introduce the class of binary classifiers that will be used in this section.

Definition 53 (Binary piecewise linear classifier). *We use the term* binary piecewise linear classifier *to denote a classifier PLC defined as follows:*

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ PLC(\boldsymbol{f}) := \begin{cases} c_1 & \text{if } \boldsymbol{f}^\top \boldsymbol{\beta}_j \geq 0 \text{ for all } j \in \{1, \ldots N\}, \\ c_2 & \text{otherwise}, \end{cases} \tag{4.2}$$

with labels c_1, c_2 and $\boldsymbol{\beta}_j \in \mathbb{R}^n$ for all j, $N \geq 1$.

Without loss of generality, in what follows we limit ourselves to binary piecewise linear classifiers with $c_1 = 1$ and $c_2 = -1$.

Definition 54 (Piecewise linear separability). *A pair of sets of gambles (A, B) is* piecewise linearly separable *if and only if there exists a binary piecewise linear classifier PLC such that $PLC(A) = 1$ and $PLC(B) = -1$. We indicate the set of these classifiers with* PLC(A, B).

Proposition 7. *Given a pair of finite sets of gambles (A, R), there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, if and only if $(A \cup T, R \cup F)$ is piecewise linearly separable.*

In particular, if there exists a finitely generated coherent set $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, it is possible to construct a classifier $PLC \in$ PLC$(A \cup T, R \cup F)$ such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\}$. Vice versa, the region classified as 1 by a binary piecewise linear classifier is always a convex cone. However, if there exists a classifier $PLC \in$ PLC$(A \cup T, R \cup F)$, the region $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\}$ is a finitely generated coherent set such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Notice moreover that if there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, the smallest such set is $\overline{\mathscr{E}(A)}$, which is finitely generated. Binary piecewise linear classifiers PLC such that $\{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\} = \overline{\mathscr{E}(A)}$, in particular, identify the minimal 1-region containing A. Hence, they correspond to make the minimal assumptions about the agent's dispositions to accept gambles.

Like before, it is then possible to transform the problem of classifying the two non-finite sets of gambles T and F into a set of constraints on the parameters of the classifier. Thus, establishing the rationality of an agent boils down to classifying only two finite sets.

Proposition 8. *Consider a pair of finite sets of gambles (A,R). Every classifier $PLC \in PLC(A \cup T, R \cup F)$ is a classifier $PLC \in PLC(A,R)$ with parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$ such that $\boldsymbol{\beta}_j \gneq 0$ for every $j \in \{1,\dots,N\}$, and vice versa.*

Corollary 6. *Given a pair of finite sets of gambles (A,R), there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, if and only if (A,R) is piecewise linearly separable and there exists a classifier $PLC \in PLC(A,R)$ with parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$ such that $\boldsymbol{\beta}_j \gneq 0$ for every $j \in \{1,\dots,N\}$.*

If the parameters of a piecewise linear classifier satisfy the constraints expressed by these last results, i.e., they are non-negative and non-vanishing, we can give them a probabilistic interpretation. Each of them, indeed, can be normalised and regarded as a probability mass function on Ω. This leads to the following reformulation of Corollary 6.

Corollary 7. *Given a pair of finite sets of gambles (A,R), there exists a coherent set of almost desirable gambles (not necessarily finitely generated) $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$, such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, if and only if equivalently:*

- *there exists a binary piecewise linear classifier $PLC \in PLC(A,R)$ with (normalised) parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$ such that $\boldsymbol{\beta}_j \gneq 0$ for all j;*
- *there exists a finite set of probability mass functions on Ω, $\{\pi_j\}_{j=1}^N$, such that $E_{\pi_j}(\boldsymbol{f}) \geq 0$ for every $\boldsymbol{f} \in A$, $j \in \{1,\dots,N\}$ and, for every $\boldsymbol{f} \in R$, there exists $k \in \{1,\dots,N\}$ such that $E_{\pi_k}(\boldsymbol{f}) < 0$.*

Specifically, in this case, we can essentially regard $\{\boldsymbol{\beta}_j\}_{j=1}^N$ as the extreme points of the credal set associated to $\overline{\mathscr{D}} := \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\}$, see Section 1.1.1.

These considerations lead to different representations of finitely generated coherent sets of almost desirable gambles. We reported them in the following diagram pointing out the connections existing among them. As before, we highlight in black the relations already known in literature and in green the ones established by us.

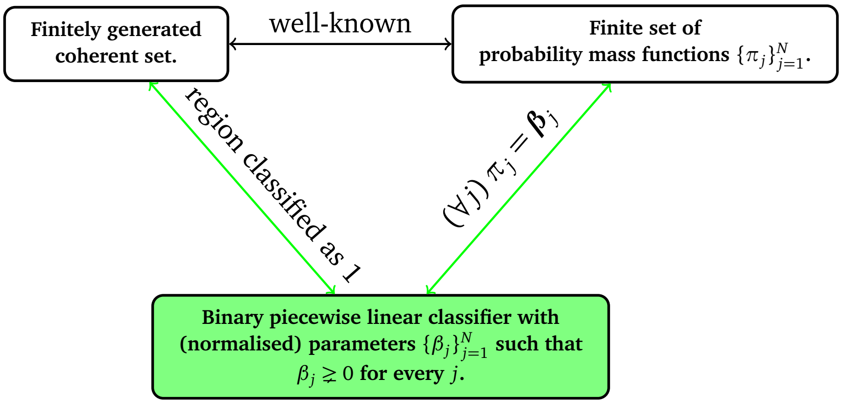

Figure 4.6. Diagram showing equivalent models for representing finitely generated coherent sets of almost desirable gambles.

The following example illustrates the results obtained in this section considering a specific numerical case and reports a numerical optimization procedure to solve the classification problem and find the coefficients of the classifer.

Example 13. *Consider again the framework of Example 12 and another agent, Claire, who is disposed to accept/reject gambles on the basis of the coherent lower prevision $\underline{P}_{Claire}$, defined as $\underline{P}_{Claire}(f) := \min\{E_{\{1/3,2/3\}}(f), E_{\{2/3,1/3\}}(f)\}$, for every $f \in \mathbb{R}^n$ (this can be interpreted as she believes that the probability of t lies in the interval $[1/3, 2/3]$). That is, she is characterised by the following coherent set of almost desirable gambles:*

$$\overline{\mathscr{D}}_{Claire} := \{f \in \mathbb{R}^n : \underline{P}_{Claire}(f) \geq 0\} := \{f \in \mathbb{R}^n : \min\{E_{\{1/3,2/3\}}(f), E_{\{2/3,1/3\}}(f)\} \geq 0\}.$$

Then, for example, she is disposed to accept gambles in $A_{Claire} = \{[-1,2]^\top, [2,-1]^\top, [4,-2]^\top, [1,-0.5]^\top\}$, and she is disposed to reject gambles in $R_{Claire} = \{[-3,2]^\top, [1,-1]^\top\}$. Notice that, differently from Anne, Claire is disposed to reject the gamble $[1,-1]^\top$.

Suppose now that the modeller Bruce sets D1'–D5' as his basic rationality axioms. If Bruce receives, as information on Claire, only the sets A_{Claire} and R_{Claire}, he believes Claire is rational. Indeed, (A_{Claire}, R_{Claire}) is piecewise linearly separable through at least the binary piecewise linear classifier:

$$(\forall f \in \mathbb{R}^n)\ PLC_C(f) := \begin{cases} 1 & \text{if } E_{\{1/3,2/3\}}(f) \geq 0, E_{\{2/3,1/3\}}(f) \geq 0, \\ -1 & \text{otherwise,} \end{cases}$$

which is a classifier that identifies, as the region classified as 1*, the set* $\overline{\mathcal{E}(A_{Claire})}$*, i.e., it corresponds to make the minimal assumptions on Claire's willingness to accept gambles. In this case, in particular, the latter identifies the set of almost desirable gambles of Claire, indeed* $\overline{\mathcal{E}(A_{Claire})} = \{\boldsymbol{f} \in \mathbb{R}^n : PLC_C(\boldsymbol{f}) = 1\} = \{\boldsymbol{f} \in \mathbb{R}^n : E_{\{1/3,2/3\}}(\boldsymbol{f}) \geq 0, E_{\{2/3,1/3\}}(\boldsymbol{f}) \geq 0\} = \{\boldsymbol{f} \in \mathbb{R}^n : \underline{P}_{Claire}(\boldsymbol{f}) \geq 0\}$*. Therefore, Bruce, can also completely reconstruct her probabilistic beliefs from the parameters of the classifier* PLC_C*.*

Figure 4.7 below shows gambles in A_{Claire}*, represented again as (blue) points in the plane* (f_1, f_2)*, gambles in* R_{Claire}*, represented as (red) triangles, and, in blue, the region classified as* 1 *by the classifier* PLC_C*.*

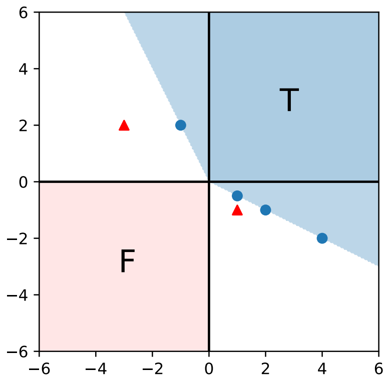

Figure 4.7. Gambles in A_{Claire}, R_{Claire} and the region classified as 1 by PLC_C.

The classifier parameters can be found numerically by solving a sequence of linear programming problems:

$$LP_j = \begin{cases} c_j = \min_{\{\boldsymbol{\beta}_j \in \mathbb{R}^2\}} (\boldsymbol{f}^j)^\top \boldsymbol{\beta}_j \\ s.t. \\ \quad -\boldsymbol{f}^\top \boldsymbol{\beta}_j \leq 0 \quad \forall \boldsymbol{f} \in A_{Claire}, \\ \displaystyle\sum_{i=1}^{2} \boldsymbol{\beta}_j^T \boldsymbol{e}_i = 1, \\ \qquad \boldsymbol{\beta}_j \geq 0, \end{cases}$$

for each $\boldsymbol{f}^j \in A_{Claire} \cup \{\boldsymbol{e}_j\}_{j=1}^n$*, where* $\{\boldsymbol{e}_j\}_{j=1}^n$ *is the canonical basis of* $\mathbb{R}^n$*. Assuming each problem is feasible and denoting the optimal solution of* LP_j *by* $\hat{\boldsymbol{\beta}}_j$*, the classifier parameters are given by all unique* $\hat{\boldsymbol{\beta}}_j$ *such that* $c_j = 0$*. The binary piecewise linear classifier having parameters* $\{\hat{\boldsymbol{\beta}}_j\}_j$ *is a valid classifier (that is, it separates* A_{Claire} *and* R_{Claire}*) provided that for each* $\boldsymbol{f}^j \in R_{Claire}$ *there exists* $\hat{\boldsymbol{\beta}}_k$ *such that* $(\boldsymbol{f}^j)^\top \hat{\boldsymbol{\beta}}_k < 0$*.*

4.2.1 Feature mapping

The binary classification problem introduced in the previous section, which is nonlinear in general, can be reformulated as a linear one in a higher dimensional space using a suitable feature mapping.

Let $\{\mathscr{B}_j\}_{j=1}^N$ denote a partition of $\mathbb{R}^n$ [Piga et al., 2020],[11] where for every j:

$$\mathscr{B}_j := \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f}^\top \boldsymbol{\omega}_j \leq \boldsymbol{f}^\top \boldsymbol{\omega}_k \text{ for } k = 1, \dots, N, \quad j \neq k\}, \tag{4.3}$$

where the vectors $\boldsymbol{\omega}_j \in \mathbb{R}^n$ are parameters defining the partition. We can introduce the feature mapping $\boldsymbol{\phi} : \mathbb{R}^n \to \mathbb{R}^{nN}$, defined as $\boldsymbol{\phi}(\boldsymbol{f}) := [\boldsymbol{\phi}_1(\boldsymbol{f}), \dots, \boldsymbol{\phi}_N(\boldsymbol{f})]^\top$ for every $\boldsymbol{f} \in \mathbb{R}^n$, where $\boldsymbol{\phi}_j : \mathbb{R}^n \to \mathbb{R}^n$ is defined in turn as:

$$\boldsymbol{\phi}_j(\boldsymbol{f}) := \mathbb{I}_{\mathscr{B}_j}(\boldsymbol{f})\boldsymbol{f}, \tag{4.4}$$

where

$$\mathbb{I}_{\mathscr{B}_j}(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \boldsymbol{f} \in \mathscr{B}_j, \\ 0 & \text{otherwise,} \end{cases}$$

for every $\boldsymbol{f} \in \mathbb{R}^n$ and $j \in \{1, \dots, N\}$. Further, we define the following classifier corresponding to a linear classifier in the feature space:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\; LC_{\boldsymbol{\phi}}(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \sum_{j=1}^N \boldsymbol{\phi}_j(\boldsymbol{f})^\top \boldsymbol{\beta}'_j \geq 0, \\ -1 & \text{otherwise,} \end{cases} \tag{4.5}$$

with $\boldsymbol{\beta}'_j \in \mathbb{R}^n$ for all $j = 1, \dots, N$. In what follows, we consider both $\{\boldsymbol{\beta}'_j\}_{j=1}^N$ and $\{\boldsymbol{\omega}_j\}_{j=1}^N$ as parameters of $LC_{\boldsymbol{\phi}}$.

Finally, we introduce the following definition to simplify notation.

Definition 55 ($\boldsymbol{\Phi}$-separability)**.** *A pair of sets of gambles* (A, B) *is* $\boldsymbol{\Phi}$-separable *if and only if there exists a classifier* $LC_{\boldsymbol{\phi}}$ *of type* (4.5) *such that* $LC_{\boldsymbol{\phi}}(A) = 1$ *and* $LC_{\boldsymbol{\phi}}(B) = -1$*. We indicate the set of these classifiers with* $LC_{\boldsymbol{\Phi}}(A, B)$*.*

We can now state the following result.

Proposition 9. *A binary piecewise linear classifier with parameters* $\{\boldsymbol{\beta}_j\}_{j=1}^N$ *and a classifier of type* (4.5) *with parameters* $\{\boldsymbol{\omega}_j, \boldsymbol{\beta}'_j\}_{j=1}^N$ *such that* $\boldsymbol{\beta}'_j = \boldsymbol{\omega}_j = \boldsymbol{\beta}_j$ *for every* $j \in \{1, .., N\}$ *classify gambles in the same way.*

[11] We call it *partition* with a little abuse of notation. Indeed, we guarantee only that $\operatorname{int} \mathscr{B}_j \cap \operatorname{int} \mathscr{B}_k = \emptyset$, for every $j, k \in \{1, \dots, N\}$, $j \neq k$, where $\operatorname{int} \mathscr{B}_j$ represents the interior of $\mathscr{B}_j$ under the usual topology of $\mathbb{R}^n$. Instead, it is guaranteed that $\cup_{j=1}^N \mathscr{B}_j = \mathbb{R}^n$ (for every $\boldsymbol{f}$, $\{\boldsymbol{f}^\top \boldsymbol{\omega}_j\}_{j \in N}$ is a finite set of real values so the minimum always exists). We can instead obtain a real partition if we combine our definition with a lexicographic order on $\{1, ...N\}$, so that for every j, $\mathscr{B}_j$ becomes the set of gambles for which the minimum is obtained with ω_j but not with ω_k for $k < j$. A similar reasoning can be applied also for the other definitions of partition given in Section 4.3.1 and Section 4.4.1. Nevertheless, for the sake of simplicity, we have chosen to adopt this particular definition.

The intuitive reasoning behind Proposition 9 is the following. Partitions $\{\mathscr{B}_j\}_{j=1}^N$ allow us to divide $\mathbb{R}^n$ in subsets where binary piecewise linear classifiers boil down to linear ones. Consider, indeed, a binary piecewise linear classifier PLC with parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$. For every $\boldsymbol{f} \in \mathbb{R}^n$, $PLC(\boldsymbol{f})$ depends on the value in $\boldsymbol{f}$ of the nonlinear functional $plc : \mathbb{R}^n \to \mathbb{R}$, defined as $plc(\boldsymbol{h}) := \min\{\boldsymbol{h}^\top \boldsymbol{\beta}_1, ..., \boldsymbol{h}^\top \boldsymbol{\beta}_N\}$, for every $\boldsymbol{h} \in \mathbb{R}^n$. Consider now the partition $\{\mathscr{B}_j\}_{j=1}^N$ with parameters $\boldsymbol{\omega}_j = \boldsymbol{\beta}_j$ for every $j \in \{1, ..., N\}$. If $\boldsymbol{f} \in \mathscr{B}_j$, $plc(\boldsymbol{f}) = \boldsymbol{f}^\top \boldsymbol{\beta}_j$. Hence, if we restrict gambles on suitable members of the partition, the nonlinear functional plc collapses on a linear one. The feature mapping $\boldsymbol{\phi}$ allows us to do that. If we then consider a linear classifier of type (4.5) with parameters $\{\boldsymbol{\omega}_j, \boldsymbol{\beta}'_j\}_{j=1}^N$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}_j = \boldsymbol{\beta}_j$ for every j, we have:

$$\begin{aligned}\min(\boldsymbol{f}^\top \boldsymbol{\beta}_1, \ldots, \boldsymbol{f}^\top \boldsymbol{\beta}_N) \geq 0 &\iff \boldsymbol{f}^\top \boldsymbol{\beta}_j \geq 0 \text{ whenever } \boldsymbol{f} \in \mathscr{B}_j \\ \iff 0 \leq \sum_{j=1}^N (\mathbb{I}_{\mathscr{B}_j}(\boldsymbol{f})\boldsymbol{f})^\top \boldsymbol{\beta}_j &=: \sum_{j=1}^N \boldsymbol{\phi}_j(\boldsymbol{f})^\top \boldsymbol{\beta}_j.\end{aligned}$$

The following corollary, which can be considered the central result of this subsection, thus holds.

Corollary 8. *Given a pair of finite sets of gambles (A, R), there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, if and only if (A, R) is $\boldsymbol{\Phi}$–separable and there exists a classifier $LC_{\boldsymbol{\phi}} \in LC_{\boldsymbol{\Phi}}(A, R)$ with parameters $\{\boldsymbol{\omega}_j, \boldsymbol{\beta}'_j\}_{j=1}^N$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}_j \gneq 0$ for every $j \in \{1, ..., N\}$.*

In particular, if $\overline{\mathscr{D}}$ is a finitely generated coherent set of almost desirable gambles, we can find $LC_{\boldsymbol{\phi}}$ of the type specified in Corollary 8 such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\boldsymbol{\phi}}(\boldsymbol{f}) = 1\}$, and vice versa.

Analogous considerations will be made for the feature mappings introduced for the other definitions of coherence analysed in Section 4.3 and in Section 4.4.

Example 14. *Consider again the framework of Example 13. In this context, let us construct the classifier $LC^C_{\boldsymbol{\phi}}$ of type (4.5) with parameters $\{\boldsymbol{\omega}_j, \boldsymbol{\beta}'_j\}_{j=1}^2$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}_j = \hat{\boldsymbol{\beta}}_j$ for every j, where $\{\hat{\boldsymbol{\beta}}_j\}_{j=1}^2$ are the parameters of PLC_C introduced in Example 13: $\hat{\boldsymbol{\beta}}_1 = [\frac{1}{3}, \frac{2}{3}]^\top$, $\hat{\boldsymbol{\beta}}_2 = [\frac{2}{3}, \frac{1}{3}]^\top$. By Proposition 9, we know that $LC^C_{\boldsymbol{\phi}}$ is a reformulation of PLC_C. It therefore classifies A_{Claire} as 1 and R_{Claire} as −1. Moreover, $\{\boldsymbol{f} \in \mathbb{R}^n : LC^C_{\boldsymbol{\phi}}(\boldsymbol{f}) = 1\} = \overline{\mathscr{E}(A_{Claire})}$.*

4.3 Convex coherence

In this section we consider a weaker form of rationality made up of the axioms D1', D2', D5' and a relaxed version of the axioms D3' and D4':

(CNV) : $f, g \in \overline{\mathscr{D}} \Rightarrow (\forall \gamma \in [0,1])\, \gamma f + (1-\gamma) g \in \overline{\mathscr{D}}$ [Convexity].

We define a set $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ satisfying D1', D2', (CNV), D5' a *convex coherent* set of almost desirable gambles. For simplicity, we indicate it with the same notation used for the standard coherent sets of almost desirable gambles.

As before, we can assume that if an agent declares two finite sets of gambles, A and R, as acceptable and rejectable respectively, a modeller who subscribes to the above notion of rationality will deem the agent to be rational if and only if there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. If this is the case, the minimal such set is $\overline{\mathrm{ch}(A \cup T)}$, i.e., the closed convex hull of $A \cup T$,[12]. Geometrically, the latter corresponds to a convex polyhedron:[13]

$$\overline{\mathrm{ch}(A \cup T)} = \mathrm{ch}^+(A \cup \{\mathbf{0}\}) := \{f \in \mathbb{R}^n : f \geq g,\ g \in \mathrm{ch}(A \cup \{\mathbf{0}\})\}.$$

This consideration suggests the following definition.

Definition 56 (Finitely generated convex coherent set). *Let us consider a convex coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$. If $\overline{\mathscr{D}} = \overline{\mathrm{ch}(A \cup T)}$ for some finite set $A \subseteq \mathbb{R}^n$, we say that $\overline{\mathscr{D}}$ is* finitely generated.

We claim that, also here, proving whether an agent is rational on the basis of a pair of finite sets (A, R) is equivalent to solving a binary classification task.

Let us first introduce the classifiers used in what follows.

Definition 57 (Binary piecewise affine classifier). *We use the term* binary piecewise affine classifier *to denote a classifier PAC defined as follows:*

$$(\forall f \in \mathbb{R}^n)\ PAC(f) := \begin{cases} c_1 & \text{if } f^\top \boldsymbol{\beta}_j + \alpha_j \geq 0 \text{ for all } j \in \{1, \ldots N\}, \\ c_2 & \text{otherwise,} \end{cases} \tag{4.6}$$

with labels c_1, c_2, and $\boldsymbol{\beta}_j \in \mathbb{R}^n$, $\alpha_j \in \mathbb{R}$ for all j, $N \geq 1$.

Again, without loss of generality, we assume $c_1 = 1$ and $c_2 = -1$.

Definition 58 (Piecewise affine separability). *A pair of sets of gambles (A, B) is* piecewise affine separable *if and only if there exists a binary piecewise affine classifier PAC such that $PAC(A) = 1$ and $PAC(B) = -1$. We indicate the set of these classifiers with $PAC(A, B)$.*

[12] See Lemma 27 in Appendix D.
[13] See Lemma 28 in Appendix D.

Now we can state the main result.

Proposition 10. *Given a pair of finite sets of gambles* (A,R)*, there exists a convex coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$*, such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if* $(A \cup T, R \cup F)$ *is piecewise affine separable.*

In particular, if there exists a finitely generated convex coherent set $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, it is possible to construct a classifier $PAC \in \mathrm{PAC}(A \cup T, R \cup F)$ such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\}$. Notice that, if there exists a convex coherent set $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, the smallest such set is $\overline{\mathrm{ch}(A \cup T)}$, which is in particular finitely generated. Binary piecewise affine classifiers PAC such that $\overline{\mathrm{ch}(A \cup T)} = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\}$ identify the minimal 1-region, then correspond to make the minimal assumptions about the agent's dispositions. Vice versa, if there exists a classifier $PAC \in \mathrm{PAC}(A \cup T, R \cup F)$, the region $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\}$ is a convex coherent set *not* necessarily finitely generated (see Example 21 in Appendix D).

Also in this case, we can transform the problem of classifying the two non-finite sets of gambles T and F into constraints to be required on the parameters of the classifier.

Proposition 11. *Consider a pair of finite sets of gambles* (A,R)*. Every classifier* $PAC \in PAC(A \cup T, R \cup F)$ *is a classifier* $PAC \in PAC(A,R)$ *with parameters* $\{\boldsymbol{\beta}_j, \alpha_j\}_{j=1}^N$ *such that* $\boldsymbol{\beta}_j \gneq 0, \alpha_j \geq 0,$ *for every* $j \in \{1,...,N\}$ *with at least an* $\alpha_k = 0$ *for some* $k \in \{1,\dots,N\}$*, and vice versa.*

Corollary 9. *Given a pair of finite sets of gambles* (A,R)*, there exists a convex coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$*, such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if* (A,R) *is piecewise affine separable and there exists a classifier* $PAC \in PAC(A,R)$ *with parameters* $\{\boldsymbol{\beta}_j, \alpha_j\}_{j=1}^N$ *such that* $\boldsymbol{\beta}_j \gneq 0, \alpha_j \geq 0,$ *for every* $j \in \{1,...,N\}$ *with at least an* $\alpha_k = 0$ *for some* $k \in \{1,\dots,N\}$*.*

Parameters $\{\boldsymbol{\beta}_j\}_j$ satisfying these constraints can still be interpreted as probability mass functions on Ω. A possible interpretation for parameters $\{\alpha_j\}_j$ instead, is the one that regards them as non-negative (with at least a zero one) 'penalty terms' due to, for example, limited financial resources of an agent, as the following Example 15 shows. These considerations lead to the following reformulation of Corollary 9.

Corollary 10. *Given a pair of finite sets of gambles* (A,R)*, there exists a convex coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$*, such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if equivalently:*

- *there exists a binary piecewise affine classifier PAC $\in PAC(A,R)$ with parameters $\{\boldsymbol{\beta}_j, \alpha_j\}_{j=1}^N$ such that $\boldsymbol{\beta}_j \gneq 0, \alpha_j \geq 0$, ($\boldsymbol{\beta}_j$ normalised) for every $j \in \{1,\dots,N\}$ with at least an $\alpha_k = 0$ for some $k \in \{1,\dots,N\}$;*

- *there exists a finite set of probability mass functions on Ω, $\{\pi_j\}_{j=1}^N$, and a finite set of non-negative penalty terms (with at least a zero one) $\{\tilde{\alpha}_j\}_{j=1}^N$, such that $E_{\pi_j}(\boldsymbol{f})+\tilde{\alpha}_j \geq 0$ for every $\boldsymbol{f} \in A$, $j \in \{1,\dots,N\}$ and, for every $\boldsymbol{f} \in R$, there exists $k \in \{1,\dots,N\}$ such that $E_{\pi_k}(\boldsymbol{f}) + \tilde{\alpha}_k < 0$.*

The following diagram summarizes the implications among different models for representing convex coherent sets found in this section.

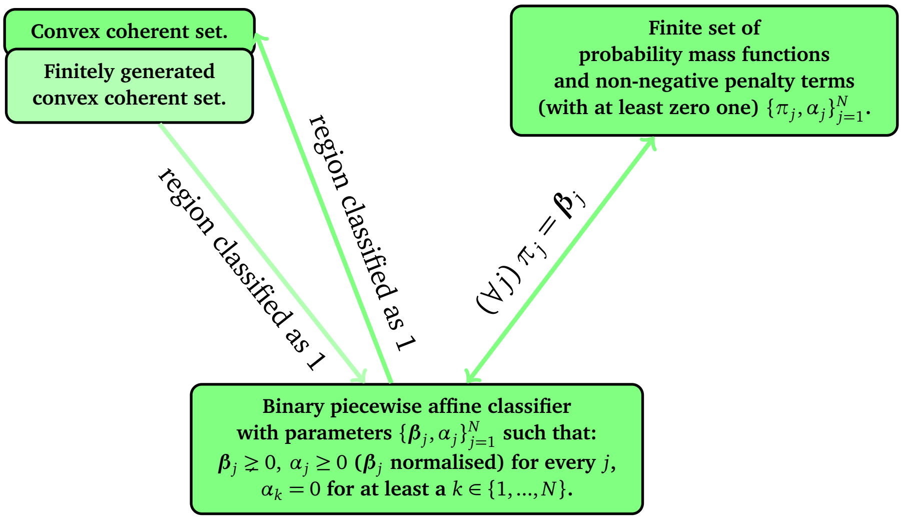

Figure 4.8. Diagram showing the implications among different models for representing convex coherent sets of almost desirable gambles (finitely generated and more general ones).

The following example illustrates the results presented in this section considering a specific numerical case and reports a numerical optimization procedure to solve the classification problem specified above.

Example 15. *Consider again the same framework of Example 12 and Example 13. Suppose now that another agent, Diana, judges gambles on the basis of Claire's coherent lower prevision $(\forall \boldsymbol{f} \in \mathbb{R}^n)\ \underline{P}_{Claire}(\boldsymbol{f}) := \min\{E_{\{1/3,2/3\}}(\boldsymbol{f}), E_{\{2/3,1/3\}}(\boldsymbol{f})\}$, but with the further constraint to not lose more than one thousand dollars. So, she* *is characterised by the following convex coherent set of almost desirable gambles:*

$$\overline{\mathscr{D}}_{Diana} := \{\boldsymbol{f} \in \mathbb{R}^n : \underline{P}_{Claire}(\boldsymbol{f}) \geq 0\} \cap \{\boldsymbol{f} \in \mathbb{R}^n : \min_i f_i \geq -1\} =$$
$$= \{\boldsymbol{f} \in \mathbb{R}^n : \min\{E_{\{1/3,2/3\}}(\boldsymbol{f}), E_{\{2/3,1/3\}}(\boldsymbol{f}), f_1 + 1, f_2 + 1\} \geq 0\}.$$

She is then disposed to accept, for example, gambles in $A_{Diana} = \{[-1,2]^\top, [2,-1]^\top, [1,-0.5]^\top\}$ *and reject gambles in* $R_{Diana} = \{[-3,2]^\top, [1,-1]^\top, [4,-2]^\top\}$. *Notice that Diana is not disposed to accept, for example, the gamble* $\boldsymbol{f}' := [4,-2]^\top$ *that instead is acceptable for Claire. The reason is that it can lead to a loss greater than one thousand dollars. Notice moreover that* $\boldsymbol{f}' = 2[2,-1]^\top$ *where* $[2,-1]^\top \in A_{Diana}$. $\overline{\mathscr{D}}_{Diana}$ *therefore* does not *respect axiom* D3', *hence Diana does not judge rewards of gambles with a linear utility scale.*

Suppose now the modeller Bruce sets D1', D2', *(CNV),* D5' *as his basic rationality axioms. If Bruce receives, as information about Diana, only the sets* A_{Diana} *and* R_{Diana}, *he thinks she is rational.* (A_{Diana}, R_{Diana}) *is indeed piecewise affine separable through, at least, the binary piecewise affine classifier:*

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\, PAC_D(\boldsymbol{f}) := \begin{cases} 1 & \text{if } E_{\{1/3,2/3\}}(\boldsymbol{f}) \geq 0, E_{\{2/3,1/3\}}(\boldsymbol{f}) \geq 0, \\ & \quad E_{\{1,0\}}(\boldsymbol{f}) + 1 \geq 0, E_{\{0,1\}}(\boldsymbol{f}) + 1 \geq 0, \\ -1 & \text{otherwise,} \end{cases}$$

which identifies, as the region classified as 1, $\overline{\mathrm{ch}(A_{Diana} \cup T)} = \mathrm{ch}^+(A_{Diana} \cup \{\mathbf{0}\})$ *that corresponds to the minimal set of assumptions the modeller can made on the agent willingness to accept gambles. It is possible to notice that, in particular, it also corresponds to Diana's set of almost desirable gambles* $\overline{\mathscr{D}}_{Diana}$.

The following Figure 4.9 *shows gambles in* A_{Diana} *represented again as (blue) points in the plane* (f_1, f_2), *gambles in* R_{Diana} *represented as (red) triangles and, in blue, the region classified as* 1 *by the classifier* PAC_D, *i.e., Diana's set of almost desirable gambles.*

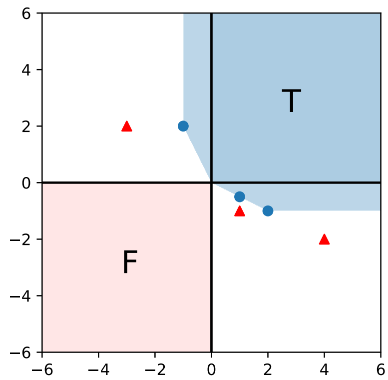

Figure 4.9. Gambles in A_{Diana}, R_{Diana} and the region classified as 1 by PAC_D.

The classifier PAC_D can be found numerically as follows: for any gamble $\boldsymbol{f}$, $PAC_D(\boldsymbol{f})$ is equal to 1 if the following linear programming problem is feasible, otherwise it is equal to -1.

$$LP = \begin{cases} \min_{\{\gamma_1,\ldots,\gamma_{|A_{Diana}\cup\{\mathbf{0}\}|},\lambda_1,\lambda_2\in\mathbb{R}\}} \lambda_1+\lambda_2 \\ s.t. \\ \boldsymbol{f} = \sum\limits_{\boldsymbol{f}_j\in A_{Diana}\cup\{\mathbf{0}\}} \gamma_j \boldsymbol{f}_j + [\lambda_1,\lambda_2]^\top, \\ \gamma_j \geq 0, \textit{ for every } j\in\{1,\ldots,|A_{Diana}\cup\{\mathbf{0}\}|\}, \\ \sum\limits_{j=1}^{|A_{Diana}\cup\{\mathbf{0}\}|} \gamma_j = 1, \\ \lambda_1,\lambda_2 \geq 0. \end{cases}$$

The classifier is contained in $PAC(A_{Diana}\cup T, R_{Diana}\cup F)$ provided that $A_{Diana}\cap F=\emptyset$, $PAC_D(\boldsymbol{f})=-1$ for each $\boldsymbol{f}\in R_{Diana}$ and $PAC_D([-\epsilon,-\epsilon]^T)=-1$ for some small ϵ. Note that, if $\boldsymbol{f}\in \mathrm{ch}(A_{Diana}\cup\{\mathbf{0}\})$ then $\lambda_1,\lambda_2=0$, otherwise either λ_1 or λ_2 must be different from zero.

4.3.1 Feature mapping

We can now reformulate the previous classification problem as a linear classification task in a higher dimensional space using a feature mapping similar to the one introduced in Section 4.2.1. We can indeed define new partitions $\{\mathscr{B}'_j\}_{j=1}^N$ of $\mathbb{R}^n$, where $\mathscr{B}'_j$ is defined as follows:

$$\mathscr{B}'_j := \{\boldsymbol{f}\in\mathbb{R}^n : \begin{bmatrix}\boldsymbol{f}\\1\end{bmatrix}^T \boldsymbol{\omega}'_j \leq \begin{bmatrix}\boldsymbol{f}\\1\end{bmatrix}^T \boldsymbol{\omega}'_k, \quad \text{for } k=1,\ldots,N, \quad j\neq k\} \tag{4.7}$$

with $\boldsymbol{\omega}'_j\in\mathbb{R}^{n+1}$ for every $j\in\{1,\ldots,N\}$. We can introduce the feature mapping $\boldsymbol{\psi}:\mathbb{R}^n\to\mathbb{R}^{(n+1)N}$, defined as $\boldsymbol{\psi}(\boldsymbol{f}) := [\boldsymbol{\psi}_1(\boldsymbol{f}),\ldots,\boldsymbol{\psi}_N(\boldsymbol{f})]^\top$ for every $\boldsymbol{f}\in\mathbb{R}^n$, where $\boldsymbol{\psi}_j:\mathbb{R}^n\to\mathbb{R}^{n+1}$ is defined in turn as:

$$\boldsymbol{\psi}_j(\boldsymbol{f}) := \begin{bmatrix}\mathbb{I}_{\mathscr{B}'_j}(\boldsymbol{f})f_1\\ \cdots \\ \mathbb{I}_{\mathscr{B}'_j}(\boldsymbol{f})f_n \\ \mathbb{I}_{\mathscr{B}'_j}(\boldsymbol{f})\end{bmatrix} \tag{4.8}$$

for every $\boldsymbol{f} \in \mathbb{R}^n$ and $j \in \{1,\ldots,N\}$. Further, we define the following classifier corresponding to a linear classifier in the feature space:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\; LC_{\psi}(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \sum_{j=1}^{N} \boldsymbol{\psi}_j(\boldsymbol{f})^{\top}\boldsymbol{\beta}'_j \geq 0, \\ -1 & \text{otherwise,} \end{cases} \tag{4.9}$$

with $\boldsymbol{\beta}'_j \in \mathbb{R}^{n+1}$ for all $j = 1,\ldots,N$. We consider both $\{\boldsymbol{\beta}'_j\}_{j=1}^N$ and $\{\boldsymbol{\omega}'_j\}_{j=1}^N$ as parameters of LC_{ψ}. We can then introduce the following definition.

Definition 59 ($\boldsymbol{\Psi}$-separability)**.** *A pair of sets of gambles (A,B) is $\boldsymbol{\Psi}$-*separable *if and only if there exists a classifier LC_{ψ} of type (4.9) such that $LC_{\psi}(A) = 1$ and $LC_{\psi}(B) = -1$. We indicate the set of these classifiers with $LC_{\boldsymbol{\Psi}}(A,B)$.*

We can now state the main result of this subsection.

Proposition 12. *A binary piecewise affine classifier with parameters $\{\boldsymbol{\beta}_j,\alpha_j\}_{j=1}^N$ and a classifier of type (4.9) with parameters $\{\boldsymbol{\omega}'_j,\boldsymbol{\beta}'_j\}_{j=1}^N$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}'_j = \begin{bmatrix}\boldsymbol{\beta}_j \\ \alpha_j\end{bmatrix}$ for every $j \in \{1,..,N\}$ classify gambles in the same way.*

The proof of Proposition 12 is analogous to the one of Proposition 9. The following corollary also holds.

Corollary 11. *Given a pair of finite sets of gambles (A,R), there exists a convex coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, if and only if (A,R) is $\boldsymbol{\Psi}$–separable and there exists a classifier $LC_{\psi} \in LC_{\boldsymbol{\Psi}}(A,R)$, with parameters $\{\boldsymbol{\omega}'_j,\boldsymbol{\beta}'_j\}_{j=1}^N$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}'_j \geq 0$ and $(\boldsymbol{\beta}'_j)^T\boldsymbol{e}_i > 0$ for every $j \in \{1,...,N\}$ and some $i \in \{1,...,n\}$, with at least a $\boldsymbol{\beta}'_k$ with null $n+1$-th component.*

Similarly to before, we have that, given a classifier LC_{ψ} of the type specified by Corollary 11, the set $\overline{\mathscr{D}} := \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\psi}(\boldsymbol{f}) = 1\}$ is always convex coherent. Vice versa, if $\overline{\mathscr{D}}$ is a finitely generated convex coherent set, we can always find a classifier LC_{ψ} of the type specified by Corollary 11 such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\psi}(\boldsymbol{f}) = 1\}$.

Example 16. *Consider again the framework of Example 15. In this context, let us construct the classifier LC_{ψ}^D of type (4.9) with parameters $\{\boldsymbol{\omega}'_j,\boldsymbol{\beta}'_j\}_{j=1}^4$ such that $\boldsymbol{\beta}'_j = \boldsymbol{\omega}'_j = \left[\begin{smallmatrix}\boldsymbol{\beta}_j \\ \alpha_j\end{smallmatrix}\right]$ for every j, where $\{\boldsymbol{\beta}_j,\alpha_j\}_{j=1}^4$ are the parameters of PAC_D introduced in Example 15:*

$$\boldsymbol{\beta}_1 := \begin{bmatrix}1/3 \\ 2/3\end{bmatrix}, \boldsymbol{\beta}_2 := \begin{bmatrix}2/3 \\ 1/3\end{bmatrix}, \boldsymbol{\beta}_3 := \begin{bmatrix}1 \\ 0\end{bmatrix}, \boldsymbol{\beta}_4 := \begin{bmatrix}0 \\ 1\end{bmatrix},$$

$$\alpha_1 = \alpha_2 = 0, \alpha_3 = \alpha_4 = 1.$$

By Proposition 12, we know that LC^D_ψ is a reformulation of PAC_D. It therefore classifies A_{Diana} as 1 and R_{Diana} as -1. Moreover, $\{\boldsymbol{f} \in \mathbb{R}^n : LC^D_\psi(\boldsymbol{f}) = 1\} = \overline{\mathrm{ch}(A_{Diana} \cup T)}$.

4.4 Positive additive coherence

We now consider an even weaker relaxation of the linearity axioms D3' and D4':[14]

(PADD) $(\forall \boldsymbol{f} \in \mathbb{R}^n)\ \boldsymbol{f} \geq \boldsymbol{g},\ \boldsymbol{g} \in \overline{\mathscr{D}} \Rightarrow \boldsymbol{f} \in \overline{\mathscr{D}}$ [Positive Additivity].

It can be considered the weakest axiom to add to D1',D2' and D5'. Indeed, it forces an agent to find almost desirable only those gambles that are 'better' than the ones they already find almost desirable, i.e., which can provide more gains and/or less losses.

We call a set $\overline{\mathscr{D}}$ satisfying D1', D2', (PADD), D5' a *positive additive coherent* set of almost desirable gambles.

Analogously to the previous sections, a modeller assuming these as their basic rationality axioms regards an agent, providing only two finite sets A, R of respectively acceptable and rejectable gambles, as rational if and only if there exists a positive additive coherent set of gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. If this is the case, the minimal such set is the *principal up-set* of $(A \cup \{\mathbf{0}\})$ [Davey and Priestley, 2002, Section 1.27]: $\uparrow (A \cup \{\mathbf{0}\}) := \{\boldsymbol{f} \in \mathbb{R}^n : (\exists \boldsymbol{g} \in A \cup \{\mathbf{0}\})\ \boldsymbol{f} \geq \boldsymbol{g}\}$.[15] It is possible to notice that, geometrically, the latter corresponds to a union of closed orthants centered in the elements of $A \cup \{\mathbf{0}\}$.

Analogously to the previous section, we can give the following definition.

Definition 60 (**Finitely generated positive additive coherent set**). *Let us consider a positive additive coherent set of almost desirable gambles $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$. If $\overline{\mathscr{D}} = \uparrow (A \cup \{\mathbf{0}\})$ for some finite set $A \subseteq \mathbb{R}^n$, we say that $\overline{\mathscr{D}}$ is* finitely generated.

As usual, in the rest of the section we show that a way the modeller has to check the rationality of the agent is by solving a binary classification task.

Let us introduce the following definitions.

Definition 61 (**PWP classifier**). *We use the term* binary piecewise positive affine (PWP) classifier *to denote a classifier $PWPC$ defined as follows:*

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ PWPC(\boldsymbol{f}) := \begin{cases} c_1 & \text{if } \exists \boldsymbol{g}^j \in \mathscr{G} \text{ s.t. } \boldsymbol{f} \geq \boldsymbol{g}^j, \\ c_2 & \text{otherwise,} \end{cases} \tag{4.10}$$

[14] See Lemma 26 in Appendix D.
[15] See Lemma 29 in Appendix D.

with labels c_1, c_2 *and with* $\mathscr{G} := \{\boldsymbol{g}^j\}_{j=1}^N$*, a finite set of gambles (generators of the classifier).*

Again, without loss of generality, we assume $c_1 = 1$ and $c_2 = -1$.

Definition 62 (PWP separability). *A pair of sets of gambles* (A, B) *is* piecewise positive affine separable *(PWP) if and only if there exists a PWP classifier* $PWPC$ *such that* $PWPC(A) = 1$ *and* $PWPC(B) = -1$*. We indicate the set of these classifiers with* PWPC(A, B)*.*

Note that, for every j, $\{\boldsymbol{f} \geq \boldsymbol{g}^j\}$ defines an orthant centered at $\boldsymbol{g}^j$, whose border can be expressed as a piecewise affine function. It can easily be proved (by induction on the elements of $\mathscr{G}$) that the decision boundary of (4.10) is also a piecewise affine function.

We can now state the main result of this section.

Proposition 13. *Given a pair of finite sets of gambles* (A, R)*, there exists a positive additive coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$*, such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if* $(A \cup T, R \cup F)$ *is PWP separable.*

In particular, if there exists a finitely generated positive additive coherent set $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, it is possible to construct a classifier $PWPC \in$ PWPC$(A \cup T, R \cup F)$ such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PWPC(\boldsymbol{f}) = 1\}$. Vice versa, if there exists a classifier $PWPC \in$ PWPC$(A \cup T, R \cup F)$, the region $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : PWPC(\boldsymbol{f}) = 1\}$ is a finitely generated positive additive coherent set such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Notice moreover that if there exists a positive additive coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, the smallest such set is $\uparrow (A \cup \{0\})$, which is finitely generated. PWP classifiers $PWPC$ such that $\{\boldsymbol{f} \in \mathbb{R}^n : PWPC(\boldsymbol{f}) = 1\} = \uparrow (A \cup \{0\})$, in particular, identify the minimal 1-region containing A. Hence, they correspond to make the minimal assumptions about the agent's dispositions to accept gambles.

Once again, we can limit ourselves to classify only A and R.

Proposition 14. *Consider a pair of finite sets of gambles* (A, R)*. Every classifier* $PWPC \in$ *PWPC*$(A \cup T, R \cup F)$ *is a classifier* $PWPC \in$ *PWPC*(A, R) *with generators* $\{\boldsymbol{g}^j\}_{j=1}^N$ *such that* $\boldsymbol{g}^j \not< 0$ *for every* $j \in \{1, \ldots, N\}$ *with at least a* $\boldsymbol{g}^k \leq 0$ *(but* $\boldsymbol{g}^k \not< 0$*) for some* $k \in \{1, \ldots, N\}$*, and vice versa.*

Corollary 12. *Given a pair of finite sets of gambles* (A, R)*, there exists a positive additive coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ *such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if* (A, R) *is PWP separable and there exists a classifier* $PWPC \in$ *PWPC*(A, R) *with generators* $\{\boldsymbol{g}^j\}_{j=1}^N$ *such that* $\boldsymbol{g}^j \not< 0$ *for every* $j \in \{1, \ldots, N\}$ *with at least a* $\boldsymbol{g}^k \leq 0$ *(but* $\boldsymbol{g}^k \not< 0$*) for some* $k \in \{1, \ldots, N\}$*.*

The following diagram summarizes the implications among different models for representing (finitely generated) positive additive coherent sets found in this section.

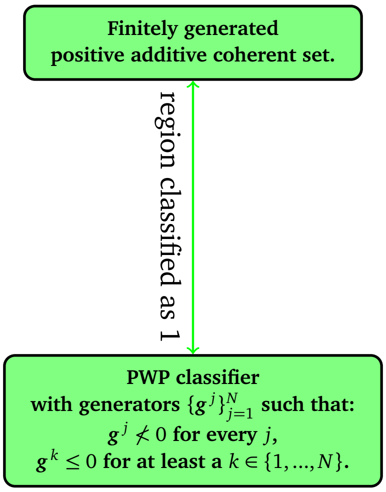

Figure 4.10. Diagram showing equivalent models for representing finitely generated positive additive coherent sets of almost desirable gambles.

The following example considers a specific numerical case and reports a numerical optimization procedure to solve the classification problem.

Example 17. *Consider again the framework of Example 12. In this context, consider also an agent Emma who is disposed to accept gambles in* $A_{Emma} = \{[-1, 2]^\top, [2, -1]^\top\}$ *and reject gambles in* $R_{Emma} = \{[-3, 2]^\top, [1, -1]^\top, [4, -2]^\top, [1, -0.5]^\top\}$.

If the modeller Bruce receives only A_{Emma} *and* R_{Emma} *as information about Emma, he can check if at least there exists a set respecting her will and the minimal rationality axioms D1',D2', (PADD),D5'. That is indeed the case since* (A_{Emma}, R_{Emma}) *is PWP separable through the PWP classifier* $PWPC_E$*:*

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ PWPC_E(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \exists \boldsymbol{g} \in A_{Emma} \cup \{\boldsymbol{0}\} \text{ s.t. } \boldsymbol{f} \geq \boldsymbol{g}, \\ -1 & \text{otherwise,} \end{cases} \tag{4.11}$$

which identifies, as the region classified as 1*, the set* $\uparrow(A_{Emma} \cup \{\boldsymbol{0}\})$*, i.e., the minimal set of assumptions the modeller can made on the agent's willingness to accept gambles.*

The following figure shows gambles in A_{Emma}*, represented again as (blue) points in the plane* (f_1, f_2)*, gambles in* R_{Emma}*, represented as (red) triangles, and, in blue, the region classified as* 1 *by the classifier* $PWPC_E$*. The numerical implementation* *of this classifier is straightforward:* $PWPC_E(\boldsymbol{f}) = 1$ *if there exists a gamble* $\boldsymbol{g}$ *in* $A_{Emma} \cup \{\boldsymbol{0}\}$ *such that* $\boldsymbol{f} \geq \boldsymbol{g}$ *and* -1 *otherwise, for every* $\boldsymbol{f} \in \mathbb{R}^n$.

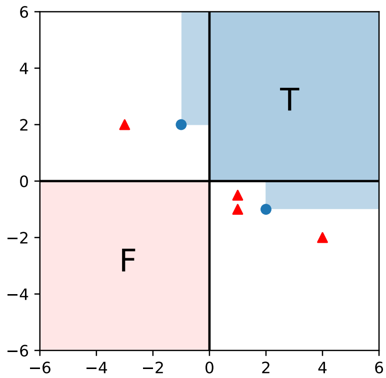

Figure 4.11. Gambles in A_{Emma}, R_{Emma} and the region classified as 1 by $PWPC_E$.

4.4.1 Feature mapping

Also in this case, the previous classification problem can be reformulated as a linear one in a higher dimensional space. Let $\{\zeta_{i,j}\}_{\{i\in\{1,\dots,n\},j\in\{1,\dots,N\}\}}$ denote another partition of $\mathbb{R}^n$, where $\zeta_{i,j}$ is defined as follows:

$$\zeta_{i,j} := \{\boldsymbol{f} \in \mathbb{R}^n : (f_i - \omega_i^j) = \max_{\{k\in\{1,\dots,N\}\}} (\min_{\{l\in\{1,\dots,n\}\}} (f_l - \omega_l^k))\} \tag{4.12}$$

with $\boldsymbol{\omega}^j \in \mathbb{R}^n$, for every i, j. We can introduce the feature mapping $\boldsymbol{\rho} : \mathbb{R}^n \to \mathbb{R}^{2nN}$, defined as $\boldsymbol{\rho}(\boldsymbol{f}) := [\boldsymbol{\rho}_1(\boldsymbol{f}), \dots, \boldsymbol{\rho}_N(\boldsymbol{f})]^\top$ for every $\boldsymbol{f} \in \mathbb{R}^n$, where $\boldsymbol{\rho}_j : \mathbb{R}^n \to \mathbb{R}^{2n}$ is defined in turn as:

$$\boldsymbol{\rho}_j(\boldsymbol{f}) := \begin{bmatrix} \mathbb{I}_{\zeta_{1,j}}(\boldsymbol{f}) f_1 \\ \dots \\ \mathbb{I}_{\zeta_{n,j}}(\boldsymbol{f}) f_n \\ \mathbb{I}_{\zeta_{1,j}}(\boldsymbol{f}) \\ \dots \\ \mathbb{I}_{\zeta_{n,j}}(\boldsymbol{f}) \end{bmatrix} \tag{4.13}$$

for every $\boldsymbol{f} \in \mathbb{R}^n$ and $j \in \{1, \dots, N\}$. Further, we define the following classifier corresponding to a linear classifier in the feature space:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ LC_\rho(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \sum_{j=1}^N \boldsymbol{\rho}_j(\boldsymbol{f})^\top \boldsymbol{\beta}'_j \geq 0, \\ -1 & \text{otherwise,} \end{cases} \tag{4.14}$$

with $\boldsymbol{\beta}'_j \in \mathbb{R}^{2n}$ for all $j = 1, \dots, N$. We consider both $\{\boldsymbol{\beta}'_j\}_{j=1}^N$ and $\{\boldsymbol{\omega}^j\}_{j=1}^N$ as parameters of LC_ρ. Similarly to before, we can introduce the following definition.

Definition 63 (P-separability)**.** *A pair of sets of gambles* (A, B) *is* $\boldsymbol{P}$-separable *if and only if there exists a classifier* $LC_{\boldsymbol{\rho}}$ *of type* (4.14) *such that* $LC_{\boldsymbol{\rho}}(A) = 1$ *and* $LC_{\boldsymbol{\rho}}(B) = -1$*. We indicate the set of these classifiers with* $LC_{\boldsymbol{P}}(A, B)$*.*

We can now state the main result of this subsection.

Proposition 15. *A PWP classifier characterised by a set of generators* $\mathscr{G} = \{\boldsymbol{g}^j\}_{j=1}^N$ *and a classifier of type* (4.14) *with parameters* $\{\boldsymbol{\omega}^j, \boldsymbol{\beta}'_j\}_{j=1}^N$ *such that* $\boldsymbol{\beta}'_j = [1, \ldots, 1, -\omega_1^j, \ldots, -\omega_n^j]^\top$ *and* $\boldsymbol{\omega}^j = \boldsymbol{g}^j$ *for every* $j \in \{1, ..., N\}$ *classify gambles in the same way.*

The proof is based on the following observation, analogous to the previous ones. Given a PWP classifier $PWPC$ with generators $\{\boldsymbol{g}^j\}_{j=1}^N$:

$$PWPC(\boldsymbol{f}) = 1 \iff \max_j(\min_i(f_i - g_i^j)) \geq 0$$

$$\iff \sum_{j=1}^N \left(\begin{bmatrix} \mathbb{I}_{\zeta_{1,j}}(\boldsymbol{f}) f_1 \\ \cdots \\ \mathbb{I}_{\zeta_{n,j}}(\boldsymbol{f}) f_n \\ \mathbb{I}_{\zeta_{1,j}}(\boldsymbol{f}) \\ \cdots \\ \mathbb{I}_{\zeta_{n,j}}(\boldsymbol{f}) \end{bmatrix} \right)^\top \begin{bmatrix} 1 \\ \cdots \\ 1 \\ -g_1^j \\ \cdots \\ -g_n^j \end{bmatrix} \geq 0,$$

where $\zeta_{i,j} := \{\boldsymbol{f} \in \mathbb{R}^n : (f_i - g_i^j) = \max_k(\min_l(f_l - g_l^k))\}$, for all i, j.

Corollary 13. *Given a pair of finite sets of gambles* (A, R)*, there exists a positive additive coherent set of almost desirable gambles* $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ *such that* $\overline{\mathscr{D}} \supseteq A$ *and* $\overline{\mathscr{D}} \cap R = \emptyset$*, if and only if* (A, R) *is* $\boldsymbol{P}$*−separable and there exists a classifier* $LC_{\boldsymbol{\rho}} \in LC_{\boldsymbol{P}}(A, R)$ *with parameters* $\{\boldsymbol{\omega}^j, \boldsymbol{\beta}'_j\}_{j=1}^N$ *such that* $\boldsymbol{\beta}'_j = [1, \ldots, 1, -\omega_1^j, \ldots, -\omega_n^j]^\top$ *with* $\boldsymbol{\omega}^j \not< 0$ *for all* $j \in \{1, ..., N\}$*, and with at least a* $\boldsymbol{\omega}^k$ *such that* $\boldsymbol{\omega}^k \leq 0$ *(but* $\boldsymbol{\omega}^k \not< 0$*) for some* $k \in \{1, .., N\}$*.*

In particular, if $\overline{\mathscr{D}}$ is a finitely generated positive additive coherent set we can find $LC_{\boldsymbol{\rho}}$ of the type specified in Corollary 13 such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\boldsymbol{\rho}}(\boldsymbol{f}) = 1\}$ and vice versa.

Example 18. *Consider again the framework of Example 17. In this context, let us construct the classifier* $LC_{\boldsymbol{\rho}}^E$ *of type* (4.14) *with parameters* $\{\boldsymbol{\omega}^j, \boldsymbol{\beta}'_j\}_{j=1}^3$ *such that* $\boldsymbol{\beta}'_j = [1, 1, -\omega_1^j, -\omega_2^j]^\top$ *for every* j *with* $\boldsymbol{\omega}^1 = \left[\begin{smallmatrix} -1 \\ 2 \end{smallmatrix}\right]$*,* $\boldsymbol{\omega}^2 = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$*,* $\boldsymbol{\omega}^3 = \left[\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right]$ *(i.e,* $\{\boldsymbol{\omega}^j\}_{j=1}^3 = (A_{Emma} \cup \{\boldsymbol{0}\})$*). By Proposition 15, we know that* $LC_{\boldsymbol{\rho}}^E$ *is a reformulation of* $PWPC_E$*. It therefore classifies* A_{Emma} *as* 1 *and* R_{Emma} *as* -1*. Moreover,* $\{\boldsymbol{f} \in \mathbb{R}^n : LC_{\boldsymbol{\rho}}^E(\boldsymbol{f}) = 1\} = \uparrow(A_{Emma} \cup \{\boldsymbol{0}\})$*.*

4.5 Lower previsions

In this section we analyze properties satisfied by a lower prevision over gambles induced by sets of gambles satisfying each of the sets of axioms considered in the previous sections.

Before starting with this analysis, we recall in the following diagram the existing relations among the different sets of rationality axioms considered. In particular, it is important to notice that {D1',D2',(PADD),D5'} corresponds to the weakest set of axioms considered. Therefore, maximal coherent sets, coherent and convex coherent ones are in particular positive additive coherent sets.

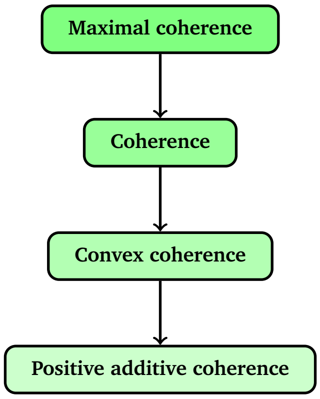

Figure 4.12. A graphical summary of the existing relations among the different concepts of coherence analysed.

Let us consider a positive additive coherent set of almost desirable gambles $\overline{\mathcal{D}}$. Analogously to what is done in Section 1.1.1 for coherent sets of almost desirable gambles, we can associate to $\overline{\mathcal{D}}$ the lower prevision $\underline{P} := \sigma(\overline{\mathcal{D}})$ given by

$$(\forall \boldsymbol{f} \in dom(\underline{P}))\ \underline{P}(\boldsymbol{f}) := \sup\{\mu \in \mathbb{R} : \boldsymbol{f} - \mu \in \overline{\mathcal{D}}\}. \tag{4.15}$$

It satisfies, in particular, the following properties.

Proposition 16. *Consider a set of gambles $\overline{\mathcal{D}} \subseteq \mathbb{R}^n$ satisfying D1',D2',(PADD),D5'. The lower prevision $\underline{P} := \sigma(\overline{\mathcal{D}})$ it induces by means of Eq. (4.15) satisfies the following properties.*

- $dom(\underline{P}) = \mathbb{R}^n$;
- *for any* $\boldsymbol{f} \in \mathbb{R}^n$, $\underline{P}(\boldsymbol{f}) \in \mathbb{R}$;
- *for any* $\boldsymbol{f} \in \mathbb{R}^n$, $\min_i f_i \leq \underline{P}(\boldsymbol{f}) \leq \max_i f_i$;

- *for any* $\boldsymbol{f} \in \mathbb{R}^n$, $\underline{P}(\boldsymbol{f} - \underline{P}(\boldsymbol{f})) = 0$;
- *for any* $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^n$, $\boldsymbol{f} \geq \boldsymbol{g} \Rightarrow \underline{P}(\boldsymbol{f}) \geq \underline{P}(\boldsymbol{g})$ *[Monotonicity]*;
- *for any* $\boldsymbol{f} \in \mathbb{R}^n$ *and* $r \in \mathbb{R}$, $\underline{P}(\boldsymbol{f} + r) = \underline{P}(\boldsymbol{f}) + r$ *[Translation Invariance]*;
- $\underline{P}(\mathbf{0}) = 0$.

If $\overline{\mathscr{D}}$ *also satisfies (CNV), then* $\underline{P} := \sigma(\overline{\mathscr{D}})$ *satisfies:*

- *for any* $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^n$ *and* $\gamma \in [0,1]$, $\underline{P}(\gamma \boldsymbol{f} + (1-\gamma)\boldsymbol{g}) \geq \gamma \underline{P}(\boldsymbol{f}) + (1-\gamma)\underline{P}(\boldsymbol{g})$ *[Concavity]*.

If $\overline{\mathscr{D}}$ *also satisfies* D3' *and* D4', *then* $\underline{P} := \sigma(\overline{\mathscr{D}})$ *satisfies:*

- *for any* $\boldsymbol{f} \in \mathbb{R}^n$ *and* $\lambda > 0$, $\underline{P}(\lambda \boldsymbol{f}) = \lambda \underline{P}(\boldsymbol{f})$ *[Positive Homogeneity]*;
- *for any* $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^n$, $\underline{P}(\boldsymbol{f} + \boldsymbol{g}) \geq \underline{P}(\boldsymbol{f}) + \underline{P}(\boldsymbol{g})$ *[Superlinearity]*.

If $\overline{\mathscr{D}}$ *is in particular a maximal coherent set of almost desirable gambles, then* $\underline{P} = \sigma(\overline{\mathscr{D}})$ *also satisfies:*

- *for any* $\boldsymbol{f} \in \mathbb{R}^n$, $\underline{P}(\boldsymbol{f}) = -\underline{P}(-\boldsymbol{f})$ *[Self-Conjugacy]*.

Notice that, in particular, if $\overline{\mathscr{D}}$ satisfies the standard axioms D1'–D5', $\underline{P}$ is a coherent lower prevision, if it is a maximal coherent set, $\underline{P}$ is linear. If instead $\overline{\mathscr{D}}$ satisfies only D1',D2',(PADD),D5', $\underline{P}$ is a *centered 2-convex lower prevision* and if $\overline{\mathscr{D}}$ satisfies also (CNV), $\underline{P}$ is a *centered convex lower prevision*. Centered convex and centered 2-convex lower previsions are extensions of the concept of coherent lower prevision proposed in Pelessoni and Vicig [2016]. These types of lower previsions have important applications. In Pelessoni and Vicig, 2005 indeed, it is shown that the former can be employed to formulate many commonly used *convex risk measures*. *Risk measures* are functions ρ measuring how 'risky' a random variable X is and whether it is acceptable to buy or hold it. In essence, X is considered acceptable (or unacceptable) when $\rho(X) \leq 0$ (when $\rho(X) > 0$), and $\rho(X)$ should represent the maximum amount that can be deducted from X while still keeping it acceptable (the minimum amount that must be added to X to render it acceptable). This connection establishes a strong link with imprecise probability models. Indeed, in Pelessoni and Vicig, 2003, it is established that a risk measure for X can be interpreted as an upper prevision for $-X$. *Convex risk measures* are, in particular, special types of risk measures often preferred in a risk measurement environment because they permit to not assume positive homogeneity [Föllmer and Schied, 2002]. A motivation for not assuming this property indeed is that $\rho(\lambda X)$ may be larger than $\lambda\rho(X)$ for $\lambda > 1$ because of the *liquidity risks*: if one were to immediately sell a substantial quantity λX of a financial investment, they might be compelled to accept a smaller reward than λ times the current selling price for X. *Centered 2-convex upper previsions*, i.e., conjugates of centered 2-convex lower previsions, are instead generalizations of coherent upper previsions that, when interpreted as risk measures, encompass the well-known *Value-at-risk*, probably the widespread risk measure known in the literature [Pelessoni and Vicig, 2016].

The following example illustrates some of the results of this section.

Example 19. *Consider the following sets of almost desirable gambles introduced in the previous examples:*

- $\overline{M}_{Anne}$:=:= $\{\boldsymbol{f} \in \mathbb{R}^n : E_{\{2/3,1/3\}}(\boldsymbol{f}) \geq 0\}$ *of Example 12;*
- $\overline{\mathscr{D}}_{Claire} := \{\boldsymbol{f} \in \mathbb{R}^n : \min\{E_{\{1/3,2/3\}}(\boldsymbol{f}), E_{\{2/3,1/3\}}(\boldsymbol{f})\} \geq 0\}$ *of Example 13;*
- $\overline{\mathscr{D}}_{Diana} := \{\boldsymbol{f} \in \mathbb{R}^n : \min\{E_{\{1/3,2/3\}}(\boldsymbol{f}), E_{\{2/3,1/3\}}(\boldsymbol{f}), f_1 + 1, f_2 + 1\} \geq 0\}$ *of Example 15;*
- $\overline{\mathscr{D}}_{Emma} :=\uparrow (A_{Emma} \cup \{\mathbf{0}\}) = \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f} \geq [-1,2]^\top$ *or* $\boldsymbol{f} \geq [0,0]^\top$ *or* $\boldsymbol{f} \geq [2,-1]^\top\}$ *of Example 17.*

From each one of these sets, we can induce a lower prevision.

Consider the gambles $\mathbf{0} = [0,0]^\top$, $\boldsymbol{g} = [2,-1]^\top$ *and* $\boldsymbol{h} = [-1,0]^\top$. *The following table summarizes the values of the four lower previsions* $P_{Anne} := \sigma(\overline{M}_{Anne})$, $\underline{P}_{Claire} := \sigma(\overline{\mathscr{D}}_{Claire})$, $\underline{P}_{Diana} := \sigma(\overline{\mathscr{D}}_{Diana})$, $\underline{P}_{Emma} := \sigma(\overline{\mathscr{D}}_{Emma})$ *for* $\mathbf{0}$, $\boldsymbol{g}, \boldsymbol{h}$, $\frac{1}{2}\boldsymbol{g}$, $2\boldsymbol{g}, \boldsymbol{g}+\boldsymbol{h}$.

	P_{Anne}	$\underline{P}_{Claire}$	$\underline{P}_{Diana}$	$\underline{P}_{Emma}$
$\mathbf{0}$	0	0	0	0
$\boldsymbol{g}$	1.34	0	0	0
$\boldsymbol{h}$	−0.89	−0.89	−0.89	−1
$\frac{1}{2}\boldsymbol{g}$	0.67	0	0	−0.5
$2\boldsymbol{g}$	2.68	0	−1	−1
$\boldsymbol{g}+\boldsymbol{h}$	0.45	−0.45	−0.45	−1

From these values, it is possible to notice that $\underline{P}_{Emma}(\frac{1}{2}\boldsymbol{g}) \not\geq \frac{1}{2}\underline{P}_{Emma}(\boldsymbol{g})$, *hence it is not even a concave functional.* $\underline{P}_{Diana}$ *does not respect positive homogeneity with* $\lambda > 1$ *since* $\underline{P}_{Diana}(2\boldsymbol{g}) \neq 2\underline{P}_{Diana}(\boldsymbol{g})$. *Only* P_{Anne} *and* $\underline{P}_{Claire}$ *are standard coherent lower previsions. Moreover,* P_{Anne} *is also a linear one.*

4.6 General coherence

Consider now a new agent, Florence, again providing a finite set of acceptable gambles A and a finite set of rejectable gambles R. In the previous sections we have analysed different examples of axiomatisations of desirability and show that they can be formulated as classification problems. We would now like to outline a general framework for modelling coherence starting directly from this alternative reformulations.

Consider a general feature mapping $\hat{\boldsymbol{\phi}} : \mathbb{R}^n \to \mathbb{R}^{MN}$ with $M \geq n$ and $N \geq 1$, such that $\hat{\boldsymbol{\phi}} := [\hat{\boldsymbol{\phi}}_1, \dots, \hat{\boldsymbol{\phi}}_N]^\top$, where $\hat{\boldsymbol{\phi}}_j : \mathbb{R}^n \to \mathbb{R}^M$ for every $j \in \{1, ..., N\}$.

Starting from this feature mapping, we can define the following classifier, which corresponds to a linear classifier in the feature space:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n) \; LC_{\hat{\boldsymbol{\phi}}}(\boldsymbol{f}) := \begin{cases} 1 & \text{if } \sum_{j=1}^{N} \hat{\boldsymbol{\phi}}_j(\boldsymbol{f})^\top \boldsymbol{\beta'}_j \geq 0, \\ -1 & \text{otherwise,} \end{cases} \tag{4.16}$$

with $\boldsymbol{\beta}'_j \in \mathbb{R}^M$ for every $j \in \{1, ..., N\}$.

We then consider sets separable through this general type of classifier.

Definition 64 ($\hat{\Phi}$-separability)**.** *A pair of sets of gambles (A, B) is* $\hat{\boldsymbol{\Phi}}$-separable *if and only if there exists a classifier $LC_{\hat{\boldsymbol{\phi}}}$ of type (4.16), such that $LC_{\hat{\boldsymbol{\phi}}}(A) = 1$ and $LC_{\hat{\boldsymbol{\phi}}}(B) = -1$. We indicate the set of these classifiers with $LC_{\hat{\boldsymbol{\Phi}}}(A, B)$.*

From these elements we can derive a more general definition of coherence.

Definition 65. *Consider a general feature mapping $\hat{\boldsymbol{\phi}} : \mathbb{R}^n \to \mathbb{R}^{MN}$ with $M \geq n$ and $N \geq 1$. A set $\overline{\mathscr{D}} \subseteq \mathbb{R}^n$ is said to be a* $\hat{\boldsymbol{\phi}}$-coherent *set of almost desirable gambles if and only if there exists a classifier $LC_{\hat{\boldsymbol{\phi}}}$ constructed from $\hat{\boldsymbol{\phi}}$ through (4.16), such that $\overline{\mathscr{D}} = \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\hat{\boldsymbol{\phi}}}(\boldsymbol{f}) = 1\}$.*

By selecting the functions $\hat{\boldsymbol{\phi}}_j$ in (4.16) as in (4.4), (4.8), (4.13), and imposing conditions on the parameters of the classifier as in Proposition 9, 12 and 15, we obtain the piecewise linear, piecewise affine and, respectively, orthant-based classifers discussed in the previous sections whose coherence-notion has been linked to the desirability axioms D1',D2',D3',D4',(CNV),(PADD) and D5'. Choosing instead the identity function as a feature mapping, we obtain the linear classifiers discussed in Section 4.1.

In the more general setting treated in this section, we can provide sufficient conditions to guarantee that a $\hat{\boldsymbol{\phi}}$-coherent set $\overline{\mathscr{D}}$ satisfy at least the axioms D1',D2',D5'that we consider to be common to every notion of coherence considered.

1. D5'. A sufficient condition to guarantee that $\overline{\mathscr{D}}$ satisfies D5' is that $\hat{\boldsymbol{\phi}}$ is a continuous function. This is the case treated in the following Example 20. This condition is not necessary. Indeed it is not satisfied when $\hat{\boldsymbol{\phi}}$ coincides with one of the feature mappings considered in the previous Sections 4.2.1, 4.3.1 and 4.4.1. However, in all those cases, D5' is always satisfied.

2. D1' and D2'. To guarantee that D1' and D2' are satisfied is sufficient to ask that the classifier $LC_{\hat{\phi}}$ classifies T as 1 and F as -1. Since these are infinite sets, it can be useful to constrain the possible feature mappings in such a way that this task boils down to conditions to be required to the parameters of the classifier, as it is discussed in Example 20. This can be more easily achieved considering particular bases functions for the feature mappings, such as exponential bases or odd polynomial functions.

Example 20. *Consider, as before, a possibility space of size $n = 2$. Then, every gamble $\boldsymbol{f}$ has two components f_1 and f_2. Consider the feature mapping $\hat{\boldsymbol{\phi}} : \mathbb{R}^2 \to \mathbb{R}^3$ defined as:*

$$\hat{\boldsymbol{\phi}}(\boldsymbol{f}) = \begin{bmatrix} -e^{-d_1 f_1} \\ -e^{-d_2 f_2} \\ 1 \end{bmatrix} \tag{4.17}$$

for every $\boldsymbol{f} \in \mathbb{R}^n$, where $d_i \geq 0$ is a constant (the parameter of the classifier) for every $i \in \{1,2\}$. We consider then the classifier defined as:

$$(\forall \boldsymbol{f} \in \mathbb{R}^n)\ LC_{\hat{\phi}}(\boldsymbol{f}) := \begin{cases} 1 & \text{if } -\beta_1' e^{-d_1 f_1} - \beta_2' e^{-d_2 f_2} + \beta_3' \geq 0, \\ -1 & \text{otherwise}, \end{cases} \tag{4.18}$$

We constrain $\beta_1', \beta_2' \geq 0$, $\beta_1' + \beta_2' = 1$ and $\beta_3' = 1$ so that $LC_{\hat{\phi}}(T) = 1$ and $LC_{\hat{\phi}}(F) = -1$. Therefore, for $\overline{\mathscr{D}} := \{\boldsymbol{f} \in \mathbb{R}^n : LC_{\hat{\phi}}(\boldsymbol{f}) = 1\}$, D1',D2' hold and D5' follows by the continuity of the function $\hat{\boldsymbol{\phi}}$.

In this context suppose that Diane, the agent previously considered in Example 15, provides two finite sets of gambles A and R that are respectively acceptable and rejectable for her. Suppose moreover that they are $\hat{\boldsymbol{\phi}}$-separable through a classifier $LC_{\hat{\phi}}$ as defined in (4.18). Among the classifiers that satisfy $LC_{\hat{\phi}}(A) = 1$ and $LC_{\hat{\phi}}(R) = -1$, we select the one that minimises the objective function:

$$\sum_{f^j \in A} -\beta_1' e^{-d_1 f_1^j} - \beta_2' e^{-d_2 f_2^j} + 1. \tag{4.19}$$

Note that $-\beta_1' e^{-d_1 f_1^j} - \beta_2' e^{-d_2 f_2^j} + 1 \geq 0$ for all $\boldsymbol{f}^j \in A$ (coherence constraint). Therefore, by minimizing the objective function, we are minimizing the sum of the slack variables $s^j := -\beta_1' e^{-d_1 f_1^j} - \beta_2' e^{-d_2 f_2^j} + 1 \geq 0$, forcing the classifier to interpolate *as many points in A as possible (while satisfying $LC_{\hat{\phi}}(A) = 1$). This means we are looking for the smallest extension in the feature space: the smallest set compatible with the assessments A. The above nonlinear nonconvex optimisation problem can be solved numerically, by minimising (4.19) subject to*

$$
\begin{aligned}
-\beta_1' e^{-d_1 f_1^j} - \beta_2' e^{-d_2 f_2^j} + 1 \geq 0, \quad \forall \boldsymbol{f}^j \in A,\\
-\beta_1' e^{-d_1 f_1^j} - \beta_2' e^{-d_2 f_2^j} + 1 < 0, \quad \forall \boldsymbol{f}^j \in R,\\
d_1, d_2, \beta_1' \geq 0, \quad \beta_2' = 1 - \beta_1'.
\end{aligned} \tag{4.20}
$$

Let us show a numerical example. Suppose Diana considers to be acceptable and rejectable the finite sets of gambles already considered in Example 15: $A_{Diana} = \{[-1,2]^\top, [2,-1]^\top, [1,-0.5]^\top\}$ and $R_{Diana} = \{[-3,2]^\top, [1,-1]^\top, [4,-2]^\top\}$. The optimal nonlinear classifier is obtained for $\beta_1' = 0.082$, $d_1 = 0.90, d_2 = 0.07$.

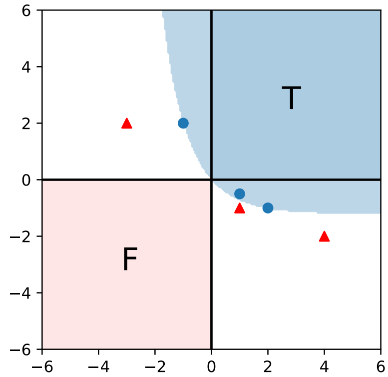

Figure 4.14. Nonlinear smooth classifier.

This classifier can be interpreted as the minimal set of assumptions the modeller Bruce can make on Diana assuming she only accepts gambles $\boldsymbol{f}$ such that $-\beta_1' e^{-d_1 f_1} - \beta_2' e^{-d_2 f_2} + 1 \geq 0$ for some probability distribution $\{\beta_1', \beta_2'\}$. The basis functions in (4.17) are commonly used to construct risk aversion models, see for instance Nau [2003]. Therefore, this example shows how our general framework allows us to express risk aversion models as binary classifiers. The set of gambles identified by this classifier satisfies also (CNV). However, it is not the smallest convex coherent set which is compatible with A_{Diana}, R_{Diana} because, as discussed in Section 4.3, the smallest convex coherent set is identified through a binary piecewise affine classifer.

4.7 Summary

Starting from the realization that the fundamental rationality axioms of *Additivity* and *Positive Homogeneity*, which underlie desirability, may impose limiting constraints on the interpretation and application of gambles, we proceed to examine alternative sets of axioms that relax these assumptions. Specifically, we demonstrate how each set of axioms considered can be associated with a distinct binary classification problem, thereby aligning more closely with practical applications.

We restrict our attention to two example of different axiomatizations of coherence that have been already proposed in the literature: namely, *convex coherence* and *positive additive coherence*. These alternatives definitions relax the standard linearity axioms of desirability, replacing them with either convexity or the minimal requirement of accepting gambles that are better than those already considered acceptable. When dealing with finite sets of both *acceptable* and *rejectable* gambles for an agent, we demonstrate that these axiomatizations, along with the standard coherence framework, can be reinterpreted as classification problems. Notably, the family of classifiers varies according to the particular axiomatization considered: *linear* for maximal coherence, *piecewise linear* for standard coherence, *piecewise affine* for convex coherence, and *piecewise positive affine* for positive additive coherence. For each classifier considered then, we show the possibility of constructing a feature mapping enabling it to be represented as a linear classifier in a higher-dimensional space. To conclude the chapter, we also establish connections with other imprecise-probabilistic models and lay the foundation for a comprehensive theory of coherence rooted in classification, thereby unifying the concepts presented earlier.

Chapter 5

Conclusions

In this thesis, we aim at enforcing even more the idea, born with de Finetti and then generalised with Walley [1991], Williams [1975] and others, that different research fields related to beliefs and behaviour of rational agents are founded on the same basic idea of rationality as consistency. This idea is made explicit by means of desirability, which has already been proven to be both a powerful theory of uncertainty and decision making, and whose main tools, used to model agents dispositions, encapsulate an idea of rationality that can be captured through the concept of logical consistency.

In the first line of research developed, we show that desirability can also be used to set a general problem of opinions aggregation. Its tools indeed allow to aggregate opinions expressed in very different forms, such as (sets of) probability distributions, (incomplete) preference relations, and so on. Moreover, their capability to model incomplete information provides a natural way to escape from traditional impossibility results and preserve desirable properties under pooling.

This approach offers many interesting research lines for the future. First, in our work we concentrate on a comparison between our formulation of the aggregation problem and traditional results of social choice and opinion pooling. In the future, it could be useful to enlarge this anaysis and investigate possible links with other formulations of the problem. In this sense, given the existing connection between desirability and classical propositional logic (see Section 1.1.3), it could be worth to investigate other connections with *belief merging* [Konieczny and Pérez, 1999, 2002] and *judgement aggregation* [Kornhauser and Sager, 1986; List and Pettit, 2002; Grossi and Pigozzi, 2014], two disciplines that study the problem of aggregating individual inputs into a global consistent outcome by using propositional logic.

The general framework we have proposed allows us also to investigate whether some axioms and results from a field are sensible in the other. In this respect, in Section 2.2, we have already discussed the axiom of *independence of irrelevant alternatives* that seems to be particularly suited to model voting problems but not so much to model problems involving aggregating (imprecise) beliefs. In a logical context, an analogue of this axiom receives some critics [Chapman, 2002], but it has also been recognised as a key factor to ensure consistent global outcomes and protect from strategic manipulation [Grossi and Pigozzi, 2014; Dietrich, 2006]. In the future, it would be of further interest to deepen this analysis including different axioms and frameworks.

In Section 2.2, moreover, we have seen that asking a social rule to generate group opinions consistent with the agents' ones leads to a linear approach to pooling. Hereafter, a similar analysis can be performed considering more general definitions of desirability and coherence, such as the ones suggested in Casanova, Benavoli and Zaffalon [2021]; Casanova et al. [2023]; Miranda and Zaffalon [2023], which can result in more general approaches to pooling as well.

Finally, it could also be useful for practical applications to move our results closer to computational social choice, an active area of research that focuses on the algorithmic tasks in social choice and their complexity [Chevaleyre et al., 2007].

As a second line of research, we analysed the connections between desirability and information algebras, generic structures to manage information providing operations and architectures for inference. We showed, in particular, that it is possible to formally induce information algebras from elements of desirability. This permits to abstract away properties of the latter that can be regarded as properties of more general algebraic structures and to further generalise results obtained with desirability, such as the ones related to the treatment of the marginal problem.

In the future, we would like to prove a sort of converse statement, i.e., that information algebras can be induced by a general non-linear reformulation of desirability (Casanova, Benavoli and Zaffalon [2021]; Miranda and Zaffalon [2023]). If this was possible, desirability would become an even more powerful formalism capable to deal with issues of uncertainty and decision making, but also subsuming all the variety of formalisms falling under the umbrella of information algebras, such as relational databases and various kind of logical formalisms. A deep comparison with the belief structures introduced by de Cooman [2005] and the models they subsume could be of further importance to reach this aim.

As a final line of research, we have analysed some instances of nonlinear desirability. For each of the different axiomatizations proposed, we provide an operational tool to check their compatibility with finite sets of acceptance/rejection statements, possibly provided by an agent. The latter is based on a different interpretation of the problem as a binary (usually nonlinear) classification task. Finally, we show the possibility to define a feature mapping that allows us to reformulate the above nonlinear classification problems as linear ones in higher-dimensional spaces. Extended notions related to the probabilistic interpretation of sets of gambles are also provided as well as a more general definition of coherence based on the algorithmic framework proposed.

In our work, we suggest examples of nonlinear desirability by proceeding in an axiomatic way. In Miranda and Zaffalon [2023], an even more general notion of nonlinear desirability is given by replacing the standard posi-operator with a more general one. This permits to not to be tight to single axiomatizations and deal with more general instances of nonlinearity.[1]

It could be interesting in the future to try to combine the two approaches. In particular, a focused examination of how integrating our feature mappings within the framework of Miranda and Zaffalon [2023] could prove beneficial in bringing it closer to the applications.

Another research line to pursue is exploring possible connections of Nau, 2003 with our work. Nau indeed characterizes sets of desirable gambles for agents with inseparable subjective probabilities and utilities, through a twice-differentiable utility function U. Specifically, he defines sets of desirable gambles as $\mathscr{D} := \{f \in \mathscr{L} : U(f) \geq 0\}$. In this framework, he also establishes *risk-neutral probabilities* for the agent, i.e., agent's subjective probabilities adjusted for risk, and provides a measure of their *local risk aversion*. It would be intriguing to investigate possible connections of Nau's findings with out work. A potential starting point for this exploration is to seek an equivalent representation of convex coherent sets and positive additive coherent sets in terms of lower previsions and credal sets, similar to the approach employed for coherent sets.

It could also be interesting to analyze the characterizations of the analogous, in desirability terms, of the most-famous examples of non-expected utility theories known in the literature [Quiggin, 1982; Chew and MacCrimmon, 1979a,b; Kahneman and Tversky, 2013]. We highlight however that this task is not straightforward since they usually work with horse-lotteries instead of gambles.

Another direction to pursue in the future to obtain a fully-operational theory of nonlinear desirability is the study of the dynamic case. At the moment indeed,

[1] It is possible to notice, moreover, that it is also the crucial point we need to invert the relationship between desirability and information algebras, as well as between desirability and logic: a broad range of logical systems, defined in quite general terms, can be considered as part of a generalised theory of desirability.

neither us or Miranda and Zaffalon [2023] account for the evolution of one's wealth in time due to the actual buying and selling of gambles. This is however fundamental for a real-world theory of decision making. Careful analysis is necessary since standard formulation of desirability operates under the assumption of "act-state independence", which implies that rewards of gambles are not influenced by the actions a subject takes.

Appendix A

Proofs of Chapter 1

Proof of Lemma 1. Consider a maximal set of strictly desirable gambles M^+. It is not strictly included in any other coherent set of strictly desirable gambles if and only if adding any gamble f not contained in M^+ to M^+ makes sure we can no longer extend the result $M^+ \cup \{f\}$ to a coherent set of strictly desirable gambles.

To prove this fact, let us also consider $P := \sigma(M^+)$. If $f \notin M^+$, then $P(f) \leq 0$. If $P(f) < 0$, then $P(-f) = -P(f) > 0$ and $-f \in M^+$ by definition. So there can not be a coherent set containing $M^+ \cup \{f\}$. If $P(f) = 0$, then $f \notin \mathscr{L}^+$ (because $f \notin M^+$ by hypothesis) and $P(f - \delta) = P(f) - \delta < 0$ for every $\delta > 0$. Thus, following the previous reasoning, $-f + \delta \in M^+$ for every $\delta > 0$. So, there cannot be a coherent set containing $M^+ \cup \{f\}$ also satisfying Eq. 1.5. $\square$

Proof of Lemma 2. Consider a maximal set of almost desirable gambles $\overline{M}$. It is not strictly included in any other coherent set of almost desirable gambles if and only if adding any gamble f not contained in $\overline{M}$ to $\overline{M}$ makes sure we can no longer extend the result $\overline{M} \cup \{f\}$ to a coherent set of almost desirable gambles.

To prove this fact, let us consider also $P := \sigma(\overline{M})$. If $f \notin \overline{M}$, then $P(f) < 0$. Hence, there exists $\epsilon > 0$ such that $P(f + \epsilon) = P(f) + \epsilon \leq 0$. Hence, $P(-f - \epsilon) = -P(f + \epsilon) \geq 0$ implying that $-f - \epsilon \in \overline{M}$ by definition. So, there can not be a coherent set containing $\overline{M} \cup \{f\}$. $\square$

Appendix B

Proofs of Chapter 2

Proof of Lemma 5. Suppose $\mathscr{G}$ and $\mathscr{G}'$ are both oligarchies, and assume that $i^* \in \mathscr{G}' \setminus \mathscr{G}$. Let $f \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$, and let us consider a profile (if needed made up only by maximal coherent sets of gambles) $[\mathscr{D}_i]$ such that $f \in \mathscr{D}_i$ for all $i \in \mathscr{G}$ and $-f \in \mathscr{D}_i$ for all $i \in \mathscr{G}' \setminus \mathscr{G}$: it suffices to make

$$(\forall i \in \mathscr{V})\ \mathscr{D}_i := \begin{cases} \mathscr{E}(\{f\}) := \operatorname{posi}(\{f\} \cup \mathscr{L}^+) & \text{or a maximal superset, if } i \in \mathscr{G} \\ \mathscr{E}(\{-f\}) := \operatorname{posi}(\{-f\} \cup \mathscr{L}^+) & \text{or a maximal superset, if } i \notin \mathscr{G}. \end{cases}$$

Then since $\mathscr{G}'$ is an oligarchy we should have $f \notin \Gamma([\mathscr{D}_i])$, but $\mathscr{G}$ decisive implies $f \in \Gamma([\mathscr{D}_i])$, a contradiction. □

We establish now some preliminary results needed to prove the subsequent ones.

Lemma 13. *Given a set* $\mathscr{K} \subseteq \mathscr{L}$, $\operatorname{ch}(\mathscr{K}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset \Leftrightarrow 0 \notin \operatorname{posi}(\mathscr{K} \cup \mathscr{L}^+) =: \mathscr{E}(\mathscr{K})$.[1] *As a consequence, if any of the two conditions above hold, it follows that* $\mathscr{K}$ *has a coherent (maximal) superset.*

Proof. It is equivalent to prove that

$$\operatorname{ch}(\mathscr{K}) \cap (\mathscr{L}^- \cup \{0\}) \neq \emptyset \Leftrightarrow 0 \in \operatorname{posi}(\mathscr{K} \cup \mathscr{L}^+).$$

[1] We indicate with $\operatorname{ch}(\mathscr{K})$ the convex hull of a set of gambles $\mathscr{K}$, as we will see in Section 1.

To prove that this is the case, note that

$$
\begin{aligned}
&0 \in \mathrm{posi}(\mathscr{K} \cup \mathscr{L}^+)\\
&\quad\Leftrightarrow (\exists f \in \mathscr{L}^+ \cup \{0\}, f_j \in \mathscr{K}, \lambda_j > 0, j = 1,\dots,r, r \geq 1)\ 0 = f + \sum_{j=1}^{r} \lambda_j f_j\\
&\quad\Leftrightarrow (\exists f \in \mathscr{L}^+ \cup \{0\}, f_j \in \mathscr{K}, \lambda_j > 0, j = 1,\dots,r, r \geq 1)\ -f = \sum_{j=1}^{r} \lambda_j f_j\\
&\quad\Leftrightarrow (\exists f \in \mathscr{L}^+ \cup \{0\}, f_j \in \mathscr{K}, \lambda_j > 0, j = 1,\dots,r, r \geq 1)\ -\frac{f}{\sum_{j=1}^{r}\lambda_j} = \sum_{j=1}^{r} \frac{\lambda_j}{\sum_{j=1}^{r}\lambda_j} f_j,
\end{aligned}
$$

and since the left-hand side belongs to $\mathscr{L}^- \cup \{0\}$ and the right-hand side belongs to $\mathrm{ch}(\mathscr{K})$, we deduce that this is equivalent to $\mathrm{ch}(\mathscr{K}) \cap (\mathscr{L}^- \cup \{0\}) \neq \emptyset$.

For the second statement, if any of these conditions hold we deduce that the set $\mathrm{posi}(\mathscr{K} \cup \mathscr{L}^+)$ is a coherent set of gambles that trivially includes $\mathscr{K}$. Since any coherent set of gambles has a coherent maximal superset, see Section 1, it follows in particular that $\mathscr{K}$ has a coherent maximal superset. □

Lemma 14. *Assume that $|\mathscr{X}| \geq 3$. Then for every gamble f such that $f \notin (\mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\})$ there always exist two gambles h, t such that $-f = f + h + t$ and*

- $\mathrm{ch}(\{f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, $\mathrm{ch}(\{-f, h, f + h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ *and* $\mathrm{ch}(\{-f, h, -(f + h)\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$*;*
- $\mathrm{ch}(\{f + h, t\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, $\mathrm{ch}(\{-(f + h), t, -f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ *and* $\mathrm{ch}(\{-(f + h), t, f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$*.*

Proof. Since f is such that $f \notin (\mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\})$ then there exist at least two points $z_1, z_2 \in \mathscr{X}$ with $z_1 \neq z_2$ such that $f(z_1) = f_1 > 0$ and $f(z_2) = f_2 < 0$. We consider a number of possibilities:

1. If there is some $z_3 \neq z_1, z_2$ such that $f(z_3) \geq 0$, then we define, for a fixed $\epsilon > 0$

$$
(\forall z \in \mathscr{X})\ h(z) := \begin{cases} f_1 > 0 & \text{if } z = z_1,\\ -f_2/2 > 0 & \text{if } z = z_2,\\ -2f(z) - \epsilon & \text{otherwise,} \end{cases}
$$

and

$$
(\forall z \in \mathscr{X})\ t(z) := \begin{cases} -3f_1 < 0 & \text{if } z = z_1,\\ -3f_2/2 > 0 & \text{if } z = z_2,\\ \epsilon & \text{otherwise.} \end{cases}
$$

Then by construction $-f = f + h + t$. To see that they fulfil the conditions in the lemma, note that:

- $\mathrm{ch}(\{f,h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_1);
- $\mathrm{ch}(\{-f,h,f+h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$: if $0 \geq \lambda_{-f}(-f)+\lambda_h h+\lambda_{f+h}(f+h)$, then it should be on the one hand $\lambda_{-f} \geq \lambda_h + 2\lambda_{f+h}$ because of z_1, and on the other $2\lambda_{-f} + \lambda_h \leq \lambda_{f+h}$ because of z_2; these two conditions are incompatible.
- $\mathrm{ch}(\{-f,h,-(f+h)\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_2);

and also

- $\mathrm{ch}(\{f+h,t\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$: if $\lambda_{f+h}(f+h)+\lambda_t t \leq 0$, it should be on the one hand $\lambda_{f+h} \leq 0.6$ because of z_1 and on the other $\lambda_{f+h} \geq 0.75$ because of z_2;
- $\mathrm{ch}(\{-(f+h),t,-f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_2);
- $\mathrm{ch}(\{-(f+h),t,f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_3 and z_1).

2. If $f(z_3) < 0$ for every $z_3 \neq z_1, z_2$, then we define:

$$(\forall z \in \mathscr{Z})\ h(z) := \begin{cases} -f_1 < 0 & \text{if } z = z_1, \\ -f_2/2 > 0 & \text{if } z = z_2, \\ -2f(z) & \text{otherwise,} \end{cases}$$

and

$$(\forall z \in \mathscr{Z})\ t(z) := \begin{cases} -f_1 < 0 & \text{if } z = z_1, \\ -3f_2/2 > 0 & \text{if } z = z_2, \\ 0 & \text{otherwise.} \end{cases}$$

Then by construction $-f = f + h + t$. To see that they fulfil the conditions in the lemma, note that:

- $\mathrm{ch}(\{f,h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$: if $\lambda_f(f) + \lambda_h h \leq 0$, it should be on the one hand $\lambda_f \leq 0.5$ because of z_1 and on the other $\lambda_f \geq \frac{2}{3}$ because of z_3;
- $\mathrm{ch}(\{-f,h,f+h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_3);
- $\mathrm{ch}(\{-f,h,-(f+h)\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_2);

and also

- $\mathrm{ch}(\{f+h,t\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$ (use z_2 and z_3);
- $\mathrm{ch}(\{-(f+h),t,-f\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$ (use z_2);
- $\mathrm{ch}(\{-(f+h),t,f\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$: if $0\geq-\lambda_{-(f+h)}(f+h)+\lambda_t t+\lambda_f f$, then it should be $\lambda_t\geq\lambda_f$ so that this combination is non-positive on z_1, but then this implies that it is positive on z_2. □

Lemma 15 (Field expansion lemma). *Assume that $|\mathscr{X}|\geq 3$. If a social rule satisfies the following properties:*

- *unlimited domain or unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*

and it admits a group $\mathscr{G}$ that is almost decisive for a gamble f, then it is decisive.

Proof. Note that we can assume without loss of generality that $\mathscr{G}$ is a proper subset of $\mathscr{V}$, since otherwise the thesis follows immediately from the property of weak Pareto. In addition, we can then assume that the gamble f does not belong to $\mathscr{L}^+\cup\mathscr{L}^-\cup\{0\}$.

We must prove that $(\forall[\mathscr{D}_i])\ \cap_{i\in\mathscr{G}}\mathscr{D}_i\subseteq\Gamma([\mathscr{D}_i])$. Consider thus a profile $[\mathscr{D}_i]$ (eventually in $\mathbb{M}^n$), and let $g\in\cap_{i\in\mathscr{G}}\mathscr{D}_i$. We can assume without loss of generality that $g\notin\mathscr{L}^+$, since in that case it trivially belongs to $\Gamma([\mathscr{D}_i])$. There are a number of possibilities:

1. Assume first of all that $g=f+h$, where:

$$\mathrm{ch}(\{f,h\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset, \tag{B.1}$$

$$\mathrm{ch}(\{-f,h,g\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset, \tag{B.2}$$

$$\mathrm{ch}(\{-f,h,-g\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset. \tag{B.3}$$

   Hence, it holds that:

   - By Lemma 13 in Appendix B, if $\mathrm{ch}(\{f,h\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$, there exists a coherent set of gambles $\mathscr{D}^1$ (possibly maximal) that includes both f,h, and as a consequence also g by additivity.
   - Analogously, we can construct the coherent set $\mathscr{D}^2:=\mathrm{posi}(\{-f,h,g\}\cup\mathscr{L}^+)$ (or a maximal set that contains it). It includes h,g but not f.

- Finally, the coherent set $\mathscr{D}^3 := \operatorname{posi}(\{-f, h, -g\} \cup \mathscr{L}^+)$ (or a maximal set that contains it) includes h and not f and g.

Let us now define the following profile $[\mathscr{D}'_i]$:

- For every $i \in \mathscr{G}$, let $\mathscr{D}'_i = \mathscr{D}^1$ (possibly the maximal one). Then $f, h \in \mathscr{D}'_i$, whence also $g \in \mathscr{D}'_i$ by additivity.
- Given $i \in \mathscr{V} \setminus \mathscr{G}$, if $g \in \mathscr{D}_i$ then we let $\mathscr{D}'_i = \mathscr{D}^2$ (possibly the maximal one). Then $h, g \in \mathscr{D}'_i$ and $f \notin \mathscr{D}'_i$.
- Given $i \in \mathscr{V} \setminus \mathscr{G}$, if $g \notin \mathscr{D}_i$ then we let $\mathscr{D}'_i = \mathscr{D}^3$ (possibly the maximal one). Then $h \in \mathscr{D}'_i$ and $f, g \notin \mathscr{D}'_i$.

It follows then that $h \in \cap_i \mathscr{D}'_i$, whence $h \in \Gamma([\mathscr{D}'_i])$, using the weak Pareto property; moreover, $f \in \cap_{i \in \mathscr{G}} \mathscr{D}'_i$ and $f \notin \cup_{i \in \mathscr{V} \setminus \mathscr{G}} \mathscr{D}'_i$. Since $\mathscr{G}$ is almost decisive for f, we deduce that $f \in \Gamma([\mathscr{D}'_i])$. Since the latter is a coherent set of gambles, we deduce by additivity that $g = f + h \in \Gamma([\mathscr{D}'_i])$. And since $(\forall i \in \mathscr{V})\ g \in \mathscr{D}_i \iff g \in \mathscr{D}'_i$ by construction, we deduce from the independence of irrelevant alternatives that also $g \in \Gamma([\mathscr{D}_i])$.

Now note that the reasoning above can be applied in particular for those $g \gneq f$ (i.e., when $h \in \mathscr{L}^+$). To see that this is the case, note that when $h \in \mathscr{L}^+$:

(C.1) $\operatorname{ch}(\{f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ is equivalent to $0 \notin \operatorname{posi}(\{f\} \cup \mathscr{L}^+)$, which holds because $f \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$ by assumption.

(C.2) $\operatorname{ch}(\{-f, h, g\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ becomes equivalent to $\operatorname{ch}(\{-f, g\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$. But if this intersection was non-empty we would deduce that for some $\lambda_{-f} + \lambda_g = 1$ it holds that $0 \geq \lambda_{-f}(-f) + \lambda_g g \geq (\lambda_g - \lambda_{-f}) f$, which contradicts the assumption that $f \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$.

(C.3) $\operatorname{ch}(\{-f, h, -g\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ becomes equivalent to $0 \notin \operatorname{posi}(\{-g\} \cup \mathscr{L}^+)$, which holds because $g \notin \mathscr{L}^+ \cup \{0\}$.

This allows us to deduce that

$$\mathscr{G} \text{ is decisive for any } g \gneq f. \tag{B.4}$$

Next we establish that:

$$\mathscr{G} \text{ almost decisive for } g \Rightarrow \mathscr{G} \text{ decisive for } -g. \tag{B.5}$$

To prove this, note that we can assume without loss of generality that $g \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$. Applying Lemma 14 in Appendix B, we can find two gambles h, t such that $-g = g + h + t$. Moreover, by construction in the proof of that lemma we have that $\mathrm{ch}(\{g,h\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$, $\mathrm{ch}(\{-g,h,(g+h)\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$ and $\mathrm{ch}(\{-g,h,-(g+h)\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$. Applying point 1, we deduce that $\mathscr{G}$ is decisive for $g+h$, and therefore also almost decisive. Moreover, we also have from the proof of the lemma that $\mathrm{ch}(\{g+h,t\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$, $\mathrm{ch}(\{-(g+h),t,-g\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$ and $\mathrm{ch}(\{-(g+h),t,g\})\cap(\mathscr{L}^-\cup\{0\})=\emptyset$. Applying again point 1, we deduce that $\mathscr{G}$ is decisive for $g+h+t=-g$.

Thus, Eq. (B.5) holds. Applying twice this condition we deduce in particular that $\mathscr{G}$ is decisive for $-f$, and also for f.

2. Assume now that $g=f+h$, with $h\in\mathscr{L}^-$. Then $g\lnsim f$, or, equivalently $-f\lnsim -g$. Since $\mathscr{G}$ is decisive for $-f$, it is in particular almost decisive for this gamble, whence by Eq.(B.4) $\mathscr{G}$ is also almost decisive for $-g\gnsim -f$. Applying now Eq. (B.5), we deduce that $\mathscr{G}$ is also decisive for g.

3. Finally, consider a gamble $h\notin\mathscr{L}^+\cup\mathscr{L}^-\cup\{0\}$, and let $g:=f+h$. Then we can rewrite g as $g=f-h^-+h^+$, where h^+ and h^- are respectively the positive and the negative part function of h, i.e. $h^+:=\max\{0,h\}$ and $h^-:=-\min\{0,h\}$. It then follows from point 2 that $\mathscr{G}$ is decisive for $g=f-h^-$ because $-h^-\in\mathscr{L}^-$, and as a consequence it is also almost decisive for this gamble. If we now apply Eq. (B.4) we deduce that $\mathscr{G}$ is decisive for $g=(f-h^-)+h^+\gnsim f-h^-$. □

Lemma 16 (Group contraction lemma). *Assume that $|\mathscr{X}|\geq 3$. If a social rule satisfies the following properties:*

- *completeness,*
- *unlimited maximal domain,*
- *independence of irrelevant alternatives,*
- *weak Pareto,*

then if a group $\mathscr{G}$ containing at least two individuals is decisive, then it contains a proper subset of individuals that is also decisive.

Proof. Let us prove that there is a proper subset $\mathscr{G}'$ of $\mathscr{G}$ that is also decisive. For this reason, we consider a partition of $\mathscr{G}$ into non-empty and disjoint subsets $\mathscr{G}_1$ and $\mathscr{G}_2$, and we proceed to establish that one of these two sets is also decisive.

Consider $z_1, z_2 \in \mathscr{Z}$ with $z_1 \neq z_2$, and let us define the gambles f_1, f_2, f_3 as:

$$(\forall z \in \mathscr{Z})\ f_1(z) := \begin{cases} -1 & \text{if } z = z_1, \\ 1 & \text{if } z = z_2, \\ 0 & \text{otherwise,} \end{cases} \qquad (\forall z \in \mathscr{Z})\ f_2(z) := \begin{cases} 0 & \text{if } z = z_1, \\ -2 & \text{if } z = z_2, \\ 1 & \text{otherwise,} \end{cases}$$

$$f_3 := f_1 + f_2.$$

Let z_3 denote an element different from z_1, z_2, existing because $|\mathscr{Z}| \geq 3$. These gambles satisfy the following conditions:

- $\mathrm{ch}(\{f_1, f_2, f_3\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_3 and z_2);
- $\mathrm{ch}(\{-f_1, f_2, -f_3\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_1 and z_3);
- $\mathrm{ch}(\{f_1, -f_3\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_2).

Then, thanks to Lemma 13 in Appendix B we can consider the following profile $[M_i] \in \mathbb{M}^n$:

- for all $i \in \mathscr{G}_1, M_i := M^1 \supseteq \mathrm{posi}(\{f_1, f_2, f_3\} \cup \mathscr{L}^+)$, with $M^1 \in \mathbb{M}$;
- for all $i \in \mathscr{G}_2, M_i := M^2 \supseteq \mathrm{posi}(\{-f_1, f_2, -f_3\} \cup \mathscr{L}^+)$, with $M^2 \in \mathbb{M}$;
- for all $i \notin \mathscr{G}, M_i := M^3 \supseteq \mathrm{posi}(\{f_1, -f_3\} \cup \mathscr{L}^+)$, with $M^3 \in \mathbb{M}$.

Note then that $f_3 \notin M_i$ for any $i \notin \mathscr{G}_1$, and that $-f_1 \notin M_i$ for any $i \notin \mathscr{G}_2$. Since $f_2 \in \cap_{i \in \mathscr{G}} M_i$, it follows from decisiveness of $\mathscr{G}$ that $f_2 \in \Gamma([M_i])$. Since this set is maximal, then it either includes f_1 or $-f_1$.

1. If $f_1 \in \Gamma([M_i])$, then $f_3 \in \Gamma([M_i])$ by additivity. Let us prove that in that case $\mathscr{G}_1$ is almost decisive for f_3.

   Consider any profile $[M_i'] \in \mathbb{M}^n$, such that $f_3 \in \cap_{i \in \mathscr{G}_1} M_i'$ and $f_3 \notin \cup_{i \notin \mathscr{G}_1} M_i'$. Then, $(\forall i \in \mathscr{V})\ f_3 \in M_i' \iff f_3 \in M_i$ and applying independence of irrelevant alternatives we deduce that $f_3 \in \Gamma([M_i'])$, thus $\mathscr{G}_1$ is almost decisive for f_3. Applying Lemma 15 in Appendix B, we deduce that it is also decisive.

2. If $-f_1 \in \Gamma([M_i])$, we can analogously show that $\mathscr{G}_2$ is almost decisive for $-f_1$ and thus, applying again Lemma 15 in Appendix B, we deduce that it is also decisive.

$\square$

Proof of Theorem 3. The case of $|\mathscr{V}| = 1$ is trivial, so let us assume $|\mathscr{V}| \geq 2$. By weak Pareto condition we know that $\mathscr{V}$ is decisive. By repeatedly applying Lemma 16 in Appendix B, we can eventually arrive to a decisive individual, who must, thus, be a dictator. □

Proof of Proposition 2. Let j be the dictator. Since we have unlimited maximal domain, the social rule must be given by $\Gamma([M_i]) = M_j$ for any profile $[M_i]$, and as a consequence it satisfies independence of irrelevant alternatives. □

Proof of Theorem 4. Since by Lemma 5 there can be at most one oligarchy, we only need to establish its existence.

By the weak Pareto condition, the set $\mathscr{V}$ of all individuals is decisive. Let $\mathscr{G}$ be a decisive set of minimal size, meaning that there does not exist $\mathscr{G}' \subset \mathscr{G}$ that is also decisive. We shall demonstrate that $\mathscr{G}$ is an oligarchy. If $\mathscr{G}$ contains a single member, then $\mathscr{G}$ is trivially an oligarchy because any dictatorship is. Let us then consider the case in which $|\mathscr{G}| \geq 2$. Note that we only need to prove that O2 holds, because O1 follows immediately because $\mathscr{G}$ is decisive.

We consider first of all the case where Γ has unlimited domain.

Consider a profile $[\mathscr{D}_i]$, and let $f \in \cup_{i \in \mathscr{G}} \mathscr{D}_i$. We must prove that $-f \notin \Gamma([\mathscr{D}_i])$. We may assume without loss of generality that $f \notin \mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\}$; the result otherwise is trivial or impossible. We can partition the group $\mathscr{G}$ into the following sets:

- $A := \{i \in \mathscr{G} | f \in \mathscr{D}_i\}$,
- $B := \{i \in \mathscr{G} | -f \in \mathscr{D}_i\}$,
- $C := \{i \in \mathscr{G} | f \notin \mathscr{D}_i, -f \notin \mathscr{D}_i\}$,

where by assumption A is non-empty, but B, C may be.

Since $f \notin (\mathscr{L}^+ \cup \mathscr{L}^- \cup \{0\})$, there exist two points $z_1, z_2 \in \mathscr{Z}$ with $z_1 \neq z_2$ such that $f(z_1) = f_1 > 0$ and $f(z_2) = f_2 < 0$. Let us prove that there always exist two gambles $\{h, g\}$ such that $f = h + g$ and

- $\mathrm{ch}(\{f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, and $g, -g \notin \mathscr{D}^1 := \mathrm{posi}(\{f, h\} \cup \mathscr{L}^+)$;
- $\mathrm{ch}(\{-f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, and $-g \in \mathscr{D}^2 := \mathrm{posi}(\{-f, h\} \cup \mathscr{L}^+)$;
- $\mathrm{ch}(\{-g, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, and $f, -f \notin \mathscr{D}^3 := \mathrm{posi}(\{-g, h\} \cup \mathscr{L}^+)$;
- $\mathrm{ch}(\{-f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$, whence $-g$ and as a consequence $h \notin \mathscr{D}^4 := \mathrm{posi}(\{-f\} \cup \mathscr{L}^+)$.

We consider a number of possibilities:

1. If there is some $z_3 \neq z_1, z_2$ such that $f(z_3) > 0$, then we define for every $z \in \mathscr{Z}$:

$$h(z) := \begin{cases} 3f_1/2 > 0 & \text{if } z = z_1, \\ f_2/2 < 0 & \text{if } z = z_2, \\ f(z) - \epsilon & \text{otherwise,} \end{cases} \quad \text{and} \quad g(z) := \begin{cases} -f_1/2 < 0 & \text{if } z = z_1, \\ f_2/2 < 0 & \text{if } z = z_2, \\ \epsilon & \text{otherwise,} \end{cases}$$

where $\epsilon > 0$ is small enough for the conditions to be satisfied.

Then by construction $f = h + g$. To see that they fulfil the conditions, note that:

- $\mathrm{ch}(\{f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_1);
  - $g \notin \mathscr{D}^1$ (use z_1);
  - $-g \notin \mathscr{D}^1$ (use z_3 with $0 < \epsilon < f(z_3)$);
- $\mathrm{ch}(\{-f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$: If $\lambda_{-f}(-f) + \lambda_h h \leq 0$ it should be on the one hand $\lambda_{-f} \geq 0.6$ because of z_1 and on the other $\lambda_{-f} \leq \frac{1}{3}$ because of z_2. Then $-g \in \mathscr{D}^2$ by additivity;
- $\mathrm{ch}(\{-g, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_1);
  - $f \notin \mathscr{D}^3$: If $f \geq \lambda_{-g}(-g) + \lambda_h h$ for some non-negative λ_{-g}, λ_h (with at least one positive), we have that
  $$f_2 \geq -\lambda_{-g} f_2/2 + \lambda_h f_2/2,$$
  whence
  $$-1 \geq \lambda_{-g}/2 - \lambda_h/2$$
  dividing by the positive number $-f_2$. This means that
  $$\lambda_h \geq 2 + \lambda_{-g} \geq 2.$$
  Now, this means that
  $$\lambda_h h(z_1) + \lambda_{-g}(-g(z_1)) = \lambda_h 3f_1/2 + \lambda_{-g} f_1/2 \geq 3f_1 > f_1,$$
  a contradiction;
  - $-f \notin \mathscr{D}^3$ (use z_1);
- $\mathrm{ch}(\{-f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ by definition of f;
  - $-g \notin \mathscr{D}^4$: If $-g \geq \lambda(-f)$ it should be on the one hand $\lambda \leq 0.5$ because of z_2 and on the other $\lambda \geq \epsilon/f(z_3) > 0.5$ if we choose $0 < f(z_3)/2 < \epsilon < f(z_3)$, because of z_3.

2. If there is some $z_3 \neq z_1, z_2$ such that $f(z_3) = 0$, then we define for every $z \in \mathscr{Z}$:
$$h(z) := \begin{cases} 3f_1/2 > 0 & \text{if } z = z_1, \\ f_2/2 < 0 & \text{if } z = z_2, \\ f(z) - \epsilon & \text{otherwise,} \end{cases} \quad \text{and} \quad g(z) := \begin{cases} -f_1/2 < 0 & \text{if } z = z_1, \\ f_2/2 < 0 & \text{if } z = z_2, \\ \epsilon & \text{otherwise,} \end{cases}$$
where $\epsilon > 0$ is small enough for the conditions to be satisfied.

   Notice that h and g are the same of the previous case, so to show that they fulfil the conditions we indicate only the cases that involve the value of $f(z_3)$.

   - It only changes $-g \notin \mathscr{D}^1$: If $-g \geq \lambda_f f + \lambda_h h$ for some non-negative λ_f, λ_h (with at least one positive), we have that
   $$-\epsilon \geq \lambda_f(f(z_3)) + \lambda_h(f(z_3) - \epsilon) = -\epsilon\lambda_h,$$
   from which follows $\lambda_h \geq 1$, which does not work for z_1 because $-g(z_1) < h(z_1)$ and $h(z_1), f_1$ are both positive;
   - same as before;
   - same as before;
   - it only changes $-g \notin \mathscr{D}^4$ (use z_3).

3. Finally, if $f(z_3) < 0$ for every $z_3 \neq z_1, z_2$, then we define for every $z \in \mathscr{Z}$:
$$h(z) := \begin{cases} f_1 > 0 & \text{if } z = z_1, \\ 3f_2/2 < 0 & \text{if } z = z_2, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad g(z) := \begin{cases} 0 & \text{if } z = z_1, \\ -f_2/2 > 0 & \text{if } z = z_2, \\ f(z) & \text{otherwise.} \end{cases}$$
Then by construction $f = h + g$. To see that they fulfil the conditions, note that:

   - $\mathrm{ch}(\{f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_1);
     - $g \notin \mathscr{D}^1$ (use z_1);
     - $-g \notin \mathscr{D}^1$ (use z_1);
   - $\mathrm{ch}(\{-f, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_3 and z_1) and $-g \in \mathscr{D}_2$ by additivity;
   - $\mathrm{ch}(\{-g, h\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ (use z_3 and z_1);
     - $f \notin \mathscr{D}^3$ (use z_3);
     - $-f \notin \mathscr{D}^3$ (use z_1);
   - $\mathrm{ch}(\{-f\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$ by definition of f;
     - $-g \notin \mathscr{D}^4$: (use z_2).

Let us consider now a profile $[\mathscr{D}'_i]$ where $\mathscr{D}'_i = \mathscr{D}^1$ if $i \in A$, $\mathscr{D}'_i = \mathscr{D}^2$ if $i \in B$, $\mathscr{D}'_i = \mathscr{D}^3$ if $i \in C$ and $\mathscr{D}'_i$ is either $\mathscr{L}^+$ or $\mathscr{D}^4$ if $i \notin \mathscr{G}$ (as needed below so as to include $-f$ if necessary). We shall use these sets to establish that $-f \notin \Gamma([\mathscr{D}_i])$. We have a number of possible scenarios:

- If A, B, C are all non-empty and $-f \in \Gamma([\mathscr{D}_i])$, then we can consider a profile $[\mathscr{D}'_i]$ as above so that $(\forall i \in \mathscr{V})\ -f \in \mathscr{D}_i \iff -f \in \mathscr{D}'_i$. Applying independence of irrelevant alternatives, it follows that $-f \in \Gamma([\mathscr{D}'_i])$. Since $\mathscr{G}$ is decisive and $h \in \cap_{i\in\mathscr{G}}\mathscr{D}'_i$, we also have that $h \in \Gamma([\mathscr{D}'_i])$, we deduce that $-g = -f + h$ belongs to $\Gamma([\mathscr{D}'_i])$. By construction, $-g \in \cap_{i\in B\cup C}\mathscr{D}'_i$. Applying independence of irrelevant alternatives, for any other profile $[\mathscr{D}''_i]$ such that $-g \in \cap_{i\in B\cup C}\mathscr{D}''_i$ and $-g \notin \cup_{i\notin B\cup C}\mathscr{D}''_i$, it holds that $-g \in \Gamma([\mathscr{D}''_i])$. But this means that $B \cup C$ is almost decisive for $-g$, and applying Lemma 15 in Appendix B, it is decisive. This contradicts that $\mathscr{G}$ is a decisive set of minimal size.

- If $A \neq \emptyset \neq B$ and $C = \emptyset$, we reason as in the previous case, and end up concluding that B is a decisive set.

- If $A \neq \emptyset \neq C$ and $B = \emptyset$, we reason as in the first case, and end up concluding that C is a decisive set.

- Finally, if $A \neq \emptyset$ and $B = C = \emptyset$, it holds that $f \in \cap_{i\in\mathscr{G}}\mathscr{D}_i$, whence, since $\mathscr{G}$ is decisive, $f \in \Gamma([\mathscr{D}_i])$. Since the latter is a coherent set, this means that $-f \notin \Gamma([\mathscr{D}_i])$.

We consider next the case of unlimited maximal domain. Recall that since $\mathscr{G}$ is decisive, condition O1 is satisfied. Let us consider next condition O2. Consider a profile $[M_i] \in \mathbb{M}^n$. Assume that $f \in \cup_{i\in\mathscr{G}}M_i$. Since all the sets in the profile are maximal, we can partition again the group $\mathscr{G}$ into the following sets:

- $A := \{i \in \mathscr{G} | f \in M_i\}$,

- $B := \{i \in \mathscr{G} | -f \in M_i\}$,

where by assumption A is non-empty, but B may be. We need to show that $-f \notin \Gamma([M_i])$.

Let us consider again h, g of the form considered before such that $f = h + g$. It is possible to notice that regardless of the sign of $f(z_3)$:

- $\operatorname{ch}(\{h, g\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$. It can be shown by using z_3 if $f(z_3) > 0$; z_1 and z_2 if $f(z_3) < 0$. If instead $f(z_3) = 0$, if $\lambda_h h + \lambda_g g \leq 0$ it should be on the one hand $\lambda_h \leq 0.25$ because of z_1 and on the other hand $\lambda_h \geq 0.5$ because of z_3. Hence, $\text{posi}(\{h, g, f\} \cup \mathscr{L}^+)$ is a coherent set of gambles for which there exists at least a maximal coherent superset M^1 such that $-g \notin M^1$ by coherence.

- $\text{ch}(\{-f, g\}) \cap (\mathscr{L}^- \cup \{0\}) = \emptyset$. It can be shown by using z_2 or z_3 if $f(z_3) = 0$; z_2 if $f(z_3) < 0$. If instead $f(z_3) > 0$, if $\lambda_{-f}(-f) + \lambda_g g \leq 0$ it should be on the one hand $\lambda_{-f} \leq 1/3$ because of z_2 and $\lambda_{-f} > 1/3$ given that $\epsilon > f(z_3)/2$, because of z_3. Hence, $\text{posi}(\{-f, g\} \cup \mathscr{L}^+)$ is a coherent set of gambles for which there exists at least a maximal coherent superset M^3 such that $-g \notin M^3$ by coherence.

Let us consider now a profile $[M_i']$ where $M_i' = M^1$ if $i \in A$, $M_i' = M^2$ if $i \in B$, where M^2 is a maximal superset of $\mathscr{D}^2$ (hence $-f, h, -g \in M^2$), and M_i' is either M^1 or M^3 if $i \notin \mathscr{G}$ (as needed below so as to include $-f$ if necessary). We shall use these sets to establish that $-f \notin \Gamma([M_i])$. There are two possible scenarios similar to the ones considered for a social rule satisfying unlimited domain.

- If both A, B are non-empty and $-f \in \Gamma([M_i])$, then we can consider a profile $[M_i']$ as above so that $(\forall i \in \mathscr{V})\ -f \in M_i \iff -f \in M_i'$. Applying independence of irrelevant alternatives, it follows that $-f \in \Gamma([M_i'])$. Since $\mathscr{G}$ is decisive and $h \in \cap_{i \in \mathscr{G}} M_i'$, we also have that $h \in \Gamma([M_i'])$, we deduce that $-g = -f + h$ belongs to $\Gamma([M_i'])$. By construction, $-g \in \cap_{i \in B} M_i'$. Applying independence of irrelevant alternatives, for any other profile $[M_i'']$ such that $-g \in \cap_{i \in B} M_i''$ and $-g \notin \cup_{i \notin B} M_i''$, it holds that $-g \in \Gamma([M_i''])$. But this means that B is almost decisive for $-g$, and applying Lemma 15 in Appendix B, it is decisive. This contradicts that $\mathscr{G}$ is a decisive set of minimal size.

- if instead $A \neq \emptyset$ and $B = \emptyset$, it holds that $f \in \cap_{i \in \mathscr{G}} M_i$, whence, since $\mathscr{G}$ is decisive, $f \in \Gamma([M_i])$. Since the latter is a coherent set, this means that $-f \notin \Gamma(M_i)$. □

Proof of Proposition 3. Assume ex-absurdo that there exists a profile $[\mathscr{D}_i]$ such that $\Gamma([\mathscr{D}_i]) \supsetneq \cap_{i \in \mathscr{G}} \mathscr{D}_i$, and let us consider $f \in \Gamma([\mathscr{D}_i]) \setminus \cap_{i \in \mathscr{G}} \mathscr{D}_i$, so that $A = \{i \in \mathscr{G} : f \in \mathscr{D}_i\} \neq \mathscr{G}$. For every $j \in \mathscr{G} \setminus A$, let $\mathscr{D}_j'$ be a maximal set of gambles M that includes $-f$, and let $\mathscr{D}_i' := \mathscr{D}_i$ for every $i \in A \cup \mathscr{G}^c$. Then since $-f \in \cup_{i \in \mathscr{G}} \mathscr{D}_i'$, it follows from O2 that $f \notin \Gamma([\mathscr{D}_i'])$. But on the other hand f belongs to the same sets in the profiles $[\mathscr{D}_i]$ and $[\mathscr{D}_i']$, so $f \in \Gamma([\mathscr{D}_i]), f \notin \Gamma([\mathscr{D}_i'])$ is a contradiction of independence of irrelevant alternatives. □

Proof of Theorem 5. From Theorem 4 there exists a unique oligarchy $\mathscr{G}$. Then we must show that if the social rule satisfies also anonymity, $\mathscr{G} = \mathscr{V}$. Assume ex-absurdo that $\mathscr{G} \neq \mathscr{V}$, and take $i^* \notin \mathscr{G}$.

By the hypothesis of unlimited (maximal) domain we can consider a profile (possibly composed only by maximal coherent sets of gambles) $[\mathscr{D}_i]$ such that there exists $f \in \mathscr{D}_{i^*}$ and $-f \in \mathscr{D}_j$ $\forall j \neq i^*$. Since $\mathscr{G}$ is an oligarchy, it follows that $-f \in \Gamma([\mathscr{D}_i])$.

Consider now a permutation σ of $\mathscr{V}$ such that $\sigma(i^*) \neq i^*$, $\sigma(i^*) \in \mathscr{G}$, and let $[\mathscr{D}'_i] := [\mathscr{D}_{\sigma(i)}]$ denote the associated profile. Then there exists some $j \in \mathscr{G}$ such that $f \in \mathscr{D}'_j$, whence, by definition of oligarchy, $-f \notin \Gamma([\mathscr{D}'_i])$. But this means that $\Gamma([\mathscr{D}_{\sigma(i)}]) \neq \Gamma([\mathscr{D}_i])$, meaning that anonymity is violated. □

Proof of Proposition 4. Thanks to Proposition 3 we know that for every profile (possibly composed only by maximal coherent sets of gambles) $[\mathscr{D}_i]$, $\Gamma([\mathscr{D}_i]) = \cap_i \mathscr{D}_i$. Hence it satisfies anonymity. □

Proof of Lemma 6. That $\mathscr{E}$ is coherent follows from [Miranda et al., 2012, Prop. 29].

From Eq. (2.2), any $f \in \mathscr{E}$ can be written as

$$f = \sum_{i \in \mathscr{V}} \lambda_i \mathbb{I}_i \otimes f_i + f_0,$$

with $f_0 \in \mathscr{L}^+(\mathscr{V} \times \mathscr{X}) \cup \{0\}, f_i \in \mathscr{D}_i \cup \{0\}, \lambda_i > 0, f \neq 0$. Then $f \in \mathrm{Marg}_{\mathscr{X}}(\mathscr{E})$ iff $f \in \mathscr{E}$ and

$(\forall i, j \in \mathscr{V})(\forall z \in \mathscr{X})\, f(i,z) := \lambda_i \mathbb{I}_i \otimes f_i(i,z) + f_0(i,z) = \lambda_j \mathbb{I}_j \otimes f_j(j,z) + f_0(j,z) =: f(j,z)$.

This is equivalent to

$$(\forall i, j \in \mathscr{V})(\forall z \in \mathscr{X})\ \lambda_i f_i(z) + f_0(i,z) = \lambda_j f_j(z) + f_0(j,z).$$

You can observe that, fixing i and j, the left term of this equation is a gamble in $\mathscr{D}_i$ and the right term is a gamble in $\mathscr{D}_j$, considering also that $f \neq 0$. Thus, $f \in \mathrm{Marg}_{\mathscr{X}}(\mathscr{E})$ if and only if it depends only on $z \in \mathscr{X}$ and, as a function of only $z \in \mathscr{X}$, it belongs to $\cap_{i \in \mathscr{V}} \mathscr{D}_i$. Hence we have the thesis. □

Proof of Theorem 6. Let us address the points of the statement in turn.

1. The equality of the sets in Eqs. (2.2) and (6) is well known, see for instance [Miranda et al., 2012, Prop. 29].

2. The converse implication is trivial: if such an $\mathscr{E}'$ exists, then $\Gamma([\mathscr{D}_i]) = \mathrm{Marg}_{\mathscr{X}}(\mathscr{E}') \supseteq \mathrm{Marg}_{\mathscr{X}}(\mathscr{E})$.

   For the direct implication, consider $\mathscr{E}'$ as defined in Eq. (2.3). It includes $\mathscr{E}$ by definition. To prove that it is coherent, note that $\mathscr{E}' = \mathrm{posi}(\Gamma([\mathscr{D}_i]) \otimes \mathscr{V} \cup \cup_i \mathscr{D}|i)$ (this is due again to the transformation referenced in point 1).

It includes $\mathscr{L}^+(\mathscr{V} \times \mathscr{X})$ because for every $f \in \mathscr{L}^+(\mathscr{V} \times \mathscr{X})$, the restriction of $\mathbb{I}_i \otimes f$ on $\mathscr{L}(\mathscr{X})$ belongs to $\mathscr{D}_i \cup \{0\}$ for all $i \in \mathscr{V}$. Thus, coherence holds if and only if $0 \notin \text{posi}(\Gamma([\mathscr{D}_i]) \otimes \mathscr{V} \cup \cup_i \mathscr{D}|i)$.

To prove that this is the case, let us reason by contradiction. Note that the zero gamble can in principle be produced only by adding $\mathbb{I}_{\mathscr{V}} \otimes f_0$ with $\sum_{i\in\mathscr{V}} \mathbb{I}_i \otimes f_i$, for some $f_0 \in \Gamma([\mathscr{D}_i]) \cup \{0\}, (\forall i \in \mathscr{V}) f_i \in \mathscr{D}_i \cup \{0\}$, not all of them zero. In order for the sum to yield zero, f_0 must be different from zero. And since $\mathbb{I}_{\mathscr{V}} \otimes f_0$ is $\mathscr{X}$-measurable, in order to yield zero it must hold that $f_1 = \cdots = f_n = -f_0$. This implies that $-f_0 \in \cap_{i\in\mathscr{V}} \mathscr{D}_i$. We are showing the direct implication, so it follows by hypothesis $\Gamma([\mathscr{D}_i]) \supseteq \text{Marg}_{\mathscr{X}}(\mathscr{E})$ and Lemma 6, that $-f_0 \in \Gamma([\mathscr{D}_i])$. But then $\Gamma([\mathscr{D}_i])$ contains both f_0 and $-f_0$. This contradicts the coherence of $\Gamma([\mathscr{D}_i])$. Therefore the sum of two gambles in $\mathscr{E}'$ must be different from zero.

To conclude, we show that $\Gamma([\mathscr{D}_i]) = \text{Marg}_{\mathscr{X}}(\mathscr{E}')$. Since the direct inclusion is trivial, we focus on the converse inclusion. Remember that $\text{Marg}_{\mathscr{X}}(\mathscr{E}') = \mathscr{E}' \cap \mathscr{L}_{\mathscr{X}}(\mathscr{V} \times \mathscr{X})$. Let us consider a $\mathscr{X}$-measurable gamble $f \in \mathscr{E}'$: $f = \mathbb{I}_{\mathscr{V}} \otimes f_0 + \sum_{i\in\mathscr{V}} \mathbb{I}_i \otimes f_i$. The case $f_1 = \cdots = f_n = 0$ is trivial, so let us assume that there is $i \in \mathscr{V}$ such that $f_i \neq 0$. Since f is $\mathscr{X}$-measurable, this means that

$$(\forall i, j \in \mathscr{V})(\forall z \in \mathscr{X})\ f(i,z) = f_0(z) + f_i(z) = f_0(z) + f_j(z) = f(j,z).$$

This implies $f_j = f_i$ for all $j \neq i$. As a consequence, $f_i \in \text{Marg}_{\mathscr{X}}(\mathscr{E})$ and by hypothesis then $f_i \in \Gamma([\mathscr{D}_i])$. It follows that $f_0 + f_1 + \cdots + f_n \in \Gamma([\mathscr{D}_i])$, whence, as a function of only z, $f \in \Gamma([\mathscr{D}_i])$.

3. Assume by contradiction that there is such a set $\mathscr{E}''$: $\mathscr{E}' \supsetneq \mathscr{E}'' \supseteq \mathscr{E}$. Since $\mathscr{E}'$ strictly contains $\mathscr{E}''$, there must be some $f_0 \in \Gamma([\mathscr{D}_i])$ such that $\mathbb{I}_{\mathscr{V}} \otimes f_0 \notin \mathscr{E}''$. Note that $f_0 \notin \cap_{i\in\mathscr{V}} \mathscr{D}_i$, otherwise it would belong to $\mathscr{E}''$ given that $\mathscr{E}'' \supseteq \mathscr{E}$. Whence $f_0 \in \Gamma([\mathscr{D}_i]) \setminus \cap_{i\in\mathscr{V}} \mathscr{D}_i$ and by marginalising $\mathscr{E}''$ we obtain a set that does not contain f_0; therefore $\Gamma([\mathscr{D}_i]) \neq \text{Marg}_{\mathscr{X}}(\mathscr{E}'')$. This is a contradiction.

4. Remember that $\mathscr{E}'|i = \{f \in \mathscr{E}' : f = \mathbb{I}_i f\}$. Since $\mathbb{I}_{\mathscr{V}} \otimes f_0$, in the definition of $\mathscr{E}'$, is constant on the elements of $\mathscr{V}$, any gamble $f = \mathbb{I}_i f$ must be such that $f_0 = 0$. The thesis then follows immediately.

$\square$

Theorem 19. *Let Γ be a social rule defined on a set of profiles $\mathscr{A}$ composed by coherent lower previsions (resp., credal sets) $[\underline{P}_i]$ (resp., $[\mathscr{M}_i]$). Then:*

1. *Γ satisfies weak Pareto* $\Leftrightarrow (\forall[\underline{P}_i] \in \mathscr{A})\ \Gamma([\underline{P}_i]) =: \underline{P} \geq \min_i \underline{P}_i$, $\Leftrightarrow$ $(\forall[\mathscr{M}_i] \in \mathscr{A})\ \Gamma([\mathscr{M}_i]) =: \mathscr{M} \subseteq \overline{\mathrm{ch}(\cup_i \mathscr{M}_i)}$.

2. *Γ satisfies strict completeness* $\Leftrightarrow (\forall[\underline{P}_i] \in \mathscr{A})\ \Gamma([\underline{P}_i]) =: \underline{P}$ *linear* $\Leftrightarrow$ $(\forall[\mathscr{M}_i] \in \mathscr{A})\ |\Gamma([\mathscr{M}_i])| =: |\mathscr{M}| = 1$.

3. *A set $\mathscr{G} \subseteq \mathscr{V}$ is decisive* $\Leftrightarrow (\forall[\underline{P}_i] \in \mathscr{A})\ \Gamma([\underline{P}_i]) =: \underline{P} \geq \min_{i\in\mathscr{G}} \underline{P}_i$ $\Leftrightarrow$ $(\forall[\mathscr{M}_i] \in \mathscr{A})\ \Gamma([\mathscr{M}_i]) =: \mathscr{M} \subseteq \overline{\mathrm{ch}(\cup_{i\in\mathscr{G}} \mathscr{M}_i)}$.

4. *A set $\mathscr{G} \subseteq \mathscr{V}$ is an oligarchy* $\Leftrightarrow (\forall[\underline{P}_i] \in \mathscr{A})\ \Gamma([\underline{P}_i]) =: \underline{P} \geq \min_{i\in\mathscr{G}} \underline{P}_i$ *and* $\overline{P} \geq \max_{i\in\mathscr{G}} \underline{P}_i$, *where $\overline{P}$ is the upper prevision associated to $\underline{P}$.*

5. *An individual $j \in \mathscr{V}$ is a dictator* $\Leftrightarrow (\forall[\underline{P}_i] \in \mathscr{A})\ \Gamma([\underline{P}_i]) =: \underline{P} \geq \underline{P}_j$ $\Leftrightarrow$ $\Gamma([\mathscr{M}_i]) =: \mathscr{M} \subseteq \mathscr{M}_j$.

6. *Γ satisfies anonymity* $\Leftrightarrow (\forall[\underline{P}_i], [\underline{P}_{\sigma(i)}] \in \mathscr{A})\ \underline{P} = \underline{P}_\sigma \Leftrightarrow (\forall[\mathscr{M}_i], [\mathscr{M}_{\sigma(i)}] \in \mathscr{A})\ \mathscr{M} = \mathscr{M}_\sigma$, *where σ is a permutation of $\mathscr{V}$ and $\underline{P}$, $\underline{P}_\sigma$ (resp. $\mathscr{M}$, $\mathscr{M}_\sigma$) are the coherent lower previsions (resp. credal set) given by $\Gamma([\underline{P}_i])$, $\Gamma([\underline{P}_{\sigma(i)}])$ (resp. $\Gamma([\mathscr{M}_i])$, $\Gamma([\mathscr{M}_{\sigma(i)}])$).*

7. *Γ satisfies independence of irrelevant alternatives* $\Leftrightarrow (\forall[\underline{P}_i], [\underline{P}'_i] \in \mathscr{A})(\forall f \notin \mathscr{L}^+)$, *given $\underline{P} := \Gamma([\underline{P}_i])$ and $\underline{P}' := \Gamma([\underline{P}'_i])$:*
$$(\forall i \in \mathscr{V})\underline{P}_i(f)\underline{P}'_i(f) > 0 \Rightarrow \underline{P}(f)\underline{P}'(f) > 0.$$

Proof of Theorem 19. For every statement, we shall establish the first equivalence. The second one follows taking into account that:
$$\underline{P} \geq \underline{Q} \iff \mathscr{M}(\underline{P}) \subseteq \mathscr{M}(\underline{Q});$$
the credal set associated with $\min_i \underline{P}_i$ is $\overline{\mathrm{ch}(\cup_i \mathscr{M}(\underline{P}_i))}$; the one associated with $\max_i \underline{P}_i$ is $\cap_i \mathscr{M}(\underline{P}_i)$.

We shall denote by $\mathscr{D}_i^+$ the coherent set of strictly desirable gambles associated with $\underline{P}_i$ by means of Eq. (1.8), and by $\mathscr{D}^+$ the coherent set of strictly desirable gambles associated with $\underline{P} = \Gamma([\underline{P}_i])$.

1. Let us establish the direct implication. Assume that Γ satisfies weak Pareto, and as a consequence that $\cap_{i\in\mathscr{V}} \mathscr{D}_i^+ \subseteq \Gamma'([\mathscr{D}_i^+])$, where Γ' is associated with Γ by means of Eq. (2.4). It follows that, for any $f \in \mathscr{L}$,
$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \Gamma'[\mathscr{D}_i^+]\} \geq \sup\{\mu \in \mathbb{R} : f - \mu \in \cap_{i\in\mathscr{V}} \mathscr{D}_i^+\}.$$
Let us prove that the right-hand side is greater than or equal to $\min_i \underline{P}_i(f)$. Assume that $\min_i \underline{P}_i(f) = \underline{P}_j(f)$. Then it follows from Eq. (1.6) that for any $\epsilon > 0$, $f - \underline{P}_j(f) + \epsilon$ belongs to $\mathscr{D}_j^+$, and for any $j' \neq j$ it holds that

$f - \underline{P}_j(f) + \epsilon \geq f - \underline{P}_{j'}(f) + \epsilon \in \mathcal{D}^+_{j'}$. As a consequence, $f - \underline{P}_j(f) + \epsilon \in \cap_i \mathcal{D}^+_i$, whence

$$\sup\{\mu \in \mathbb{R} : f - \mu \in \cap_{i\in\mathcal{V}} \mathcal{D}^+_i\} \geq \underline{P}_j(f) - \epsilon.$$

Since this holds for any $\epsilon > 0$, it follows that

$$\sup\{\mu \in \mathbb{R} : f - \mu \in \cap_{i\in\mathcal{V}} \mathcal{D}^+_i\} \geq \underline{P}_j(f) = \min_i \underline{P}_i(f).$$

We conclude that $\underline{P}(f) \geq \min_i \underline{P}_i(f)$ for every gamble f.

To see the converse, given a gamble $f \in (\cap_i \mathcal{D}^+_i) \setminus \mathcal{L}^+$, it holds that $\underline{P}_i(f) > 0$ for every i, whence $0 < \min \underline{P}_i(f) \leq \underline{P}(f)$, and as a consequence f belongs to the set of strictly desirable gambles associated with $\underline{P}$. Since trivially $(\cap_i \mathcal{D}^+_i) \cap \mathcal{L}^+ \subseteq \mathcal{L}^+ \subseteq \mathcal{D}^+$, we deduce that Γ' satisfies weak Pareto and therefore so does Γ.

2. This follows from Eqs. (1.6), Definition 13 and Definition 24.

3. This follows applying the first statement to $\mathcal{G}$ instead of $\mathcal{V}$.

4. The first part is a consequence of the third statement. Let us now prove that

$$\left[(\exists i \in \mathcal{G})(f \in \mathcal{D}^+_i)) \Rightarrow -f \notin \Gamma'([\mathcal{D}^+_i])\right] \Leftrightarrow \overline{P} \geq \max_{i\in\mathcal{G}} \underline{P}_i. \tag{B.6}$$

To see the direct implication, consider a gamble f, and let $\underline{P}_j(f) = \max_{i\in\mathcal{G}} \underline{P}_i(f)$. If $\underline{P}_j(f) > 0$, then $f \in \mathcal{D}^+_j$, whence $-f \notin \Gamma'([\mathcal{D}^+_i])$ and, from Eq. (1.6) and Eq. (1.7), it follows that $\overline{P}(f) \geq 0$. Moreover, it must be $\overline{P}(f) > 0$: otherwise by considering $f' := f - \frac{\underline{P}_j(f)}{2}$, we would obtain $\underline{P}_j(f') > 0$ and $\overline{P}(f') < 0$. Therefore, we conclude that $\max_{i\in\mathcal{G}} \underline{P}_i(f) > 0 \Rightarrow \overline{P}(f) > 0$. But since both $\max_{i\in\mathcal{G}} \underline{P}_i$ and $\overline{P}$ satisfy constant additivity, this means that $\overline{P} \geq \max_{i\in\mathcal{G}} \underline{P}_i$.

Conversely, if there is a gamble f such that $f \in \mathcal{D}^+_j$ for some $j \in \mathcal{G}$ while $-f \in \Gamma'([\mathcal{D}^+_i])$, it necessarily must be $f \notin \mathcal{L}^+$. It then follows that $\max_{i\in\mathcal{G}} \underline{P}_i(f) \geq \underline{P}_j(f) > 0$, while $0 < \underline{P}(-f) = -\overline{P}(f)$, meaning that $\overline{P}(f) < 0$. This is a contradiction.

As a consequence, Eq. (B.6) holds. This concludes the proof of this statement.

5. This is a particular case of the third statement with $|\mathcal{G}| = 1$.

6. Γ satisfies anonymity if and only if $\Gamma'([\mathcal{D}^+_i]) = \Gamma'([\mathcal{D}^+_{\sigma(i)}])$. It follows that, for any $f \in \mathcal{L}$,

$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \Gamma'([\mathcal{D}^+_i])\} = \sup\{\mu \in \mathbb{R} : f - \mu \in \Gamma'([\mathcal{D}^+_{\sigma(i)}])\} =: \underline{P}_\sigma(f).$$

To see the converse, let us suppose without loss of generality that there is a gamble $f \in \Gamma'([\mathscr{D}_i^+]) \cap (\Gamma'([\mathscr{D}_{\sigma(i)}^+])^c$ for some permutation σ of $\mathscr{V}$. Then it follows that $\underline{P}(f) > 0$ and $\underline{P}_\sigma(f) \leq 0$ and this contradicts $\underline{P}(f) = \underline{P}_\sigma(f)$.

7. Let us consider $f \notin \mathscr{L}^+$ and two profiles $[\underline{P}_i], [\underline{P}'_i]$ such that $(\forall i \in \mathscr{V})\underline{P}_i(f)\underline{P}'_i(f) > 0$. If Γ satisfies independence of irrelevant alternatives, then $f \in \Gamma'([\mathscr{D}_i^+]) \Leftrightarrow \Gamma'([\mathscr{D}_i'^+])$. As a consequence, we have $\underline{P}(f) > 0 \Leftrightarrow \underline{P}'(f) > 0$, whence $\underline{P}(f)\underline{P}'(f) > 0$.

   Conversely, if $\underline{P}(f)\underline{P}'(f) > 0$ then $f \in \Gamma'([\mathscr{D}_i^+]) \Leftrightarrow \Gamma'([\mathscr{D}_i'^+])$ by Eq. (1.8), whence Γ satisfies independence of irrelevant alternatives. □

Corollary 14. *Let Γ be a strict complete social rule defined on a profile $[P_i]$ given by linear previsions. Then:*

1. *A set $\mathscr{G} \subseteq \mathscr{V}$ is decisive $\Leftrightarrow (\forall [P_i] \in \mathscr{A})\ \Gamma([P_i]) \in \mathrm{ch}(\{P_i \mid i \in \mathscr{G}\})$.*

2. *A set $\mathscr{G} \subseteq \mathscr{V}$ is an oligarchy $\Leftrightarrow (\forall [P_i] \in \mathscr{A}), P_j = P_{j'}\ \forall j, j' \in \mathscr{G}$ and $\Gamma([P_i]) = P_j$.*

3. *An individual $j \in \mathscr{V}$ is a dictator $\Leftrightarrow (\forall [P_i] \in \mathscr{A})\ \Gamma([P_i]) = P_j$.*

Proof of Corollary 2. Weak Pareto means that if a gamble f is strictly desirable for any $i = 1, \ldots, n$, then it should also be strictly desirable for the group, or, equivalently, that $P \geq \min_i P_i$ (item 1 of Theorem 19 in Appendix B). But this means that P belongs to the credal set associated with the coherent lower prevision $\underline{P} := \min_i P_i$, and as a consequence that it is a convex combination of $P_1, \ldots, P_n$. □

Proof of Theorem 7. If π is not degenerate, then we can find $j_1 \neq j_2$ in $\mathscr{V}$ such that $\pi(j_1), \pi(j_2) > 0$. Consider a profile $[P_i] \in \mathscr{A}$ such that P_i assigns all the mass to (s_1, x_1) if $i = j_1$ and all the mass to (s_2, x_2) if $i \neq j_1$, where $s_1 \neq s_2$ and $j_1 \neq j_2$. Then we obtain

$$P(s_1, x_1) = \pi(j_1) = P(s_1) = P(x_1),$$

meaning that Γ is not state independent. This is a contradiction. □

Appendix C

Proofs of Chapter 3

C.1 Proofs of Section 3.1

We begin this appendix by recalling some preliminary results concerning domain-free information algebras.

Lemma 17. *Consider a domain-free information algebra* $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$*. The following properties are valid. For any* $\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\mu} \in \boldsymbol{\Phi}$ *and* $S, T \subseteq I$*, we have*

1. $\mathbf{1} \leq \boldsymbol{\phi} \leq \mathbf{0}$;
2. $\boldsymbol{\phi}, \boldsymbol{\psi} \leq \boldsymbol{\phi} \cdot \boldsymbol{\psi}$;
3. $\boldsymbol{\phi} \leq \boldsymbol{\psi} \Rightarrow \boldsymbol{\phi} \cdot \boldsymbol{\mu} \leq \boldsymbol{\psi} \cdot \boldsymbol{\mu}$;
4. $\boldsymbol{\epsilon}_S(\boldsymbol{\phi}) \leq \boldsymbol{\phi}$;
5. $\boldsymbol{\phi} \leq \boldsymbol{\psi} \Rightarrow \boldsymbol{\epsilon}_S(\boldsymbol{\phi}) \leq \boldsymbol{\epsilon}_S(\boldsymbol{\psi})$;
6. $T \subseteq S \Rightarrow \boldsymbol{\epsilon}_T(\boldsymbol{\phi}) \leq \boldsymbol{\epsilon}_S(\boldsymbol{\phi})$.

Proof. These properties are proven in [Kohlas, 2003, Lemma 6.2]. □

Lemma 18. *Consider a domain-free information algebra* $(\boldsymbol{\Phi}, I; \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$*. The following properties are valid. For any* $\boldsymbol{\phi}, \boldsymbol{\psi} \in \boldsymbol{\Phi}$ *and* $S, T \subseteq I$*, we have*

1. *any* S *is a support of the null* $\mathbf{0}$ *and the unit* $\mathbf{1}$ *elements,*
2. S *is a support of* $\boldsymbol{\epsilon}_S(\boldsymbol{\phi})$*,*
3. S *support of* $\boldsymbol{\phi} \Rightarrow S$ *support of* $\boldsymbol{\epsilon}_T(\boldsymbol{\phi})$*,*

4. *S and T supports of $\phi \Rightarrow S \cap T$ support of ϕ,*

5. *S support of $\phi \Rightarrow \epsilon_T(\phi) = \epsilon_{S\cap T}(\phi)$,*

6. *T support of ϕ and $T \subseteq S \Rightarrow S$ support of ϕ,*

7. *S support of ϕ and $\psi \Rightarrow S$ support of $\phi \cdot \psi$,*

8. *S support of ϕ and T support of $\psi \Rightarrow S \cup T$ support of $\phi \cdot \psi$.*

Proof. These properties are proven in [Kohlas, 2003, Lemma 3.6]. □

In what follows, we consider instead the same framework of Section 3.1. Specifically, we assume a possibility space $\Omega := \bigtimes_{i\in I} \Omega_i$, where I is an index set. We also consider $\mathscr{L}_{\mathscr{P}_S}(\Omega)$, denoted also as $\mathscr{L}_S$, where $\mathscr{P}_S$ is the partition whose blocks are composed by elements ω, ω' of Ω such that $\omega|S = \omega'|S$.

Lemma 19. *If Cl is any closure operator on $(\mathrm{P}(\mathscr{L}), \subseteq)$ then, for any $\mathscr{K}_1, \mathscr{K}_2 \subseteq \mathscr{L}$:*

$$Cl(\mathscr{K}_1 \cup \mathscr{K}_2) = Cl(Cl(\mathscr{K}_1) \cup \mathscr{K}_2).$$

Proof. Obviously, $Cl(\mathscr{K}_1 \cup \mathscr{K}_2) \subseteq Cl(Cl(\mathscr{K}_1) \cup \mathscr{K}_2)$. On the other hand, $\mathscr{K}_1, \mathscr{K}_2 \subseteq \mathscr{K}_1 \cup \mathscr{K}_2$, hence $Cl(\mathscr{K}_1) \subseteq Cl(\mathscr{K}_1 \cup \mathscr{K}_2)$ and $\mathscr{K}_2 \subseteq Cl(\mathscr{K}_1 \cup \mathscr{K}_2)$. This implies $Cl(Cl(\mathscr{K}_1) \cup \mathscr{K}_2) \subseteq Cl(Cl(\mathscr{K}_1 \cup \mathscr{K}_2)) = Cl(\mathscr{K}_1 \cup \mathscr{K}_2)$. □

Lemma 20. *For any subsets S and T of I:*

$$\mathscr{L}_{S\cap T} = \mathscr{L}_S \cap \mathscr{L}_T.$$

Proof. Consider firstly $f \in \mathscr{L}_{S\cap T}$. Consider two elements $\omega, \mu \in \Omega$ so that $\omega|S = \mu|S$. Then we have also $\omega|S \cap T = \mu|S \cap T$ and $f(\omega) = f(\mu)$. So we see that $f \in \mathscr{L}_S$ and similarly $f \in \mathscr{L}_T$.

Conversely, assume $f \in \mathscr{L}_S \cap \mathscr{L}_T$. Consider two elements ω, $\mu \in \Omega$. so that $\omega|S \cap T = \mu|S \cap T$. Consider then the element $\lambda \in \Omega$ defined as

$$(\forall i \in I)\ \lambda_i := \begin{cases} \omega_i = \mu_i, & i \in (S \cap T), \\ \omega_i, & i \in (S \setminus T) \cup (S \cup T)^c, \\ \mu_i, & i \in T \setminus S. \end{cases}$$

Then $\lambda|S = \omega|S$ and $\lambda|T = \mu|T$. Since f is both S- and T-measurable we have $f(\omega) = f(\mu)$. It follows that $f \in \mathscr{L}_{S\cap T}$ and this concludes the proof. □

Proof of Theorem 8. 1. That $(\Phi; \cdot)$ is a commutative semigroup follows from $\mathscr{D}_1 \cdot \mathscr{D}_2 := \mathscr{D}_1 \vee \mathscr{D}_2$, for any $\mathscr{D}_1, \mathscr{D}_2$ in the complete lattice induced by $(\Phi; \subseteq)$, see Section 1. As stated above, $0 = \mathscr{L}$ is the null element and $1 = \mathscr{L}^+$ the unit element of the semigroup (null and unit in a semigroup are always unique).

2. We have
$$\epsilon_S(0) = \epsilon_S(\mathscr{L}) := \mathscr{C}(\mathscr{L} \cap \mathscr{L}_S) = \mathscr{C}(\mathscr{L}_S) = \mathscr{L} = 0,$$
for any $S \subseteq I$.

3. It follows since $\mathscr{D} \cap \mathscr{L}_S \subseteq \mathscr{D}$ and $\mathscr{C}(\mathscr{D} \cap \mathscr{L}_S) \subseteq \mathscr{D}$, for any $\mathscr{D} \in \Phi$, $S \subseteq I$.

4. Let us define, using Lemma 19 in Appendix C.1,
$$\begin{aligned} A &:= \mathscr{C}(\mathscr{C}(\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2) \cap \mathscr{L}_S = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2) \cap \mathscr{L}_S, \\ B &:= \mathscr{C}(\mathscr{C}(\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{C}(\mathscr{D}_2 \cap \mathscr{L}_S)) = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup (\mathscr{D}_2 \cap \mathscr{L}_S)). \end{aligned}$$
Then we have $B := \epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2)$ and $\mathscr{C}(A) := \epsilon_S(\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2)$. Note that $B \subseteq \mathscr{C}(A)$.

   We claim first that:
$$\epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2) = \mathscr{L} \iff \epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2 = \mathscr{L}. \tag{C.1}$$
Indeed, $\epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2) = \mathscr{L}$ implies a fortiori $\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2 = \mathscr{L}$.

   Assume therefore that $\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2 = \mathscr{L}$. This implies $\mathscr{L} = \mathscr{C}(\mathscr{C}(\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2) = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2)$, by Lemma 19 in Appendix C.1. Now, if $\mathscr{D}_1 = \mathscr{L}$ or $\mathscr{D}_2 = \mathscr{L}$ we have immediately the result, otherwise we claim that $0 = f + g'$ with $f \in \mathscr{D}_1 \cap \mathscr{L}_S$ and $g' \in \mathscr{D}_2 \cap \mathscr{L}_S$. Indeed, from $\mathscr{L} = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2)$, we know that $0 \in \mathscr{E}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2)$ therefore $0 = f + g + h'$ with $f \in \mathscr{D}_1 \cap \mathscr{L}_S$, $g \in \mathscr{D}_2$, $h' \in \mathscr{L}^+(\Omega) \subseteq \mathscr{D}_2$ or $h' = 0$. Then, if we introduce $g' = g + h'$, we have $0 = f + g'$ with $f \in \mathscr{D}_1 \cap \mathscr{L}_S$, $g' \in \mathscr{D}_2$. However, this implies $g' = -f \in \mathscr{L}_S$ and then $g' \in \mathscr{D}_2 \cap \mathscr{L}_S$. Notice that $\epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2) =: B = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup (\mathscr{D}_2 \cap \mathscr{L}_S))$. Therefore, we have the result.

   So, if $\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2 = \mathscr{L}$ or $B = \mathscr{L}$, then $\mathscr{C}(A) = \mathscr{L}$ and $B = \mathscr{L}$. Therefore we have $\mathscr{C}(A) \subseteq B$.

   Viceversa, assume both $\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2$ and $\epsilon_S(\mathscr{D}_1) \cdot \epsilon_S(\mathscr{D}_2)$ coherent. Therefore $\epsilon_S(\mathscr{D}_1) \cdot \mathscr{D}_2 = \mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2) = \mathscr{E}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup \mathscr{D}_2)$. Then we have
$$A = \{f \in \mathscr{L}_S : f \geq \lambda g + \mu h, g \in \mathscr{D}_1 \cap \mathscr{L}_S, h \in \mathscr{D}_2, \lambda, \mu \geq 0, f \neq 0\}.$$

   Consider $f \in A$. Then $f = \lambda g + \mu h + h'$, where $h' \in \mathscr{L}^+ \cup \{0\}$. Since f and g are S-measurable, $\mu h + h'$ must be S-measurable. Now, if $\mu h + h' = 0$ then $f \in \mathscr{D}_1 \cap \mathscr{L}_S \subseteq B$. Otherwise, $\mu h + h' \in \mathscr{D}_2 \cap \mathscr{L}_S$. So in any case $f \in B$, hence we have $\mathscr{C}(A) \subseteq \mathscr{C}(B) = B$.

5. Note first that $\epsilon_S(\epsilon_T(\mathscr{D})) = 0$ and $\epsilon_{S \cap T}(\mathscr{D}) = 0$ if and only if $\mathscr{D} = 0$. So assume $\mathscr{D}$ to be coherent. Then we have, by Lemma 20 in Appendix C.1,
$$\begin{aligned} \epsilon_S(\epsilon_T(\mathscr{D})) &:= \mathscr{C}(\mathscr{C}(\mathscr{D} \cap \mathscr{L}_T) \cap \mathscr{L}_S), \\ \epsilon_{S \cap T}(\mathscr{D}) &:= \mathscr{C}(\mathscr{D} \cap \mathscr{L}_{S \cap T}) = \mathscr{C}(\mathscr{D} \cap \mathscr{L}_T \cap \mathscr{L}_S). \end{aligned}$$

Obviously, $\epsilon_{S\cap T}(\mathscr{D}) \subseteq \epsilon_S(\epsilon_T(\mathscr{D}))$. Consider then $f \in \mathscr{C}(\mathscr{D} \cap \mathscr{L}_T) \cap \mathscr{L}_S = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_T) \cap \mathscr{L}_S$. If $f \in \mathscr{L}_S^+$ then clearly $f \in \epsilon_{S\cap T}(\mathscr{D})$. Otherwise,

$$f \in \mathscr{L}_S, \quad f \geq g, \quad g \in \mathscr{D} \cap \mathscr{L}_T.$$

Define

$$g'(\omega) := \sup_{\lambda|S=\omega|S} g(\lambda).$$

Then we have $f \geq g'$. Clearly g' is S-measurable and belongs to $\mathscr{D}$, $g' \in \mathscr{D} \cap \mathscr{L}_S$. We claim that g' is also T-measurable. Consider two elements ω and μ so that $\omega|S \cap T = \mu|S \cap T$. Note that we may write

$$g'(\omega) := \sup_{\lambda|S=\omega|S} g(\lambda) = \sup_{\lambda|I\setminus S} g(\omega|S \cap T, \omega|S \setminus T, \lambda|T \setminus S, \lambda|R),$$

where $R = (S \cup T)^c$. Similarly, we have

$$g'(\mu) := \sup_{\lambda'|S=\mu|S} g(\lambda') = \sup_{\lambda'|I\setminus S} g(\omega|S \cap T, \mu|S \setminus T, \lambda'|T \setminus S, \lambda'|R).$$

Since g is T-measurable, we have:

$$g'(\mu) = \sup_{\lambda'|I\setminus S} g(\omega|S \cap T, \omega|S \setminus T, \lambda'|T \setminus S, \lambda'|R),$$

that clearly coincides with $g'(\omega)$.

This shows that g' is $S\cap T$-measurable, therefore both S- and T-measurable by Lemma 20 in Appendix C.1. So we have $g' \in \mathscr{D} \cap \mathscr{L}_S \cap \mathscr{L}_T$, hence $f \in \mathscr{C}(\mathscr{D} \cap \mathscr{L}_T \cap \mathscr{L}_S)$. And hence $\mathscr{C}(\mathscr{C}(\mathscr{D} \cap \mathscr{L}_T) \cap \mathscr{L}_S) \subseteq \mathscr{C}(\mathscr{C}(\mathscr{D} \cap \mathscr{L}_T \cap \mathscr{L}_S)) = \mathscr{C}(\mathscr{D} \cap \mathscr{L}_T \cap \mathscr{L}_S)$.

Analogously, we can prove that $\epsilon_T(\epsilon_S(\mathscr{D})) = \epsilon_{S\cap T}(\mathscr{D})$.

6. It is obvious.

□

To introduce the following result, we recall the operation ${}^{\downarrow S} : \mathscr{L}_S(\Omega_R) \to \mathscr{L}(\Omega_S)$ with $S \subseteq R \subseteq I$, such that $(\forall \omega_S \in \Omega_S)\, f^{\downarrow S}(\omega_S) := f(\omega_R)$, where $\omega_R \in \Omega_R$ and $\omega_R|S = \omega_S$, for every $f \in \mathscr{L}_S(\Omega_R)$. We also recall its inverse ${}^{\uparrow R} : \mathscr{L}(\Omega_S) \to \mathscr{L}_S(\Omega_R)$ such that $(\forall \omega_R \in \Omega_R)\, f^{\uparrow R}(\omega_R) := f(\omega_R|S)$, for every $f \in \mathscr{L}(\Omega_S)$. These operations can be easily extended to sets of gambles, as follows. Given $\mathscr{K} \subseteq \mathscr{L}_S(\Omega_R)$:

$$\mathscr{K}^{\downarrow S} := \{f' \in \mathscr{L}(\Omega_S) : f' = f^{\downarrow S} \text{ for some } f \in \mathscr{K}\}.$$

Analogously, given $\mathscr{K} \subseteq \mathscr{L}(\Omega_S)$:

$$\mathscr{K}^{\uparrow R} := \{f' \in \mathscr{L}_S(\Omega_R) : f' = (f)^{\uparrow R} \text{ for some } f \in \mathscr{K}\}. \tag{C.2}$$

Lemma 21. *Consider $T \subseteq S \subseteq R \subseteq I$. The following properties are valid.*

1. $\mathscr{K} \subseteq \mathscr{L}_T(\Omega_R) \Rightarrow \mathscr{K}^{\downarrow S} = (\mathscr{K}^{\downarrow T})^{\uparrow S}$. *So, in particular,* $S = R \Rightarrow \mathscr{K} = (\mathscr{K}^{\downarrow T})^{\uparrow R}$.
2. $\mathscr{K} \subseteq \mathscr{L}_T(\Omega_S) \Rightarrow \mathscr{K}^{\downarrow T} = (\mathscr{K}^{\uparrow R})^{\downarrow T}$.
3. $\mathscr{K}_1, \mathscr{K}_2 \subseteq \mathscr{L}_T(\Omega_R) \Rightarrow \mathscr{K}_1^{\downarrow T} \cap \mathscr{K}_2^{\downarrow T} = (\mathscr{K}_1 \cap \mathscr{K}_2)^{\downarrow T}$.
4. $\mathscr{K}_1, \mathscr{K}_2 \subseteq \mathscr{L}_T(\Omega_R) \Rightarrow \mathscr{K}_1^{\downarrow T} \cup \mathscr{K}_2^{\downarrow T} = (\mathscr{K}_1 \cup \mathscr{K}_2)^{\downarrow T}$.
5. $\mathscr{K} \subseteq \mathscr{L}_T(\Omega_R) \Rightarrow (\mathscr{C}(\mathscr{K}) \cap \mathscr{L}_T)^{\downarrow T} = \mathscr{C}(\mathscr{K}^{\downarrow T})$.

Proof.

1. Notice that $\mathscr{L}_T(\Omega_R) \subseteq \mathscr{L}_S(\Omega_R)$, hence $\mathscr{K}^{\downarrow S}$ is well-defined. The result then follows by definition.
2. Notice that $\mathscr{L}_T(\Omega_S) \subseteq \mathscr{L}(\Omega_S)$, hence $\mathscr{K}^{\uparrow R}$ is well-defined. Notice moreover that $\mathscr{K}^{\uparrow R} \subseteq \mathscr{L}_T(\Omega_R)$, hence also $(\mathscr{K}^{\uparrow R})^{\downarrow T}$ is well-defined. The result then follows by definition.
3. The result follows by definition.
4. The result follows by definition.
5. $0 \in \mathscr{E}(\mathscr{K}) \iff 0 \in \mathscr{E}(\mathscr{K}^{\downarrow T})$. Therefore, we need to show only that $(\mathscr{E}(\mathscr{K}) \cap \mathscr{L}_T)^{\downarrow T} = \mathscr{E}(\mathscr{K}^{\downarrow T})$ with $0 \notin \mathscr{E}(\mathscr{K}), \mathscr{E}(\mathscr{K}^{\downarrow T})$. So, consider $f' \in (\mathscr{E}(\mathscr{K}) \cap \mathscr{L}_T)^{\downarrow T}$. Then $f' = f^{\downarrow T}$, for some $f \in \mathscr{E}(\mathscr{K}) \cap \mathscr{L}_T$, so, for every $\omega_T \in \Omega_T$, $f'(\omega_T) = f(\omega_R) = \sum_{i=1}^r \lambda_i g_i(\omega_R) + \mu h(\omega_R)$, with $\lambda_i, \mu \geq 0$, $\forall i$ not all equal to 0, $r \geq 0$, $g_i \in \mathscr{K} \subseteq \mathscr{L}_T(\Omega_R)$, $h \in \mathscr{L}^+$, for every $\omega_R \in \Omega_R$ such that $\omega_R|T = \omega_T$. Therefore $h \in \mathscr{L}_T^+$. So, $f' = \sum_{i=1}^r \lambda_i g_i^{\downarrow T} + \mu h^{\downarrow T}$, therefore $f' \in \mathscr{E}(\mathscr{K}^{\downarrow T})$. The other inclusion can be proven analogously, hence we have the result.

$\square$

Proof of Theorem 9.

1. Since $\mathscr{D}_1$ has support S and $\mathscr{D}_2$ has support T, we have

$$\begin{aligned} h((\mathscr{D}_1, S) \cdot (\mathscr{D}_2, T)) &:= h(\mathscr{D}_1 \cdot \mathscr{D}_2, S \cup T) \\ &:= ((\mathscr{D}_1 \cdot \mathscr{D}_2 \cap \mathscr{L}_{S \cup T})^{\downarrow S \cup T}, S \cup T) \\ &= ((\mathscr{C}((\mathscr{D}_1 \cap \mathscr{L}_S) \cup (\mathscr{D}_2 \cap \mathscr{L}_T)) \cap \mathscr{L}_{S \cup T})^{\downarrow S \cup T}, S \cup T) \\ &= (\mathscr{C}(((\mathscr{D}_1 \cap \mathscr{L}_S) \cup (\mathscr{D}_2 \cap \mathscr{L}_T))^{\downarrow S \cup T}), S \cup T) \end{aligned}$$

thanks to Lemma 19 in Appendix C.1, and item 5 of Lemma 21 in Appendix C.1. On the other hand, thanks again to Lemma 19 in Appendix C.1, we have

$$\begin{aligned} h(\mathscr{D}_1, S) \cdot h(\mathscr{D}_2, T) &:= ((\mathscr{D}_1 \cap \mathscr{L}_S)^{\downarrow S}, S) \cdot ((\mathscr{D}_2 \cap \mathscr{L}_T)^{\downarrow T}, T) := \\ &(\mathscr{C}(((\mathscr{D}_1 \cap \mathscr{L}_S)^{\downarrow S})^{\uparrow S \cup T} \cup ((\mathscr{D}_2 \cap \mathscr{L}_T)^{\downarrow T})^{\uparrow S \cup T}), S \cup T). \end{aligned}$$

Now, using again properties of Lemma 21 in Appendix C.1, we have

$$
\begin{aligned}
h(\mathscr{D}_1,S)\cdot h(\mathscr{D}_2,T)&\\
&:= (\mathscr{C}(((\mathscr{D}_1\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T}\cup((\mathscr{D}_2\cap\mathscr{L}_T)^{\downarrow T})^{\uparrow S\cup T}),S\cup T)\\
&= (\mathscr{C}((\mathscr{D}_1\cap\mathscr{L}_S)^{\downarrow S\cup T}\cup(\mathscr{D}_2\cap\mathscr{L}_T)^{\downarrow S\cup T}),S\cup T)\\
&= (\mathscr{C}(((\mathscr{D}_1\cap\mathscr{L}_S)\cup(\mathscr{D}_2\cap\mathscr{L}_T))^{\downarrow S\cup T}),S\cup T)=h((\mathscr{D}_1,S)\cdot(\mathscr{D}_2,T)).
\end{aligned}
$$

2. Obviously, $(\mathscr{L}(\Omega),S)$ maps to $(\mathscr{L}(\Omega_S),S)$.

3. Similarly, $(\mathscr{L}^+(\Omega),S)$ maps to $(\mathscr{L}^+(\Omega_S),S)$.

4. We have, again by definition,

$$
\begin{aligned}
h(\pi_T(\mathscr{D},S)) &:= h(\epsilon_T(\mathscr{D}),T)\\
&:= ((\epsilon_T(\mathscr{D})\cap\mathscr{L}_T)^{\downarrow T},T)\\
&= ((\mathscr{D}\cap\mathscr{L}_T)^{\downarrow T},T).
\end{aligned}
$$

Indeed, $\mathscr{D}\cap\mathscr{L}_T\subseteq\mathscr{C}(\mathscr{D}\cap\mathscr{L}_T)\cap\mathscr{L}_T =: \epsilon_T(\mathscr{D})\cap\mathscr{L}_T\subseteq\mathscr{D}\cap\mathscr{L}_T$. However, from $T\subseteq S$, it follows $\mathscr{L}_T\subseteq\mathscr{L}_S$. Therefore we have

$$
\begin{aligned}
h(\pi_T(\mathscr{D},S)) &= ((\mathscr{D}\cap\mathscr{L}_T)^{\downarrow T},T)\\
&= ((\mathscr{D}\cap\mathscr{L}_S)\cap\mathscr{L}_T)^{\downarrow T},T).
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
\pi_T(h(\mathscr{D},S)) &:= \pi_T((\mathscr{D}\cap\mathscr{L}_S)^{\downarrow S},S)\\
&:= ((\epsilon_T((\mathscr{D}\cap\mathscr{L}_S)^{\downarrow S})\cap\mathscr{L}_T(\Omega_S))^{\downarrow T},T)\\
&= (((\mathscr{D}\cap\mathscr{L}_S)^{\downarrow S}\cap\mathscr{L}_T(\Omega_S))^{\downarrow T},T)\\
&= (((\mathscr{D}\cap\mathscr{L}_S)^{\downarrow S}\cap(\mathscr{L}_T(\Omega))^{\downarrow S})^{\downarrow T},T)\\
&= (((\mathscr{D}\cap\mathscr{L}_S)\cap\mathscr{L}_T)^{\downarrow S})^{\downarrow T},T)\\
&= ((((\mathscr{D}\cap\mathscr{L}_S)\cap\mathscr{L}_T)^{\downarrow S})^{\uparrow I})^{\downarrow T},T)\\
&= ((\mathscr{D}\cap\mathscr{L}_S)\cap\mathscr{L}_T)^{\downarrow T},T)=h(\pi_T(\mathscr{D},S)),
\end{aligned}
$$

thanks to Lemma 21 in Appendix C.1.

5. Suppose $h(\mathscr{D}_1,S)=h(\mathscr{D}_2,T)$. Then we have $S=T$ and $(\mathscr{D}_1\cap\mathscr{L}_S)^{\downarrow S}=(\mathscr{D}_2\cap\mathscr{L}_S)^{\downarrow S}$, from which we derive that $\mathscr{D}_1\cap\mathscr{L}_S=\mathscr{D}_2\cap\mathscr{L}_S$ and therefore, $\mathscr{D}_1=\mathscr{C}(\mathscr{D}_1\cap\mathscr{L}_S)=\mathscr{C}(\mathscr{D}_2\cap\mathscr{L}_S)=\mathscr{D}_2$. So the map h is injective.

   Moreover, for any $(\tilde{\mathscr{D}},S)\in\tilde{\Phi}$ we have that $(\tilde{\mathscr{D}},S)=h(\mathscr{D},S)$ where $(\mathscr{D},S)=(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}),S)\in\hat{\Phi}$. Indeed:

   - $(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}),S)\in\hat{\Phi}$. In fact, $\epsilon_S(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})):=\mathscr{C}(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})\cap\mathscr{L}_S)$. Now, $\tilde{\mathscr{D}}^{\uparrow I}\subseteq\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})\cap\mathscr{L}_S$, therefore $\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})\subseteq\mathscr{C}(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})\cap\mathscr{L}_S)$. On the other hand, $\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}) \cap \mathscr{L}_S \subseteq \mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})$, therefore $\mathscr{C}(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}) \cap \mathscr{L}_S) \subseteq \mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})$. Hence, $\epsilon_S(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})) := \mathscr{C}(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}) \cap \mathscr{L}_S) = \mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})$.

- $h(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}), S) = (\tilde{\mathscr{D}}, S)$. In fact, $h(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}), S) := ((\epsilon_S(\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I})) \cap \mathscr{L}_S)^{\downarrow S}, S) = ((\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}) \cap \mathscr{L}_S)^{\downarrow S}, S)$ by previous item. Moreover, $((\mathscr{C}(\tilde{\mathscr{D}}^{\uparrow I}) \cap \mathscr{L}_S)^{\downarrow S}, S) = (\tilde{\mathscr{D}}, S)$ by item 5 of Lemma 21 in Appendix C.1.

So h is surjective, hence bijective.

□

Now, we introduce some preliminary results needed to prove the subsequent Theorem 10.

Lemma 22. *Consider a non-empty set of gambles $\mathscr{K} \subseteq \mathscr{L}(\Omega)$ and its associated lower prevision $\sigma(\mathscr{K})$ defined in Eq. (1.6). The following are true.*

1. $\mathscr{K} \subseteq dom(\sigma(\mathscr{K}))$.
2. $0 \notin \mathscr{E}(\mathscr{K}) \Rightarrow (\forall f \in \mathscr{K})\, \sigma(\mathscr{K})(f) \in \mathbb{R}$.
3. $\mathscr{K} \in \mathbb{D}(\Omega) \Rightarrow dom(\sigma(\mathscr{K})) = \mathscr{L}(\Omega)$ *and* $(\forall f \in \mathscr{L}(\Omega))\, \sigma(\mathscr{K})(f) \in \mathbb{R}$.
4. $\mathscr{K} \in \overline{\mathbb{D}}(\Omega) \Rightarrow dom(\sigma(\mathscr{K})) = \mathscr{L}(\Omega)$ *and* $(\forall f \in \mathscr{L}(\Omega))\, \sigma(\mathscr{K})(f) \in \mathbb{R}$.

Proof.

1. Consider $f \in \mathscr{K}$. Then the set $\{\mu \in \mathbb{R} : f - \mu \in \mathscr{K}\}$ is not empty, since it contains at least 0.
2. Assume $f - \mu \in \mathscr{K}$. If $\mu \geq \sup f$, then $f(\omega) - \mu \leq 0$ for all ω, but then $0 \in \mathscr{E}(\mathscr{K})$, contrary to the assumption. So, the set $\{\mu \in \mathbb{R} : f - \mu \in \mathscr{K}\}$ is not empty and bounded from above for every $f \in \mathscr{K}$.
3. This result is well-known in literature. See for instance [Troffaes and de Cooman, 2014, Section 4].
4. This result is well-known in literature. See for instance [Troffaes and de Cooman, 2014, Section 4].

□

Lemma 23. *Let $\{\mathscr{D}_j\}_{j \in J}$ be any family of coherent sets. Then we have*

$$\sigma(\bigcap_{j \in J} \mathscr{D}_j) = \inf\{\sigma(\mathscr{D}_j) : j \in J\}.$$

Proof. Note that the intersection of the coherent sets $\mathcal{D}_j$ equals a coherent set $\mathcal{D}$. Moreover, $\inf\{\sigma(\mathcal{D}_j) : j \in J\}$ is coherent [Walley, 1991, Theorem 2.6.3]. We have $\sigma(\bigcap_{j\in J} \mathcal{D}_j) := \sigma(\mathcal{D}) =: \underline{P} \leq \sigma(\mathcal{D}_j)$, for all $j \in J$. So $\underline{P}(f) \leq \sigma(\mathcal{D}_j)(f)$ for all $f \in \mathcal{L}$ and $j \in J$, therefore $\underline{P} \leq \inf\{\sigma(\mathcal{D}_j) : j \in J\}$. However, given the fact that $\inf\{\sigma(\mathcal{D}_j) : j \in J\}$ is coherent, we have also $\tau^+(\inf\{\sigma(\mathcal{D}_j) : j \in J\}) \subseteq \tau^+(\sigma(\mathcal{D}_j)) \subseteq \mathcal{D}_j$ for all $j \in J$, by definition of $\inf\{\sigma(\mathcal{D}_j) : j \in J\}$, where τ^+ is defined in Section 1.1.1. Hence, $\tau^+(\inf\{\sigma(\mathcal{D}_j) : j \in J\}) \subseteq \bigcap_j \mathcal{D}_j =: \mathcal{D}$. But this implies $\inf\{\sigma(\mathcal{D}_j) : j \in J\} \leq \sigma(\mathcal{D}) =: \underline{P}$. This concludes the proof. □

Theorem 20. *Let $\mathcal{K} \subseteq \mathcal{L}$ be a non-empty set of gambles which satisfies the following two conditions:*

1. $0 \notin \mathcal{E}(\mathcal{K})$,
2. *for all $f \in \mathcal{K} \setminus \mathcal{L}^+$ there exists a $\delta > 0$ such that $f - \delta \in \mathcal{K}$.*

Then we have

$$\sigma(\mathcal{C}(\mathcal{K})) = \sigma(\mathcal{E}(\mathcal{K})) = \underline{E}^*(\sigma(\mathcal{K})) = \underline{E}(\sigma(\mathcal{K})),$$

where $\underline{E}^*$ is defined in Definition 17.

Proof. $\mathcal{E}(\mathcal{K})$ is, in particular, a coherent set of strictly desirable gambles. Therefore, we have:

$$\sigma(\mathcal{C}(\mathcal{K})) = \sigma(\mathcal{E}(\mathcal{K})) = \sigma(\bigcap\{\mathcal{D} \in \mathbb{D} : \mathcal{K} \subseteq \mathcal{D}\}) = \sigma(\bigcap\{\mathcal{D}^+ \in \mathbb{D}^+ : \mathcal{K} \subseteq \mathcal{D}^+\}).$$

Now,

$$(\forall \mathcal{D}^+ \in \mathbb{D}^+)\ \mathcal{K} \subseteq \mathcal{D}^+ \iff \sigma(\mathcal{K}) \leq \sigma(\mathcal{D}^+). \tag{C.3}$$

Indeed, $\mathcal{K} \subseteq \mathcal{D}^+ \Rightarrow \sigma(\mathcal{K}) \leq \sigma(\mathcal{D}^+)$, by definition of lower prevision (see Eq. (1.6)). Vice versa, let us consider $\sigma(\mathcal{D}^+)$ such that $\sigma(\mathcal{K}) \leq \sigma(\mathcal{D}^+)$. If $f \in \mathcal{K}$, then $\sigma(\mathcal{K})(f) \geq 0$. If $f \in \mathcal{L}^+$, then $f \in \mathcal{D}^+$. Otherwise, there is $\delta > 0$ such that $f - \delta \in \mathcal{K}$, by assumption. Hence $0 < \sigma(\mathcal{K})(f) \leq \sigma(\mathcal{D}^+)(f)$. This means $f \in \mathcal{D}^+$. So, thanks to Lemma 23 in Appendix C.1, we have:

$$\sigma(\bigcap\{\mathcal{D}^+ \in \mathbb{D}^+ : \mathcal{K} \subseteq \mathcal{D}^+\}) = \inf\{\underline{P} \in \underline{\mathbb{P}} : \sigma(\mathcal{K}) \leq \underline{P}\} = \underline{E}^*(\sigma(\mathcal{K})) = \underline{E}(\sigma(\mathcal{K})).$$

This concludes the proof. □

Proof of Theorem 10.

1. It follows from the definition.
2. Assume first that $\mathcal{D}_1^+ \cdot \mathcal{D}_2^+ = 0$ and let $\underline{P}_1 := \sigma(\mathcal{D}_1^+)$, $\underline{P}_2 := \sigma(\mathcal{D}_2^+)$. Then there can be no coherent lower prevision $\underline{P}$ dominating both $\underline{P}_1$ and $\underline{P}_2$. Indeed, otherwise we would have $\mathcal{D}_1^+ =: \tau^+(\underline{P}_1) \leq \tau^+(\underline{P})$ and $\mathcal{D}_2^+ =: \tau^+(\underline{P}_2) \leq \tau^+(\underline{P})$, where $\tau^+(\underline{P})$ is a coherent set of strictly desirable gambles. But

this is a contradiction. The vice versa is also true. Therefore, we have $\sigma(\mathscr{D}_1^+ \cdot \mathscr{D}_2^+)(f) = \infty = (\sigma(\mathscr{D}_1^+) \cdot \sigma(\mathscr{D}_2^+))(f)$, for all gambles $f \in \mathscr{L}$.

Let then $\mathscr{D}_1^+ \cdot \mathscr{D}_2^+ \neq 0$. Then $\mathscr{D}_1^+ \cdot \mathscr{D}_2^+$ as well as $\mathscr{D}_1^+ \cup \mathscr{D}_2^+$ satisfy the condition of Theorem 20 in Appendix C.1. Therefore, applying this theorem, we have

$$\begin{aligned}\sigma(\mathscr{D}_1^+ \cdot \mathscr{D}_2^+) &:= \sigma(\mathscr{C}(\mathscr{D}_1^+ \cup \mathscr{D}_2^+)) = \sigma(\mathscr{E}(\mathscr{D}_1^+ \cup \mathscr{D}_2^+)) \\ &= \underline{E}(\sigma(\mathscr{D}_1^+ \cup \mathscr{D}_2^+)) = \underline{E}(\max\{\sigma(\mathscr{D}_1^+), \sigma(\mathscr{D}_2^+)\}) =: \sigma(\mathscr{D}_1^+) \cdot \sigma(\mathscr{D}_2^+).\end{aligned}$$

3. We remark that $\mathscr{D}^+ \cap \mathscr{L}_S$ satisfies the conditions of Theorem 20 in Appendix C.1. Thus we obtain

$$\sigma(\epsilon_S(\mathscr{D}^+)) := \sigma(\mathscr{C}(\mathscr{D}^+ \cap \mathscr{L}_S)) = \sigma(\mathscr{E}(\mathscr{D}^+ \cap \mathscr{L}_S)) = \underline{E}(\sigma(\mathscr{D}^+ \cap \mathscr{L}_S)).$$

Now,

$\sigma(\mathscr{D}^+ \cap \mathscr{L}_S)(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathscr{D}^+ \cap \mathscr{L}_S\}, \ \forall f \in dom(\sigma(\mathscr{D}^+ \cap \mathscr{L}_S))$.

But, $f - \mu \in \mathscr{D}^+ \cap \mathscr{L}_S$ if and only if f is S-measurable and $f - \mu \in \mathscr{D}^+$. Therefore, we conclude that $\sigma(\mathscr{D}^+ \cap \mathscr{L}_S) = \sigma(\mathscr{D}^+)_S$. Thus, we have $\sigma(\epsilon_S(\mathscr{D}^+)) = \underline{E}(\sigma(\mathscr{D}^+)_S) =: \underline{e}_S(\sigma(\mathscr{D}^+))$.

□

Proof of Theorem 11. 1. $\mathscr{D}_1^+ \subseteq \mathscr{D}_1$ and $\mathscr{D}_2^+ \subseteq \mathscr{D}_2$, so that

$$\mathscr{D}_1^+ \cdot \mathscr{D}_2^+ = \tau^+(\sigma(\mathscr{D}_1^+ \cdot \mathscr{D}_2^+)) \subseteq \tau^+(\sigma(\mathscr{D}_1 \cdot \mathscr{D}_2)) =: (\mathscr{D}_1 \cdot \mathscr{D}_2)^+.$$

Further

$$(\mathscr{D}_1 \cdot \mathscr{D}_2)^+ := \tau^+(\sigma(\mathscr{D}_1 \cdot \mathscr{D}_2)) := \{f \in \mathscr{L} : \sigma(\mathscr{D}_1 \cdot \mathscr{D}_2)(f) > 0\} \cup \mathscr{L}^+.$$

So, if $f \in (\mathscr{D}_1 \cdot \mathscr{D}_2)^+$, then either $f \in \mathscr{L}^+$ or

$$\sigma(\mathscr{D}_1 \cdot \mathscr{D}_2)(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathscr{C}(\mathscr{D}_1 \cup \mathscr{D}_2)\} > 0. \tag{C.4}$$

In the first case obviously $f \in \mathscr{D}_1^+ \cdot \mathscr{D}_2^+$. Let us consider now $f \notin \mathscr{L}^+$, in this case there is a $\delta > 0$ so that $f - \delta \in \mathscr{C}(\mathscr{D}_1 \cup \mathscr{D}_2) = \mathscr{E}(\mathscr{D}_1 \cup \mathscr{D}_2)$. This means that $f - \delta = h + \lambda_1 f_1 + \lambda_2 f_2$, where $h \in \mathscr{L}^+ \cup \{0\}$, $f_1 \in \mathscr{D}_1$, $f_2 \in \mathscr{D}_2$ and $\lambda_1, \lambda_2 \geq 0$ and not both equal 0. But then

$$f = h + (\lambda_1 f_1 + \delta/2) + (\lambda_2 f_2 + \delta/2).$$

We have $f_1' := \lambda_1 f_1 + \delta/2 \in \mathscr{D}_1$ and $f_2' := \lambda_2 f_2 + \delta/2 \in \mathscr{D}_2$. But this, together with $\lambda_1 f_1 = f_1' - \delta/2 \in \mathscr{D}_1$ if $\lambda_1 > 0$ or otherwise $f_1' \in \mathscr{L}^+$, and $\lambda_2 f_2 = f_2' - \delta/2 \in \mathscr{D}_2$ if $\lambda_2 > 0$ or otherwise $f_2' \in \mathscr{L}^+$, show according to Eq. (1.10), that $f_1' \in \mathscr{D}_1^+$ and $f_2' \in \mathscr{D}_2^+$. So, finally, we have $f \in \mathscr{D}_1^+ \cdot \mathscr{D}_2^+ := \mathscr{C}(\mathscr{D}_1^+ \cup \mathscr{D}_2^+)$. This proves that $(\mathscr{D}_1 \cdot \mathscr{D}_2)^+ = \mathscr{D}_1^+ \cdot \mathscr{D}_2^+$.

2. $\mathscr{D}^+ \subseteq \mathscr{D}$, so that

$$\epsilon_S(\mathscr{D}^+) = \tau^+(\sigma(\epsilon_S(\mathscr{D}^+))) \subseteq \tau^+(\sigma(\epsilon_S(\mathscr{D}))) =: (\epsilon_S(\mathscr{D}))^+.$$

Further

$$(\epsilon_S(\mathscr{D}))^+ := \tau^+(\sigma(\epsilon_S(\mathscr{D}))) := \{f \in \mathscr{L} : \sigma(\epsilon_S(\mathscr{D}))(f) > 0\} \cup \mathscr{L}^+,$$

where

$$\sigma(\epsilon_S(\mathscr{D}))(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathscr{C}(\mathscr{D} \cap \mathscr{L}_S)\},$$

for every $f \in \mathscr{L}$. So, if $f \in (\epsilon_S(\mathscr{D}))^+$, then either $f \in \mathscr{L}^+$ in which case $f \in \epsilon_S(\mathscr{D}^+)$ or there is a $\delta > 0$ so that $f - \delta \in \mathscr{C}(\mathscr{D} \cap \mathscr{L}_S) = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_S)$. In the second case, if $f \notin \mathscr{L}^+$, $f - \delta = h + g$ where $h \in \mathscr{L}^+ \cup \{0\}$ and $g \in \mathscr{D} \cap \mathscr{L}_S$. Then we have $f = h + g'$ where $g' := g + \delta$ is still S- measurable and $g' \in \mathscr{D}$. But, given the fact that $g = g' - \delta \in \mathscr{D} \cap \mathscr{L}_S$, from Eq. (1.10), we have $g' \in \mathscr{D}^+ \cap \mathscr{L}_S$ and therefore $f \in \epsilon_S(\mathscr{D}^+)$. Thus, we conclude that $(\epsilon_S(\mathscr{D}))^+ = \epsilon_S(\mathscr{D}^+)$.

□

Proof of Corollary 3. These claims are immediate consequences of Theorems 10 and Theorem 11. □

We introduce now a preliminary result needed to prove the subsequent Theorem 12.

Corollary 15. *Consider $\mathscr{D} \in \Phi$. If $S \subseteq I$ is a support of $\mathscr{D}$, then it is a support also of $\sigma(\mathscr{D})$. Vice versa, starting from $\underline{P} \in \underline{\Phi}$, if $S \subseteq I$ is a support of $\underline{P}$ then it is also a support of $\tau^+(\underline{P})$.*

Proof. The first result derives directly from item 2 of Corollary 3. Regarding the second one, $\underline{e}_S(\underline{P}) = \underline{P}$ implies, thanks to item 3 of Theorem 10, $\sigma(\epsilon_S(\tau^+(\underline{P}))) = \sigma(\tau^+(\underline{P}))$. Applying then τ^+ to both the terms of the equivalence we have the result. □

For simplicity, we also introduce the following maps.

- $\hat{\sigma} : \hat{\Phi} \to \hat{\underline{\Phi}}$, defined by
$$(\mathscr{D}, S) \mapsto \hat{\sigma}(\mathscr{D}, S) := (\sigma(\mathscr{D}), S);$$
- $\tilde{\sigma} : \tilde{\Phi} \to \tilde{\underline{\Phi}}$, already introduced in Section 3.1, defined by
$$(\tilde{\mathscr{D}}, S) \mapsto \tilde{\sigma}(\tilde{\mathscr{D}}, S) := (\sigma(\tilde{\mathscr{D}}), S).$$

Thus, from Corollary 15 and Eq.(3.6), it follows:

$$\underline{h}(\underline{P}, S) = \underline{h}(\hat{\sigma}(\tau^+(\underline{P}), S)) = \tilde{\sigma}(h(\tau^+(\underline{P}), S)) \tag{C.5}$$

for every $(\underline{P}, S) \in \hat{\underline{\Phi}}$, where $\underline{h}$ is defined in Eq.(3.6) and h is defined in Eq. (3.4).

Proof of Theorem 12. 1. We have, by definition,

$$\underline{h}((\underline{P}_1,S)\cdot(\underline{P}_2,T)) := \underline{h}(\underline{P}_1\cdot\underline{P}_2,S\cup T)$$
$$= \underline{h}(\sigma(\mathscr{D}_1^+)\cdot\sigma(\mathscr{D}_2^+),S\cup T),$$

if we define $\mathscr{D}_1^+ := \tau^+(\underline{P}_1)$ and $\mathscr{D}_2^+ := \tau^+(\underline{P}_2)$. Now, thanks to Theorem 10, we have

$$\underline{h}((\underline{P}_1,S)\cdot(\underline{P}_2,T)) = \underline{h}(\sigma(\mathscr{D}_1^+)\cdot\sigma(\mathscr{D}_2^+),S\cup T)$$
$$= \underline{h}(\sigma(\mathscr{D}_1^+\cdot\mathscr{D}_2^+),S\cup T)$$
$$=: \underline{h}(\hat{\sigma}(\mathscr{D}_1^+\cdot\mathscr{D}_2^+,S\cup T)).$$

Therefore, by Eq. (C.5) and Theorem 9, we have

$$\underline{h}((\underline{P}_1,S)\cdot(\underline{P}_2,T)) = \underline{h}(\hat{\sigma}(\mathscr{D}_1^+\cdot\mathscr{D}_2^+,S\cup T))$$
$$= \tilde{\sigma}(h(\mathscr{D}_1^+\cdot\mathscr{D}_2^+,S\cup T))$$
$$=: \tilde{\sigma}(h((\mathscr{D}_1^+,S)\cdot(\mathscr{D}_2^+,T)))$$
$$= \tilde{\sigma}(h(\mathscr{D}_1^+,S)\cdot h(\mathscr{D}_2^+,T)).$$

Now, we claim that

$$\tilde{\sigma}(h(\mathscr{D}_1^+,S)\cdot h(\mathscr{D}_2^+,T)) = \tilde{\sigma}(h(\mathscr{D}_1^+,S))\cdot\tilde{\sigma}(h(\mathscr{D}_2^+,T)).$$

Indeed, on the one hand, we have

$$\tilde{\sigma}(h(\mathscr{D}_1^+,S)\cdot h(\mathscr{D}_2^+,T))$$
$$:= \tilde{\sigma}(((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S},S)\cdot((\mathscr{D}_2^+\cap\mathscr{L}_T)^{\downarrow T},T))$$
$$:= (\sigma(\mathscr{C}(((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T})\cdot\mathscr{C}(((\mathscr{D}_2^+\cap\mathscr{L}_T)^{\downarrow T})^{\uparrow S\cup T})),S\cup T)$$
$$:= (\sigma(\mathscr{C}(\mathscr{D}_{1,S\cup T}^+)\cdot\mathscr{C}(\mathscr{D}_{2,S\cup T}^+)),S\cup T),$$

where $\mathscr{D}_{1,S\cup T}^+ := ((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T}$ and $\mathscr{D}_{2,S\cup T}^+ := ((\mathscr{D}_2^+\cap\mathscr{L}_T)^{\downarrow T})^{\uparrow S\cup T}$. On the other hand instead, we have

$$\tilde{\sigma}(h(\mathscr{D}_1^+,S))\cdot\tilde{\sigma}(h(\mathscr{D}_2^+,T))$$
$$:= \tilde{\sigma}((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S},S)\cdot\tilde{\sigma}((\mathscr{D}_2^+\cap\mathscr{L}_T)^{\downarrow T},T)$$
$$:= (\sigma((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S}),S)\cdot(\sigma((\mathscr{D}_2^+\cap\mathscr{L}_T)^{\downarrow T}),T)$$
$$:= (\underline{E}^*((\underline{P}_{1,S}^{\downarrow S})^{\uparrow S\cup T})\cdot\underline{E}^*((\underline{P}_{2,T}^{\downarrow T})^{\uparrow S\cup T}),S\cup T).$$

Now, we can show that $(\underline{P}_{1,S}^{\downarrow S})^{\uparrow S\cup T} = \sigma(\mathscr{D}_{1,S\cup T}^+) = \sigma(((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T})$. Indeed,

$$\sigma(((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T})(f) := \sup\{\mu\in\mathbb{R} : f-\mu\in((\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S})^{\uparrow S\cup T}\}$$
$$= \sup\{\mu\in\mathbb{R} : f^{\downarrow S}-\mu\in(\mathscr{D}_1^+\cap\mathscr{L}_S)^{\downarrow S}\} =: (\underline{P}_{1,S}^{\downarrow S})^{\uparrow S\cup T}(f),$$

for every $f\in\mathscr{L}_S(\Omega_{S\cup T})$. Analogously, we can show that $(\underline{P}_{2,T}^{\downarrow T})^{\uparrow S\cup T} =$

$\sigma(\mathscr{D}^+_{2,S\cup T}) = \sigma(((\mathscr{D}^+_2 \cap \mathscr{L}_T)^{\downarrow T})^{\uparrow S\cup T})$. So, we have:

$$\begin{aligned}(\underline{E}^*((\underline{P}^{\downarrow S}_{1,S})^{\uparrow S\cup T}) \cdot \underline{E}^*((\underline{P}^{\downarrow T}_{2,T})^{\uparrow S\cup T}), S\cup T)\\ = (\underline{E}^*(\sigma(\mathscr{D}^+_{1,S\cup T})) \cdot \underline{E}^*(\sigma(\mathscr{D}^+_{2,S\cup T})), S\cup T) =\\ = (\sigma(\mathscr{C}(\mathscr{D}^+_{1,S\cup T})) \cdot \sigma(\mathscr{C}(\mathscr{D}^+_{2,S\cup T})), S\cup T).\end{aligned}$$

In fact, if $\mathscr{D}^+_1 = \mathscr{L}$, then $\underline{E}^*(\sigma(\mathscr{D}^+_{1,S\cup T})) = \sigma(\mathscr{C}(\mathscr{D}^+_{1,S\cup T}))$. Otherwise, $\mathscr{D}^+_{1,S\cup T}$ satisfies the hypotheses of Theorem 20 in Appendix C.1:

- $0 \notin \mathscr{E}(\mathscr{D}^+_{1,S\cup T})$. Otherwise, $0 \in \mathscr{E}((\mathscr{D}^+_1 \cap \mathscr{L}_S)^{\downarrow S\cup T})$ by Lemma 21 in Appendix C.1. Again by Lemma 21 in Appendix C.1, this means $(\mathscr{C}(\mathscr{D}^+_1 \cap \mathscr{L}_S) \cap \mathscr{L}_{S\cup T})^{\downarrow S\cup T} = 0$ that implies $0 \in \mathscr{E}(\mathscr{D}^+_1 \cap \mathscr{L}_S) \subseteq \mathscr{D}^+_1$, which is a contradiction;
- if $f \in \mathscr{D}^+_{1,S\cup T} \setminus \mathscr{L}^+(\Omega_{S\cup T})$, then $f \in (\mathscr{D}^+_1 \cap \mathscr{L}_S)^{\downarrow S\cup T} \setminus \mathscr{L}^+(\Omega_{S\cup T})$ thanks to Lemma 21 in Appendix C.1. Hence $f^{\uparrow I} \in \mathscr{D}^+_1 \cap \mathscr{L}_S \setminus \mathscr{L}^+_S$ and then, there exists $\delta > 0$ such that $f^{\uparrow I} - \delta = (f-\delta)^{\uparrow I} \in \mathscr{D}^+_1 \cap \mathscr{L}_S$ that means $f - \delta \in (\mathscr{D}^+_1 \cap \mathscr{L}_S)^{\downarrow S\cup T} = \mathscr{D}^+_{1,S\cup T}$.

Analogously, we can show that $\underline{E}^*(\sigma(\mathscr{D}^+_{2,S\cup T})) = \sigma(\mathscr{C}(\mathscr{D}^+_{2,S\cup T}))$.

Moreover, given the fact that $\mathscr{C}(\mathscr{D}^+_{1,S\cup T}), \mathscr{C}(\mathscr{D}^+_{1,S\cup T}) \in \Phi^+(\Omega_{S\cup T})$ and thanks again to Theorem 10, we have

$$\begin{aligned}\tilde{\sigma}(h(\mathscr{D}^+_1, S)) \cdot \tilde{\sigma}(h(\mathscr{D}^+_2, T)) = (\sigma(\mathscr{C}(\mathscr{D}^+_{1,S\cup T})) \cdot \sigma(\mathscr{C}(\mathscr{D}^+_{2,S\cup T})), S\cup T) =\\ = (\sigma(\mathscr{C}(\mathscr{D}^+_{1,S\cup T}) \cdot \mathscr{C}(\mathscr{D}^+_{2,S\cup T})), S\cup T) = \tilde{\sigma}(h(\mathscr{D}^+_1, S) \cdot h(\mathscr{D}^+_2, T)).\end{aligned}$$

Hence, we have finally:

$$\begin{aligned}\underline{h}((\underline{P}_1, S) \cdot (\underline{P}_2, T)) &= \tilde{\sigma}(h(\mathscr{D}^+_1, S) \cdot h(\mathscr{D}^+_2, T))\\ &= \tilde{\sigma}(h(\mathscr{D}^+_1, S)) \cdot \tilde{\sigma}(h(\mathscr{D}^+_2, T))\\ &= \underline{h}(\hat{\sigma}(\mathscr{D}^+_1, S)) \cdot \underline{h}(\hat{\sigma}(\mathscr{D}^+_2, T)) =\\ &= \underline{h}(\underline{P}_1, S) \cdot \underline{h}(\underline{P}_2, T).\end{aligned}$$

2. Obviously, $\underline{h}(\sigma(\mathscr{L}), S) = (\sigma(\mathscr{L}(\Omega_S)), S)$.

3. Similarly, $\underline{h}(\sigma(\mathscr{L}^+), S) = (\sigma(\mathscr{L}^+(\Omega_S)), S)$.

4. We have

$$\underline{h}(\underline{\pi}_T(\underline{P}, S)) := \underline{h}(\underline{e}_T(\underline{P}), T) = \underline{h}(\underline{e}_T(\sigma(\mathscr{D}^+)), T),$$

if $\mathscr{D}^+ = \tau^+(\underline{P})$. Therefore using Theorem 9, Theorem 10 and Eq. (C.5), we have

$$\begin{aligned}\underline{h}(\underline{\pi}_T(\underline{P},S)) &:= \underline{h}(\underline{e}_T(\sigma(\mathscr{D}^+)),T)\\ &= \underline{h}(\sigma(\epsilon_T(\mathscr{D}^+)),T)\\ &=: \underline{h}(\hat{\sigma}(\pi_T(\mathscr{D}^+,S)))\\ &= \tilde{\sigma}(h(\pi_T(\mathscr{D}^+,S)))\\ &= \tilde{\sigma}(\pi_T(h(\mathscr{D}^+,S)))\\ &:= \tilde{\sigma}(\pi_T((\mathscr{D}^+\cap\mathscr{L}_S)^{\downarrow S},S))\\ &:= \tilde{\sigma}((\epsilon_T((\mathscr{D}^+\cap\mathscr{L}_S)^{\downarrow S})\cap\mathscr{L}_T(\Omega_S))^{\downarrow T},T))\\ &:= (\sigma(((\mathscr{D}^+\cap\mathscr{L}_S)^{\downarrow S}\cap\mathscr{L}_T(\Omega_S))^{\downarrow T}),T).\end{aligned}$$

At the end we use the fact that $\mathscr{D}\cap\mathscr{L}_T(\Omega_S)\subseteq\epsilon_T(\mathscr{D})\cap\mathscr{L}_T(\Omega_S)\subseteq\mathscr{D}\cap\mathscr{L}_T(\Omega_S)$ for every $\mathscr{D}\in\Phi(\Omega_S)$ and $T\subseteq S$, similarly to what observed in the proof of Theorem 9. Now, we have

$$\begin{aligned}\underline{h}(\underline{\pi}_T(\underline{P},S)) &= (\sigma(((\mathscr{D}^+\cap\mathscr{L}_S)^{\downarrow S}\cap\mathscr{L}_T(\Omega_S))^{\downarrow T}),T)\\ &= ((\underline{P}_S^{\downarrow S})_T^{\downarrow T},T)\\ &= \underline{\pi}_T(\underline{h}(\underline{P},S)).\end{aligned}$$

Indeed, by a reasoning similar to the one underlying Eq.(3.6), we can show that $\sigma((\tau^+(\underline{P}_S^{\downarrow S})\cap\mathscr{L}_T)^{\downarrow T})=(\underline{P}_S^{\downarrow S})_T^{\downarrow T}$. Moreover, we have $\underline{e}_T(\underline{P}_S^{\downarrow S})_T^{\downarrow T}:=\underline{E}^*((\underline{P}_S^{\downarrow S})_T)_T^{\downarrow T}=(\underline{P}_S^{\downarrow S})_T^{\downarrow T}$. So we have the result.

5. Suppose $\underline{h}(\underline{P}_1,S)=\underline{h}(\underline{P}_2,T)$. Then we have $S=T$ and $\underline{P}_{1,S}^{\downarrow S}=\underline{P}_{2,S}^{\downarrow S}$, from which we derive that $\underline{P}_{1,S}=\underline{P}_{2,S}$ and therefore, $\underline{P}_1=\underline{e}_S(\underline{P}_1)=\underline{E}^*(\underline{P}_{1,S})=\underline{E}^*(\underline{P}_{2,S})=\underline{e}_S(\underline{P}_2)=\underline{P}_2$. So the map $\underline{h}$ is injective.

   Moreover, for any $(\tilde{\underline{P}},S)\in\tilde{\underline{\Phi}}$, if we call $\mathscr{D}^+:=\tau^+(\underline{P})$, we claim that $(\tilde{\underline{P}},S)=\underline{h}(\hat{\sigma}(\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I}),S))=\underline{h}(\underline{E}^*(\tilde{\underline{P}}^{\uparrow I}),S)$. Since $(\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I}),S)\in\hat{\Phi}$, see the proof of item 2 of Theorem 9, we have that $(\underline{E}^*(\tilde{\underline{P}}^{\uparrow I}),S)\in\hat{\underline{\Phi}}$ and hence $\underline{h}$ is surjective.

   Indeed, $h((\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I}),S))=(\tilde{\mathscr{D}}^+,S)$ again from the proof of item 2 of Theorem 9. Therefore:

$$\begin{aligned}\underline{h}(\hat{\sigma}(\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I}),S)) &= \tilde{\sigma}(h(\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I}),S)))\\ &= \tilde{\sigma}(\tilde{\mathscr{D}}^+,S)=(\sigma(\tilde{\mathscr{D}}^+),S)=(\tilde{\underline{P}},S).\end{aligned}$$

   Moreover, we have $\sigma((\tilde{\mathscr{D}}^+)^{\uparrow I})=\tilde{\underline{P}}^{\uparrow I}$. This follows from:

$$\begin{aligned}\sigma((\tilde{\mathscr{D}}^+)^{\uparrow I})(f) &:= \sup\{\mu\in\mathbb{R}: f-\mu\in(\tilde{\mathscr{D}}^+)^{\uparrow I}\} =\\ &\sup\{\mu\in\mathbb{R}: f^{\downarrow S}-\mu\in\tilde{\mathscr{D}}^+\} =: \tilde{\underline{P}}^{\uparrow I}(f)\end{aligned}$$

   for every $f\in\mathscr{L}_S(\Omega_I)$.

Finally, if $\underline{P} = \sigma(\mathscr{L}(\Omega_S))$, we already have $\sigma(\mathscr{C}((\tilde{\mathscr{D}}^+)^{\uparrow I})) = \underline{E}^*(\sigma((\tilde{\mathscr{D}}^+)^{\uparrow I})) = \underline{E}^*(\underline{\tilde{P}}^{\uparrow I})$, otherwise, to obtain this equivalence, we use Theorem 20 in Appendix C.1 (that can be applied on $(\tilde{\mathscr{D}}^+)^{\uparrow I}$).

So h is surjective, hence bijective.

□

Proof of Lemma 7. Clearly $\underline{P}(f) = +\infty$ for all $f \in \mathscr{L}$ is a possible solution. Consider instead the case in which $\underline{P}$ is coherent.

From [Walley, 1991, Section 2.6.1], we know that $\underline{P}(f) \leq \overline{P}(f)$, for all $f \in \mathscr{L}$. Then, we have:

$$\underline{P}(f) \leq \overline{P}(f) := -\underline{P}(-f) \leq -P(-f) = P(f),\ \forall f \in \mathscr{L}. \tag{C.6}$$

Given the fact that, by hypothesis, we have also $\underline{P}(f) \geq P(f)$ for all $f \in \mathscr{L}$, we have the result. □

Proof of Lemma 8. Let us suppose that $\mathscr{D}_1, \ldots, \mathscr{D}_n$ with supports $S_1, \ldots, S_n$ respectively are compatible. Then, there exists a coherent set of gambles $\mathscr{D}$ such that $\epsilon_{S_i}(\mathscr{D}) = \mathscr{D}_i$, for every i. By Idempotency axiom, we know that $\mathscr{D} \geq \epsilon_{S_i}(\mathscr{D}) = \mathscr{D}_i$ for every i. From this fact, we can prove recursively that $\mathscr{D} \geq \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$. Indeed, if $n = 1$ we already have the result. Otherwise, from $\mathscr{D} \geq \mathscr{D}_1$, we derive $\mathscr{D} = \mathscr{D} \cdot \mathscr{D}_2 \geq \mathscr{D}_1 \cdot \mathscr{D}_2$, by item 3 of Lemma 17 in Appendix C. This can be then repeated for every $\mathscr{D}_i$ with $i \leq n$, hence $\mathscr{D} \geq \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$. Thus, $\mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n \neq \mathscr{L}$. □

Proof of Lemma 9. Let us suppose that $\mathscr{D}_1, \ldots, \mathscr{D}_n$ with supports $S_1, \ldots, S_n$ respectively are compatible. Then, there exists a coherent set of gambles $\mathscr{D}$ such that $\epsilon_{S_i}(\mathscr{D}) = \mathscr{D}_i$, for every i. For every $i, j \in \{1, \ldots, n\}$, we have therefore $\epsilon_{S_i \cap S_j}(\mathscr{D}_i) = \epsilon_{S_i \cap S_j}(\epsilon_{S_i}(\mathscr{D})) = \epsilon_{S_i \cap S_j}(\mathscr{D}) = \epsilon_{S_i \cap S_j}(\epsilon_{S_j}(\mathscr{D})) = \epsilon_{S_i \cap S_j}(\mathscr{D}_j)$, by Transitivity axiom. Hence, $\mathscr{D}_1, \ldots, \mathscr{D}_n$ are also pairwise compatible. □

Proof of Theorem 14. Let $Y_i := S_{i+1} \cup \ldots \cup S_n$ for $i = 1, \ldots, n-1$ and $\mathscr{D} := \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$. Then by RIP, Combination and Transitivity axiom

$$\begin{aligned}\epsilon_{Y_1}(\mathscr{D}) &= \epsilon_{Y_1}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n = \epsilon_{Y_1}(\epsilon_{S_1}(\mathscr{D}_1)) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n \\ &= \epsilon_{S_1 \cap Y_1}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n = \epsilon_{S_1 \cap S_{p(1)}}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n.\end{aligned}$$

But by pairwise compatibility $\epsilon_{S_1 \cap S_{p(1)}}(\mathscr{D}_1) = \epsilon_{S_1 \cap S_{p(1)}}(\mathscr{D}_{p(1)})$, hence by Idempotency axiom

$$\epsilon_{Y_1}(\mathscr{D}) = \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n.$$

By induction on i, one shows exactly in the same way that

$$\epsilon_{Y_i}(\mathscr{D}) = \mathscr{D}_{i+1} \cdot \ldots \cdot \mathscr{D}_n,\ \forall i = 1, .., n-1.$$

So, we obtain $\epsilon_{Y_{n-1}}(\mathscr{D}) := \epsilon_{S_n}(\mathscr{D}) = \mathscr{D}_n$.

Now, we claim that $\epsilon_{S_i}(\mathscr{D}) = \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}) \cdot \mathscr{D}_i$ for every $i = 1, \ldots, n-1$. Indeed, we have by RIP, Transitivity and Combination axiom:

$$\mathscr{D}_i \cdot \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}) = \mathscr{D}_i \cdot \epsilon_{S_i \cap Y_i}(\mathscr{D}) = \mathscr{D}_i \cdot \epsilon_{S_i}(\epsilon_{Y_i}(\mathscr{D})) =$$
$$\mathscr{D}_i \cdot \epsilon_{S_i}(\mathscr{D}_{i+1} \cdot \ldots \cdot \mathscr{D}_n) = \epsilon_{S_i}(\mathscr{D}_i \cdot \mathscr{D}_{i+1} \cdot \ldots \cdot \mathscr{D}_n).$$

Now, if $i = 1$ we have the result, otherwise we have by Transitivity axiom

$$\mathscr{D}_i \cdot \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}) = \epsilon_{S_i}(\mathscr{D}_i \cdot \mathscr{D}_{i+1} \cdot \ldots \cdot \mathscr{D}_n) = \epsilon_{S_i}(\epsilon_{Y_{i-1}}(\mathscr{D})) = \epsilon_{S_i}(\mathscr{D}).$$

Then, by backward induction, based on the induction assumption $\epsilon_{S_j}(\mathscr{D}) = \mathscr{D}_j$ for $j > i$, and rooted in $\epsilon_{S_n}(\mathscr{D}) = \mathscr{D}_n$, for $i = n-1, \ldots 1$, we have by pairwise compatibility, Transitivity and Idempotency axiom

$$\epsilon_{S_i}(\mathscr{D}) = \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}) \cdot \mathscr{D}_i = \epsilon_{S_i \cap S_{p(i)}}(\epsilon_{S_{p(i)}}(\mathscr{D})) \cdot \mathscr{D}_i$$
$$= \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}_{p(i)}) \cdot \mathscr{D}_i = \epsilon_{S_i \cap S_{p(i)}}(\mathscr{D}_i) \cdot \mathscr{D}_i = \mathscr{D}_i.$$

This concludes the proof. □

C.2 Proofs of Section 3.2

We start this section by proving the following lemma that is a direct consequence of the axioms defining a generalised domain-free information algebra.

Lemma 24. *Consider a generalised domain-free information algebra* $(\boldsymbol{\Phi}, \mathbf{Q}; \vee, \perp, \cdot, \mathbf{0}, \mathbf{1}, \boldsymbol{\epsilon})$. *For any* $\boldsymbol{\phi}, \boldsymbol{\psi} \in \boldsymbol{\Phi}$ *and* $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{Q}$*, we have*

1. $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) = \mathbf{0} \iff \boldsymbol{\phi} = \mathbf{0}$,
2. $\boldsymbol{x}$ *is a support of* $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi})$,
3. $\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \boldsymbol{\epsilon}_{\boldsymbol{y}}(\boldsymbol{\phi}) = \boldsymbol{\epsilon}_{\boldsymbol{y}}(\boldsymbol{\phi})$,
4. $\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \boldsymbol{\epsilon}_{\boldsymbol{y}}(\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi})) = \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi})$,
5. $\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \epsilon_{\boldsymbol{x}}(\epsilon_{\boldsymbol{y}}(\boldsymbol{\phi})) = \epsilon_{\boldsymbol{x}}(\boldsymbol{\phi})$,
6. $\boldsymbol{x}$ *support of* $\boldsymbol{\phi}$ *and* $\boldsymbol{\psi} \Rightarrow \boldsymbol{x}$ *support of* $\boldsymbol{\phi} \cdot \boldsymbol{\psi}$,
7. $\boldsymbol{x}$ *support of* $\boldsymbol{\phi}$ *and* $\boldsymbol{y}$ *support of* $\boldsymbol{\psi} \Rightarrow \boldsymbol{x} \vee \boldsymbol{y}$ *support of* $\boldsymbol{\phi} \cdot \boldsymbol{\psi}$ *and* $\boldsymbol{\phi} \cdot \boldsymbol{\psi} = \boldsymbol{\epsilon}_{\boldsymbol{x} \vee \boldsymbol{y}}(\boldsymbol{\phi}) \cdot \boldsymbol{\epsilon}_{\boldsymbol{x} \vee \boldsymbol{y}}(\boldsymbol{\psi})$.

Proof.
1. $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\mathbf{0}) = \mathbf{0}$ by Nullity axiom. If instead $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) = \mathbf{0}$, then $\boldsymbol{\phi} = \boldsymbol{\phi} \cdot \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) = \boldsymbol{\phi} \cdot \mathbf{0} = \mathbf{0}$ by Idempotency axiom.
2. $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi})) = \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \mathbf{1}) = \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi}) \cdot \boldsymbol{\epsilon}_{\boldsymbol{x}}(\mathbf{1}) = \boldsymbol{\epsilon}_{\boldsymbol{x}}(\boldsymbol{\phi})$ by Combination axiom and $\boldsymbol{\epsilon}_{\boldsymbol{x}}(\mathbf{1}) = \mathbf{1}$.

3. By item 2 of this lemma, we know that $\epsilon_x(\epsilon_x(\phi)) = \epsilon_x(\phi)$. Hence, by Support axiom, we have also $\epsilon_y(\epsilon_x(\phi)) = \epsilon_x(\phi)$. Therefore, $\epsilon_y(\phi) = \epsilon_y(\phi \cdot \epsilon_x(\phi)) = \epsilon_y(\phi \cdot \epsilon_y(\epsilon_x(\phi))) = \epsilon_y(\phi) \cdot \epsilon_x(\phi)$ by Idempotency and Combination axiom.

4. It directly follows from Support axiom and item 2 of this lemma, as previously noticed.

5. By item 3 of this lemma, we have
$$\epsilon_x(\phi) \cdot \epsilon_y(\phi) = \epsilon_y(\phi).$$
So, applying ϵ_x operator to both sides of the equation, we have
$$\epsilon_x(\epsilon_x(\phi) \cdot \epsilon_y(\phi)) = \epsilon_x(\epsilon_y(\phi)).$$
However, by Idempotency and Combination axiom, we have
$$\epsilon_x(\epsilon_x(\phi) \cdot \epsilon_y(\phi)) = \epsilon_x(\phi) \cdot \epsilon_x(\epsilon_y(\phi)) = \epsilon_x(\phi \cdot \epsilon_x(\epsilon_y(\phi)) = \epsilon_x(\phi \cdot \epsilon_y(\phi) \cdot \epsilon_x(\epsilon_y(\phi)) = \epsilon_x(\phi \cdot \epsilon_y(\phi)) = \epsilon_x(\phi).$$

6. $\epsilon_x(\phi \cdot \psi) = \epsilon_x(\phi \cdot \epsilon_x(\psi)) = \epsilon_x(\phi) \cdot \epsilon_x(\psi) = \phi \cdot \psi$, by Combination axiom.

7. From Support axiom, we know $x \vee y$ is a support of both ϕ and ψ. Hence, by item 6 of this lemma we have the result.

□

We show now the following theorem, which proves that Definition 46 coincides with the definition of generalised domain-free information algebra given in Section 5.2 of Kohlas [2017].[1] In Kohlas [2017] indeed, a definition of generalised domain-free information algebra is given where, in place of the Combination axiom, we have the following:

- *Original Combination:* let us consider $\phi, \psi \in \Phi$ and $x, y, z \in Q$. If x is a support of ϕ, y of ψ and $x \perp y | z$,
$$\epsilon_z(\phi \cdot \psi) = \epsilon_z(\phi) \cdot \epsilon_z(\psi),$$

and in place of the Extraction axiom, we have:

- *Original Extraction:* let us consider $\phi \in \Phi$ and $x, y, z \in Q$. If x is a support of ϕ and $x \perp y | z$.
$$\epsilon_y(\phi) = \epsilon_y(\epsilon_z(\phi)).$$

[1] In Kohlas [2017] it is also required that $\epsilon_x(\phi) = 0$ if and only if $\phi = 0$, in place of Nullity axiom. However, this is satisfied also by a generalised domain-free information algebra as defined in this work, as Lemma 24 in Appendix C.2 proves.

Theorem 21. *Consider a generalised domain-free information algebra* $(\Phi, Q; \vee, \perp, \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$, *as defined in Definition 46.*

1. *For any* $\phi, \psi \in \Phi$ *and* $x, y, z \in Q$ *such that* $\epsilon_x(\phi) = \phi$, $\epsilon_y(\psi) = \psi$ *and* $x \perp y | z$, *we have:*
$$\epsilon_z(\phi \cdot \psi) = \epsilon_z(\phi) \cdot \epsilon_z(\psi).$$

2. *For any* $\phi \in \Phi$ *and* $x, y, z \in Q$ *such that* $\epsilon_x(\phi) = \phi$ *and* $x \perp y | z$, *we have:*
$$\epsilon_y(\phi) = \epsilon_y(\epsilon_z(\phi)).$$

Vice versa, if the structure $(\Phi, Q; \vee, \perp, \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$ *satisfies all the axioms defining a generalised domain-free information algebra as precised by Definition 46, except for the Combination and the Extraction axiom, replaced by item 1, i.e., Original Combination axiom, and item 2, i.e., Original Extraction axiom, of this theorem respectively, then it is still a generalised domain-free information algebra.*

Proof. Let us start by assuming that $(\Phi, Q; \vee, \perp, \cdot, \mathbf{0}, \mathbf{1}, \epsilon)$ is a generalised domain-free information algebra in the sense specified by Definition 46, and let us prove items 1 and 2. First of all, notice that from $x \perp y | z$ it follows by the properties of a quasi-separoid that $x \vee z \perp y \vee z | z$.

1. Let us consider $\phi, \psi \in \Phi$ and $x, y, z \in Q$ as in item 1. By Combination, Extraction and Support axiom, we have
$$\epsilon_{y \vee z}(\phi \cdot \psi) = \epsilon_{y \vee z}(\phi \cdot \epsilon_{y \vee z}(\psi)) = \epsilon_{y \vee z}(\phi) \cdot \psi = \epsilon_{y \vee z}(\epsilon_z(\phi)) \cdot \psi.$$
By Lemma 24 in in Appendix C.2, the last combination equals $\epsilon_z(\phi) \cdot \psi$. But then, again by Combination axiom and Lemma 24 in Appendix C.2, we have
$$\epsilon_z(\phi \cdot \psi) = \epsilon_z(\epsilon_{y \vee z}(\phi \cdot \psi)) =$$
$$\epsilon_z(\epsilon_z(\phi) \cdot \psi) = \epsilon_z(\phi) \cdot \epsilon_z(\psi).$$

2. Now, let us consider $\phi \in \Phi$ and $x, y, z \in Q$ as in item 2. By Extraction axiom, we have:
$$\epsilon_{y \vee z}(\phi) = \epsilon_{y \vee z}(\epsilon_z(\phi)).$$
By Lemma 24 in Appendix C.2, applying ϵ_y to both sides, we have $\epsilon_y(\epsilon_{y \vee z}(\phi)) = \epsilon_y(\phi)$ and $\epsilon_y(\epsilon_{y \vee z}(\epsilon_z(\phi)) = \epsilon_y(\epsilon_z(\phi))$. Hence $\epsilon_y(\phi) = \epsilon_y(\epsilon_z(\phi))$.

This concludes the first part of the proof. Vice versa, suppose item 1 and item 2 of this theorem are satisfied in place of the Extraction and the Combination axiom respectively.

- Combination axiom follows from [Kohlas, 2017, Lemma 5.2 (4)].
- Let us consider $\boldsymbol{\phi} \in \boldsymbol{\Phi}$ and $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{Q}$ such that $\boldsymbol{\epsilon}_x(\boldsymbol{\phi}) = \boldsymbol{\phi}$ and $\boldsymbol{x} \vee \boldsymbol{z} \perp \boldsymbol{y} \vee \boldsymbol{z} | \boldsymbol{z}$. Then, by properties of a quasi-separoid, we have also $\boldsymbol{x} \perp \boldsymbol{y} \vee \boldsymbol{z} | \boldsymbol{z}$. Hence, by item 1, we have
$$\boldsymbol{\epsilon}_{y \vee z}(\boldsymbol{\phi}) = \boldsymbol{\epsilon}_{y \vee z}(\boldsymbol{\epsilon}_z(\boldsymbol{\phi})),$$
thus the Extraction axiom is satisfied.

This concludes the proof. □

As seen in Section 3.2, in a generalised domain-free information algebra, a definition of partial order among pieces of information can be given analogous to the one established for domain-free information algebras. Analogous properties are also satisfied.

Lemma 25. *Consider a generalised domain-free information algebra* $(\boldsymbol{\Phi}, \boldsymbol{Q}; \vee, \perp, \cdot, \boldsymbol{0}, \boldsymbol{1}, \boldsymbol{\epsilon})$. *The following properties are valid. For any* $\boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\mu} \in \boldsymbol{\Phi}$ *and* $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{Q}$, *we have*

1. $\boldsymbol{1} \leq \boldsymbol{\phi} \leq \boldsymbol{0}$;
2. $\boldsymbol{\phi}, \boldsymbol{\psi} \leq \boldsymbol{\phi} \cdot \boldsymbol{\psi}$;
3. $\boldsymbol{\phi} \leq \boldsymbol{\psi} \Rightarrow \boldsymbol{\phi} \cdot \boldsymbol{\mu} \leq \boldsymbol{\psi} \cdot \boldsymbol{\mu}$;
4. $\boldsymbol{\epsilon}_x(\boldsymbol{\phi}) \leq \boldsymbol{\phi}$;
5. $\boldsymbol{\phi} \leq \boldsymbol{\psi} \Rightarrow \boldsymbol{\epsilon}_x(\boldsymbol{\phi}) \leq \boldsymbol{\epsilon}_x(\boldsymbol{\psi})$;
6. $\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \boldsymbol{\epsilon}_x(\boldsymbol{\phi}) \leq \boldsymbol{\epsilon}_y(\boldsymbol{\phi})$.

Proof. They directly follow from the definition of information order and axioms defining a generalised domain-free information algebra. □

Proof of Lemma 10. Items 1,2,4,5 are obvious. For item 6, consider $\omega \in \sigma_x(\sigma_x(S) \cap \sigma_x(T))$. Then there is a $\omega' \in \sigma_x(S) \cap \sigma_x(T)$ so that $\omega \equiv_x \omega'$. In particular, $\omega' \in \sigma_x(S)$, hence $\omega \in \sigma_x(\sigma_x(S)) = \sigma_x(S)$, by item 4. At the same time, $\omega' \in \sigma_x(T)$, hence $\omega \in \sigma_x(\sigma_x(T)) = \sigma_x(T)$. Then $\omega \in \sigma_x(S) \cap \sigma_x(T)$. By item 2 we must then have the equality.

Regarding item 3, $\sigma_x(\sigma_x(S) \cap T) \subseteq \sigma_x(S) \cap \sigma_x(T)$ by item 2, 5 and 6. So consider an element $\omega \in \sigma_x(S) \cap \sigma_x(T)$. Then, there exist $\omega' \in S$ and $\omega'' \in T$ such that $\omega \equiv_x \omega'$ and $\omega \equiv_x \omega''$. By transitivity, it follows that $\omega'' \equiv_x \omega'$ so that $\omega'' \in \sigma_x(S)$. But then $\omega \equiv_x \omega'' \in \sigma_x(S) \cap T$ implies $\omega \in \sigma_x(\sigma_x(S) \cap T)$ and this proves item 3. □

Proof of Theorem 15. From $\sigma_z(\sigma_x(S)) \supseteq \sigma_x(S)$ (Lemma 10) we obtain

$$\sigma_{y\vee z}(\sigma_z(\sigma_x(S)) \supseteq \sigma_{y\vee z}(\sigma_x(S)),$$

again by Lemma 10. Consider therefore an element $\omega \in \sigma_{y\vee z}(\sigma_z(\sigma_x(S)))$. Then, there are elements ω'', ω''' and ω' so that $\omega \equiv_{y\vee z} \omega'' \equiv_z \omega''' \equiv_x \omega'$ and $\omega' \in S$. This means that ω, ω'' belong to some block $B_{y\vee z}$ of partition $\mathscr{P}_{y\vee z}$, ω'', ω''' to some block B_z of partition $\mathscr{P}_z$ and ω''', ω' to some block B_x of partition $\mathscr{P}_x$. It follows that $B_x \cap B_z \neq \emptyset$ and $B_{y\vee z} \cap B_z \neq \emptyset$. Then, $x \vee z \perp y \vee z | z$ implies, thanks to properties of a quasi-separoid, that $x \perp y \vee z | z$. Therefore, we have $B_x \cap B_{y\vee z} \cap B_z \neq \emptyset$, and in particular, $B_x \cap B_{y\vee z} \neq \emptyset$. So there is a $v \in B_x \cap B_{y\vee z}$ such that $\omega \equiv_{y\vee z} v \equiv_x \omega' \in S$, hence $\omega \in \sigma_{y\vee z}(\sigma_x(S))$. So, we have $\sigma_{y\vee z}(\sigma_x(S)) = \sigma_{y\vee z}(\sigma_z(\sigma_x(S)))$. □

Proof of Theorem 16.

- Items 1–4 can be proven analogously to the correspondent results in Theorem 8.

- If $\mathscr{D} = \mathscr{L}$ this is obvious. So, assume $\mathscr{D} \neq \mathscr{L}$. Let

  $$A := \epsilon_{y\vee z}(\mathscr{D}) := \mathscr{C}(\mathscr{D} \cap \mathscr{L}_{y\vee z}) = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_{y\vee z}),$$
  $$B := \epsilon_{y\vee z}(\epsilon_z(\mathscr{D})) := \mathscr{C}(\mathscr{C}(\mathscr{D} \cap \mathscr{L}_z) \cap \mathscr{L}_{y\vee z}) = \mathscr{E}(\mathscr{E}(\mathscr{D} \cap \mathscr{L}_z) \cap \mathscr{L}_{y\vee z}).$$

  Then $\mathscr{D} \cap \mathscr{L}_z \subseteq \mathscr{D}$ implies $B \subseteq A$. Therefore, consider a gamble $f \in A$ so that $f \geq f'$ for a gamble $f' \in \mathscr{D} \cap \mathscr{L}_{y\vee z}$. Then, if $f' \in \mathscr{L}^+ \cap \mathscr{L}_{y\vee z}$, $f \in B$. Otherwise, since $\mathscr{D} = \mathscr{C}(\mathscr{D} \cap \mathscr{L}_x) = \mathscr{E}(\mathscr{D} \cap \mathscr{L}_x)$, we have

  $$f' \geq g, \quad g \in \mathscr{D} \cap \mathscr{L}_x, \quad f' \text{ is } y \vee z-\text{measurable}.$$

  Define for all $\omega \in \Omega$,

  $$g'(\omega) := \sup_{\omega' \equiv_{y\vee z} \omega} g(\omega').$$

  Since f' is $y \vee z$-measurable, we have $f' \geq g'$, and also $g' \in \mathscr{D}$. We claim that g' is z-measurable. Indeed, consider a pair of elements $\omega \equiv_z \omega''$ and the block B_z of partition $\mathscr{P}_z$ that contains these two elements. Then consider the blocks $B_{y\vee z} \subseteq B_z$ and $B''_{y\vee z} \subseteq B_z$ that contain elements ω and ω'' respectively. Finally consider the family of all blocks $B_{x\vee z} \subseteq B_z$. From $x \vee z \perp y \vee z | z$ we conclude that $B_{y\vee z} \cap B_{x\vee z} \neq \emptyset$ and $B''_{y\vee z} \cap B_{x\vee z} \neq \emptyset$ for all blocks $B_{x\vee z} \subseteq B_z$. Since g is also $x \vee z$-measurable, g is constant on any of these blocks. Define $g(B_{x\vee z}) = g(\omega)$ if $\omega \in B_{x\vee z}$. Then it follows that

  $$g'(\omega) = \sup_{B_{x\vee z} \cap B_{y\vee z} \neq \emptyset} g(B_{x\vee z}) = \sup_{B_{x\vee z} \subseteq B_z} g(B_{x\vee z}),$$

  and

  $$g'(\omega'') = \sup_{B_{x\vee z} \cap B''_{y\vee z} \neq \emptyset} g(B_{x\vee z}) = \sup_{B_{x\vee z} \subseteq B_z} g(B_{x\vee z}).$$

This shows that g' is z-measurable, hence $g' \in \mathscr{D} \cap \mathscr{L}_z$. So we conclude that $f' \in \mathscr{C}(\mathscr{D} \cap \mathscr{L}_z) \cap \mathscr{L}_{y \vee z}$. But this implies that $f \in B$ and so $A = B$.

- Assume that $\mathscr{D} = \mathscr{C}(\mathscr{D} \cap \mathscr{L}_x)$. If $x \leq y$, then $\mathscr{L}_x \subseteq \mathscr{L}_y$. Hence $\mathscr{D} \cap \mathscr{L}_x \subseteq \mathscr{D} \cap \mathscr{L}_y$. It follows that $\mathscr{D} = \mathscr{C}(\mathscr{D} \cap \mathscr{L}_x) \subseteq \mathscr{C}(\mathscr{D} \cap \mathscr{L}_y)$. But $\mathscr{C}(\mathscr{D} \cap \mathscr{L}_y) \subseteq \mathscr{D}$, hence $\mathscr{C}(\mathscr{D} \cap \mathscr{L}_y) = \mathscr{D}$ and so $\epsilon_y(\mathscr{D}) = \mathscr{D}$.

□

Theorem 22. *Let $\mathscr{D}_1^+$, $\mathscr{D}_2^+, \mathscr{D}^+ \subseteq \mathscr{L}$ be coherent sets of strictly desirable gambles and $x \in Q$. Then*

1. $(\forall f \in \mathscr{L})\ \sigma(\mathscr{L})(f) = +\infty$, $\sigma(\mathscr{L}^+)(f) = \inf f$,
2. $\sigma(\mathscr{D}_1^+ \cdot \mathscr{D}_2^+) = \sigma(\mathscr{D}_1^+) \cdot \sigma(\mathscr{D}_2^+)$,
3. $\sigma(\epsilon_x(\mathscr{D}^+)) = \underline{e}_x(\sigma(\mathscr{D}^+))$.

Proof. This result can be proven analogously to Theorem 10. □

Proof of Theorem 17. 1. Note that $\mathscr{D}_S^+ = \mathscr{L}^+$ or $\mathscr{D}_T^+ = \mathscr{L}^+$ if and only if $S = \Omega$ or $T = \Omega$. Clearly in this case we have immediately the result. The same is true if $\mathscr{D}_S^+ = \mathscr{L}$ or $\mathscr{D}_T^+ = \mathscr{L}$, which is equivalent to having $S = \emptyset$ or $T = \emptyset$. Now suppose $\mathscr{D}_S^+, \mathscr{D}_T^+ \neq \mathscr{L}^+$ and $\mathscr{D}_S^+, \mathscr{D}_T^+ \neq \mathscr{L}$. If $S \cap T = \emptyset$, then $\mathscr{D}_{S\cap T}^+ = \mathscr{L}$. Consider $f \in \mathscr{D}_S^+ \setminus \mathscr{L}^+$ and $g \in \mathscr{D}_T^+ \setminus \mathscr{L}^+$. Since S and T are disjoint, we have $\tilde{f} \in \mathscr{D}_S^+$ and $\tilde{g} \in \mathscr{D}_T^+$, where $\tilde{f}$, $\tilde{g}$ are defined in the following way for every $\omega \in \Omega$:

$$\tilde{f}(\omega) := \begin{cases} f(\omega) & \text{for } \omega \in S, \\ -g(\omega) & \text{for } \omega \in T, \\ 0 & \text{for } \omega \in (S \cup T)^c, \end{cases} \qquad \tilde{g}(\omega) := \begin{cases} -f(\omega) & \text{for } \omega \in S, \\ g(\omega) & \text{for } \omega \in T, \\ 0 & \text{for } \omega \in (S \cup T)^c. \end{cases}$$

However, $\tilde{f} + \tilde{g} = 0 \in \mathscr{E}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+)$, hence $\mathscr{D}_S^+ \cdot \mathscr{D}_T^+ := \mathscr{C}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+) = \mathscr{L}(\Omega) = \mathscr{D}_{S\cap T}^+$. Assume then that $S \cap T \neq \emptyset$. If $S \cap T = \Omega$, we already have the result. So suppose $S \cap T \neq \emptyset, \Omega$. Note that $\mathscr{D}_S^+ \cup \mathscr{D}_T^+ \subseteq \mathscr{E}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+) \subseteq \mathscr{D}_{S\cap T}^+$ so that $\mathscr{E}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+)$ is coherent and $\mathscr{D}_S^+ \cdot \mathscr{D}_T^+ = \mathscr{E}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+) \subseteq \mathscr{D}_{S\cap T}^+$. Consider then a gamble $f \in \mathscr{D}_{S\cap T}^+$. If $f \in \mathscr{L}^+$, clearly $f \in \mathscr{D}_S^+ \cdot \mathscr{D}_T^+$. Vice versa, suppose $f \in \mathscr{D}_{S\cap T}^+ \setminus \mathscr{L}^+$. Select a $\delta > 0$ and define the two gambles

$$f_1(\omega) := \begin{cases} 1/2 f(\omega) & \text{for } \omega \in (S \cap T), \\ \delta & \text{for } \omega \in S \setminus T, \\ f(\omega) - \delta & \text{for } \omega \in T \setminus S, \\ 1/2 f(\omega) & \text{for } \omega \in (S \cup T)^c, \end{cases} \qquad f_2(\omega) := \begin{cases} 1/2 f(\omega) & \text{for } \omega \in (S \cap T), \\ f(\omega) - \delta & \text{for } \omega \in S \setminus T, \\ \delta & \text{for } \omega \in T \setminus S, \\ 1/2 f(\omega) & \text{for } \omega \in (S \cup T)^c, \end{cases}$$

for every $\omega \in \Omega$. Then $f = f_1 + f_2$ and $f_1 \in \mathscr{D}_S^+$, $f_2 \in \mathscr{D}_T^+$. Therefore $f \in \mathscr{E}(\mathscr{D}_S^+ \cup \mathscr{D}_T^+) = \mathscr{D}_S^+ \cdot \mathscr{D}_T^+$. Hence $\mathscr{D}_S^+ \cdot \mathscr{D}_T^+ = \mathscr{D}_{S\cap T}^+$.

2. Both have been noted above.

3. If S is empty, then $\epsilon_x(\mathscr{D}^+_\emptyset) = \mathscr{L}(\Omega)$ so that item 3 holds in this case. Hence, assume $S \neq \emptyset$. Then we have that $\mathscr{D}^+_S$ is coherent, and therefore:

$$\epsilon_x(\mathscr{D}^+_S) := \mathscr{C}(\mathscr{D}^+_S \cap \mathscr{L}_x) = \mathscr{E}(\mathscr{D}^+_S \cap \mathscr{L}_x) := \operatorname{posi}(\mathscr{L}^+(\Omega) \cup (\mathscr{D}^+_S \cap \mathscr{L}_x)).$$

Consider a gamble $f \in \mathscr{D}^+_S \cap \mathscr{L}_x$. If $f \in \mathscr{L}^+(\Omega) \cap \mathscr{L}_x$ then clearly $f \in \mathscr{D}^+_{\sigma_x(S)}$. Otherwise, $\inf_S f > 0$ and f is x-measurable. If $\omega \equiv_x \omega'$ for some $\omega' \in S$ and $\omega \in \Omega$, then $f(\omega) = f(\omega')$. Therefore $\inf_{\sigma_x(S)} f = \inf_S f > 0$, hence $f \in \mathscr{D}^+_{\sigma_x(S)}$. Then $\mathscr{C}(\mathscr{D}^+_S \cap \mathscr{L}_x) \subseteq \mathscr{C}(\mathscr{D}^+_{\sigma_x(S)}) = \mathscr{D}^+_{\sigma_x(S)}$. Conversely, consider a gamble $f \in \mathscr{D}^+_{\sigma_x(S)}$. $\mathscr{D}^+_{\sigma_x(S)}$ is a coherent set of strictly desirable gambles. Hence, if $f \in \mathscr{D}^+_{\sigma_x(S)}$, $f \in \mathscr{L}^+(\Omega)$ or there is $\delta > 0$ such that $f - \delta \in \mathscr{D}^+_{\sigma_x(S)} \setminus \mathscr{L}^+(\Omega)$. If $f \in \mathscr{L}^+(\Omega)$, then $f \in \epsilon_x(\mathscr{D}^+_S)$. Otherwise, let us define for every $\omega \in \Omega$, $g(\omega) := \inf_{\omega' \equiv_x \omega} f(\omega') - \delta$. If $\omega \in S$, then $g(\omega) > 0$ since $\inf_{\sigma_x(S)}(f - \delta) > 0$. So, we have $\inf_S g \geq 0$ and g is x-measurable. However, $\inf_S(g+\delta) = \inf_S g + \delta > 0$ hence $(g+\delta) \in \mathscr{D}^+_S \cap \mathscr{L}_x$ and $f \geq g + \delta$. Therefore, $f \in \mathscr{E}(\mathscr{D}^+_S \cap \mathscr{L}_x) = \mathscr{C}(\mathscr{D}^+_S \cap \mathscr{L}_x) =: \epsilon_x(\mathscr{D}^+_S)$.

□

Theorem 23. *Let us consider $\Omega = \times_{i \in I} \Omega_i$, where Ω_i is the set of values of a variable X_i for every $i \in I$, and the domain-free information algebra $(\Phi(\Omega), \mathrm{P}(I); \vee, \perp, \cdot, \mathscr{L}(\Omega), \mathscr{L}^+(\Omega), \epsilon)$. In this context, consider a family of sets of gambles $\mathscr{D}_1, \dots, \mathscr{D}_n \in \mathbb{D}(\Omega) \subseteq \Phi(\Omega)$ having supports $S_1, \dots, S_n \subseteq I$ respectively. Then*

$$(\forall i,j)\ \epsilon_{S_i \cap S_j}(\mathscr{D}_j) = \epsilon_{S_i \cap S_j}(\mathscr{D}_i) \tag{C.7}$$

if and only if

$$(\forall i,j)\ \epsilon_{S_i}(\mathscr{D}_i \cdot \mathscr{D}_j) = \mathscr{D}_i. \tag{C.8}$$

Proof. Let us consider a family of coherent sets of gambles $\mathscr{D}_1, \dots., \mathscr{D}_n$ having supports $S_1, .., S_n$ respectively and satisfying Eq. (C.7). Then, we have

$$\epsilon_{S_i}(\mathscr{D}_i \cdot \mathscr{D}_j) = \mathscr{D}_i \cdot \epsilon_{S_i}(\mathscr{D}_j) = \mathscr{D}_i \cdot \epsilon_{S_i}(\epsilon_{S_j}(\mathscr{D}_j)) =$$
$$\mathscr{D}_i \cdot \epsilon_{S_i \cap S_j}(\mathscr{D}_j) = \mathscr{D}_i \cdot \epsilon_{S_i \cap S_j}(\mathscr{D}_i) = \mathscr{D}_i,$$

for every $i, j \in \{1, \dots, n\}$, by Idempotency and Combination axiom and axiom (3.8).

Vice versa, if they satisfy Eq. (C.8), we have

$$\epsilon_{S_i \cap S_j}(\mathscr{D}_j) = \epsilon_{S_i \cap S_j}(\epsilon_{S_j}(\mathscr{D}_i \cdot \mathscr{D}_j)) = \epsilon_{S_i \cap S_j}(\mathscr{D}_i \cdot \mathscr{D}_j),$$

thanks to Lemma 24 in Appendix C.2. Analogously, we have:

$$\epsilon_{S_i \cap S_j}(\mathscr{D}_i) = \epsilon_{S_i \cap S_j}(\epsilon_{S_i}(\mathscr{D}_i \cdot \mathscr{D}_j)) = \epsilon_{S_i \cap S_j}(\mathscr{D}_i \cdot \mathscr{D}_j).$$

Hence, we have the result. □

Proof of Lemma 11. This result can be proven analogously to Lemma 8. □

Proof of Lemma 12. Let us suppose that $\mathscr{D}_1, ..., \mathscr{D}_n$ with supports $x_1, ..., x_n$ respectively, are compatible. Then, there exists a coherent set of gambles $\mathscr{D}$ such that $\epsilon_{x_i}(\mathscr{D}) = \mathscr{D}_i$, for every i. By idempotency axiom, $\mathscr{D} \geq \epsilon_{x_i}(\mathscr{D}) = \mathscr{D}_i$, for every i. From which follows $\mathscr{D} \geq \mathscr{D}_1 \cdot .... \cdot \mathscr{D}_n$ by item 3 of Lemma 25 in Appendix C.2. Hence,

$$\mathscr{D}_i = \epsilon_{x_i}(\mathscr{D}) \geq \epsilon_{x_i}(\mathscr{D}_1 \cdot ... \cdot \mathscr{D}_n) \geq \epsilon_{x_i}(\mathscr{D}_i \cdot \mathscr{D}_j) \geq \epsilon_{x_i}(\mathscr{D}_i) = \mathscr{D}_i,$$

for every $i, j = 1, ..., n$, by Lemma 25 in Appendix C.2. □

Proof of Theorem 18. Let us suppose without loss of generality that the numbering $\{x_1, ..., x_n\}$ guarantees:

$$(\forall i = 1, ..., n-1,\ \exists\ b(i) > i)\ x_i \perp \bigvee_{j=i+1}^{n} x_j | x_{b(i)}. \tag{C.9}$$

Now, let $y_i := x_{i+1} \vee \ldots \vee x_n$ for $i = 1, \ldots, n-1$ and $\mathscr{D} := \mathscr{D}_1 \cdot \ldots \cdot \mathscr{D}_n$. Then, we have:

$$\epsilon_{y_1}(\mathscr{D}) = \epsilon_{y_1}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n,$$

thanks to Combination axiom that can be used because $\mathscr{D}_2 \cdot ... \cdot \mathscr{D}_n$ has support y_1, see Lemma 24 in Appendix C.2.

Now, by Eq.(C.9) and Theorem 21 in Appendix C.2, we have:

$$\epsilon_{y_1}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n = \epsilon_{y_1}(\epsilon_{x_{b(1)}}(\mathscr{D}_1)) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n = \epsilon_{x_{b(1)}}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n$$

where, to show the second equality we used again Lemma 24 in Appendix C.2.

Now, by Combination axiom and pairwise compatibility, we have:

$$\begin{aligned}\epsilon_{x_{b(1)}}(\mathscr{D}_1) \cdot \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n &= \epsilon_{x_{b(1)}}(\mathscr{D}_1 \cdot \mathscr{D}_{b(1)}) \cdot \ldots \cdot \mathscr{D}_{b(1)-1} \cdot \mathscr{D}_{\min\{b(1)+1,n\}} \cdot ... \cdot \mathscr{D}_n \\ &= \mathscr{D}_2 \cdot \ldots \cdot \mathscr{D}_n.\end{aligned}$$

By induction on i, one shows exactly in the same way that

$$\epsilon_{y_i}(\mathscr{D}) = \mathscr{D}_{i+1} \cdot \ldots \cdot \mathscr{D}_n,\ \forall i = 1, ..., n-1.$$

So, we obtain $\epsilon_{y_{n-1}}(\mathscr{D}) := \epsilon_{x_n}(\mathscr{D}) = \mathscr{D}_n$.

Now, we claim that $\epsilon_{x_i}(\mathscr{D}) = \epsilon_{x_i}(\epsilon_{x_{b(i)}}(\mathscr{D})) \cdot \mathscr{D}_i$ for every $i = 1, ..., n-1$. Since $x_{b(i)} \leq y_i$ and $y_i \perp x_i | x_{b(i)}$, we have:

$$\begin{aligned}&\epsilon_{x_i}(\epsilon_{x_{b(i)}}(\mathscr{D})) \cdot \mathscr{D}_i = \mathscr{D}_i \cdot \epsilon_{x_i}(\epsilon_{x_{b(i)}}(\epsilon_{y_i}(\mathscr{D}))) = \mathscr{D}_i \cdot \epsilon_{x_i}(\epsilon_{y_i}(\mathscr{D})) = \\ &\mathscr{D}_i \cdot \epsilon_{x_i}(\mathscr{D}_{i+1} \cdot ... \cdot \mathscr{D}_n) = \epsilon_{x_i}(\mathscr{D}_i \cdot ... \cdot \mathscr{D}_n),\end{aligned}$$

by Lemma 24 , Theorem 21 in Appendix C.2 and Combination axiom. Now, if $i = 1$ we have the result, otherwise, for $i \geq 2$, we have

$$\epsilon_{x_i}(\mathscr{D}_i \cdot ... \cdot \mathscr{D}_n) = \epsilon_{x_i}(\epsilon_{y_{i-1}}(\mathscr{D})) = \epsilon_{x_i}(\mathscr{D}),$$

again by Lemma 24 in Appendix C.2. Then, by backward induction, based on the induction assumption $\epsilon_{x_j}(\mathscr{D}) = \mathscr{D}_j$ for $j > i$, and rooted in $\epsilon_{x_n}(\mathscr{D}) = \mathscr{D}_n$, for

$i = n-1, \dots 1$, we have by pairwise compatibility and Combination axiom:

$$\epsilon_{x_i}(\mathscr{D}) = \epsilon_{x_i}(\epsilon_{x_{b(i)}}(\mathscr{D})) \cdot \mathscr{D}_i = \epsilon_{x_i}(\mathscr{D}_{b(i)}) \cdot \mathscr{D}_i = \epsilon_{x_i}(\mathscr{D}_{b(i)} \cdot \mathscr{D}_i) = \mathscr{D}_i.$$

$\square$

Appendix D

Proofs of Chapter 4

We first show a preliminary result valid for sets of gambles defined on any possibility space Ω, not necessarily finite.

Proposition 17. *Consider a possibility space Ω and $\mathscr{K} \subseteq \mathscr{L}(\Omega)$. If $\mathscr{K}$ is closed under the supremum norm topology, then it satisfies D5'. Vice versa, if $\mathscr{K}$ satisfies the following property:*

$$\text{(PADD) } f \geq g,\ g \in \mathscr{K} \Rightarrow f \in \mathscr{K}$$

then D5' implies the closure under the supremum norm topology.

Proof. It is well-known that $\mathscr{L}$ is a Banach space under the supremum norm, see [Walley, 1991, Appendix D].

Consider then $\mathscr{K} \subseteq \mathscr{L}$ closed under the supremum norm. Thus, given a sequence $(f_n)_{\{n \in \mathbb{N}\}}$ with $f_n \in \mathscr{K}$ for every n, convergent to $f \in \mathscr{L}$ (with respect to the supremum norm), we have $f \in \mathscr{K}$. Consider then a gamble f such that $f + \delta \in \mathscr{K}$ for every $\delta > 0$. We have $f + \frac{1}{n} \in \mathscr{K}$ for every $n \in \mathbb{N} \setminus \{0\}$. The limit of the latter sequence with respect to the supremum norm is f, hence $f \in \mathscr{K}$.

On the other hand, suppose $\mathscr{K}$ satisfies D5' and (PADD). Let us consider a sequence $(f_n)_{\{n \in \mathbb{N}\}}$ with $f_n \in \mathscr{K}$ for every n, convergent with respect to the supremum norm to a gamble $f \in \mathscr{L}$. We know that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\sup |f_n - f| < \epsilon$ for all $n \geq N$. Hence, $f_n < f + \epsilon$ for every $n \geq N$ and $f + \epsilon \in \mathscr{K}$ by (PADD). This procedure can be repeated for every $\epsilon > 0$. Thus, by D5', we have $f \in \mathscr{K}$. $\square$

Since any linear topological space with finite dimension n is isomorphic to $\mathbb{R}^n$ with its usual topology (see for example Walley, 1991, Appendix D), if a set $\mathscr{K} \subseteq \mathbb{R}^n$ satisfies (PADD) we can indifferently say it is closed or it satisfies axiom D5'.

Proof of Proposition 7. Consider a pair of finite sets of gambles (A, R) for which there exists a coherent set of almost desirable gambles $\overline{\mathscr{D}}$, such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. The minimal such set is $\overline{\mathscr{E}(A)}$ (see Proposition 17 in Appendix D). $\overline{\mathscr{E}(A)}$ is a *finitely generated cone*, see [Greer, 1984, Definition 2.3.2]. Hence, it is also *polyhedral cone* [Greer, 1984, Theorem 2.3.4], thus an intersection of a finite number of closed halfspaces whose bounding hyperspaces pass through the origin:

$$\overline{\mathscr{E}(A)} = \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f}^\top \boldsymbol{\beta}_j \geq 0,\ j = 1, \ldots, N\} \tag{D.1}$$

with $\boldsymbol{\beta}_j \in \mathbb{R}^n$ for every $j \in \{1, \ldots, N\}$, $N \geq 1$. This concludes the first part of the proof since it tells us that there exists a binary piecewise linear classifier PLC with parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$ classifying $A \cup T \subseteq \overline{\mathscr{E}(A)} = \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\}$ as 1 and $(R \cup F)$, which has empty intersection with $\overline{\mathscr{E}(A)}$, as -1. Notice that a similar reasoning can be repeated for every finitely generated coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, not necessarily equal to $\overline{\mathscr{E}(A)}$.

Vice versa, consider a piecewise linearly separable pair $(A \cup T, R \cup F)$, where A and R are finite sets of gambles. Consider then a classifier $PLC \in \mathrm{PLC}(A \cup T, R \cup F)$. Thus, we have:

$$\{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\} = \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f}^\top \boldsymbol{\beta}_j \geq 0, \text{ for all } j = 1, .., N\}, \tag{D.2}$$

for some $\boldsymbol{\beta}_j \in \mathbb{R}^n$ for every $j \in \{1, \ldots, N\}$, $N \geq 1$. Thus, $\overline{\mathscr{D}} := \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\}$ is a coherent set of almost desirable gambles. Indeed, it satisfies D1', D2' by hypothesis and D3'–D5' by Eq. (D.2). Moreover, thanks to the correspondence between finitely generated and polyhedral cones, it is also finitely generated. Finally, notice that $A \subseteq \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\} = \overline{\mathscr{D}}$ and $R \cap \{\boldsymbol{f} \in \mathbb{R}^n : PLC(\boldsymbol{f}) = 1\} = R \cap \overline{\mathscr{D}} = \emptyset$ by hypothesis. □

Proof of Proposition 8. Consider a piecewise linearly separable pair $(A \cup T, R \cup F)$, where A and R are finite sets of gambles. Consider also a classifier $PLC \in \mathrm{PLC}(A \cup T, R \cup F)$ characterised by the parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$. Let us assume without loss of generality that $\boldsymbol{\beta}_j \neq 0$ for every j.

$PLC \in \mathrm{PLC}(A, R)$. Now, assume there is a vector $\boldsymbol{\beta}_k$ with strictly negative i-th component. Consider then $\boldsymbol{t} \in T$ such that

$$t_l = \begin{cases} 1 & \text{if } l = i, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\boldsymbol{t}^\top \boldsymbol{\beta}_k < 0$ and $PLC(\boldsymbol{t}) = -1$, a contradiction. Therefore, $\boldsymbol{\beta}_j \gneq 0$ for every $j \in \{1, \ldots, N\}$.

The converse immediately follows. □

Proof of Corollary 6. The proof follows from Proposition 7 and Proposition 8. □

Proof of Proposition 9. Consider a binary piecewise linear classifier PLC with parameters $\{\boldsymbol{\beta}_j\}_{j=1}^N$ and a classifier $LC_{\boldsymbol{\phi}}$ of type (4.5) with parameters $\{\boldsymbol{\omega}_j, \boldsymbol{\beta}_j'\}_{j=1}^N$ such that $\boldsymbol{\beta}_j' = \boldsymbol{\omega}_j = \boldsymbol{\beta}_j$ for all $j = 1, ..., N$. They classify gambles in the same way. Indeed, consider $\boldsymbol{f} \in \mathbb{R}^n$ and let us define $m := \min(\boldsymbol{f}^\top \boldsymbol{\beta}_1, \ldots, \boldsymbol{f}^\top \boldsymbol{\beta}_N)$. Then:

$$\sum_{j=1}^N \boldsymbol{\phi}_j(\boldsymbol{f})^\top \boldsymbol{\beta}_j := \sum_{j=1}^N (\mathbb{I}_{\mathscr{B}_j}(\boldsymbol{f})\boldsymbol{f})^\top \boldsymbol{\beta}_j = \sum_{k=1}^K \boldsymbol{f}^\top \boldsymbol{\beta}_k = Km,$$

where $\{\mathscr{B}_j\}_{j=1}^N$ is the partition whose elements are specified by Eq.(4.3), with $\boldsymbol{\omega}_j = \boldsymbol{\beta}_j$ for every $j \in \{1, ..., N\}$ and $\boldsymbol{f}^\top \boldsymbol{\beta}_k = m$ for all $k = 1, ..., K$ with $1 \leq K \leq N$. Hence, $\boldsymbol{f}$ is classified in the same way by the classifiers PLC and $LC_{\boldsymbol{\phi}}$ because $m \geq 0 \iff \sum_{j=1}^N \boldsymbol{\phi}_j(\boldsymbol{f})^\top \boldsymbol{\beta}_j \geq 0$.

□

Proof of Corollary 8. The proof follows from Corollary 6 and Proposition 9. □

Lemma 26. *If a set $\mathscr{K} \subseteq \mathbb{R}^n$, satisfies D1', (CNV) and D5' then it satisfies (PADD).*

Proof. Consider $\boldsymbol{f} \geq \boldsymbol{g}$ with $\boldsymbol{g} \in \mathscr{K}$. If $\boldsymbol{f} = \boldsymbol{g}$, then $\boldsymbol{f} \in \mathscr{K}$. Otherwise, $\boldsymbol{f} = \boldsymbol{g} + \boldsymbol{t}$ with $\boldsymbol{t} \in T$.

Now, for every $\epsilon > 0$ there exists a $\gamma \in (0,1)$ such that $\gamma \boldsymbol{g} \leq \boldsymbol{g} + \epsilon$. Thus, we have $\boldsymbol{f} + \epsilon = \boldsymbol{g} + \epsilon + \boldsymbol{t} = \gamma \boldsymbol{g} + (1-\gamma)\frac{(\boldsymbol{g}+\epsilon-\gamma \boldsymbol{g})+\boldsymbol{t}}{1-\gamma}$. Now, $\boldsymbol{g} \in \mathscr{K}$ by hypothesis and $\frac{(\boldsymbol{g}+\epsilon-\gamma \boldsymbol{g})+\boldsymbol{t}}{1-\gamma} \in T \subseteq \mathscr{K}$ by D1'. Hence, by (CNV), $\boldsymbol{f} + \epsilon \in \mathscr{K}$. The same reasoning can be repeated for every $\epsilon > 0$, hence $\boldsymbol{f} \in \mathscr{K}$ by D5'. □

Lemma 27. *Consider a pair of finite sets (A, R) for which there exists a convex coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. The smallest such set is $\overline{\mathscr{D}} = \overline{\mathrm{ch}(A \cup T)}$.*

Proof. $\overline{\mathrm{ch}(A \cup T)}$ satisfies D1' by definition and (CNV) by [Rockafellar, 1970, Theorem 6.2]. Moreover, it also satisfies D5' by Proposition 17 in Appendix D.

Let us indicate with $\mathrm{D}(A, R)$, the class of convex coherent sets of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Every $\overline{\mathscr{D}} \in \mathrm{D}(A, R)$ clearly contains $\overline{\mathrm{ch}(A \cup T)}$. However, $\overline{\mathscr{D}} \cap (R \cup F) = \emptyset$. Hence, $\overline{\mathrm{ch}(A \cup T)} \cap (R \cup F) = \emptyset$ as well. Thus, $\overline{\mathrm{ch}(A \cup T)}$ belongs to $\mathrm{D}(A, R)$ and it is the smallest set contained in it.

This concludes the proof. □

Lemma 28. *Consider a finite set $A \subseteq \mathbb{R}^n$. Then:*

$$\overline{\mathrm{ch}(A \cup T)} = \mathrm{ch}^+(A \cup \{\mathbf{0}\}) := \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f} \geq \boldsymbol{g},\ \boldsymbol{g} \in \mathrm{ch}(A \cup \{\mathbf{0}\})\}.$$

Proof. First, we can observe that:

$$\mathrm{ch}^{+}(A \cup \{\mathbf{0}\}) := \{\boldsymbol{f} \in \mathbb{R}^{n} : \boldsymbol{f} \geq \boldsymbol{g},\ \boldsymbol{g} \in \mathrm{ch}(A \cup \{\mathbf{0}\})\} =$$
$$= \sum_{i \in I} \gamma_i \boldsymbol{f}_i + \sum_{j \in J} \lambda_j \boldsymbol{e}_j =: \mathrm{ch}(A \cup \{\mathbf{0}\}) + \mathrm{posi}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$$

with I, J finite and $\boldsymbol{f}_i \in A \cup \{\mathbf{0}\}$, $\gamma_i, \lambda_i \geq 0$ for every i and $\sum_i \gamma_i = 1$, where $\{\boldsymbol{e}_j\}_{j=1}^{n}$ is the canonical basis of $\mathbb{R}^n$ and $\mathrm{posi}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is a finitely generated cone, hence a polyhedral cone. From [Schrijver, 1998, Corollary 7.1.b], it follows that $\mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$ is a convex (closed) polyhedron. Hence, $\overline{\mathrm{ch}^{+}(A \cup \{\mathbf{0}\})} = \mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$. Now, we divide the rest of the proof into two parts.

- $\overline{\mathrm{ch}(A \cup T)} \subseteq \mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$. Notice that, thanks to the previous observation, it is sufficient to show that $\mathrm{ch}(A \cup T) \subseteq \mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$. So, let us consider $\boldsymbol{f} \in \mathrm{ch}(A \cup T)$. By definition, we have:
$$\boldsymbol{f} = \sum_{k=1}^{r} \gamma_k \boldsymbol{f}_k$$
with $\gamma_k \geq 0, \boldsymbol{f}_k \in (A \cup T)$, for all $k = 1, \ldots, r$, $r \geq 1$, $\sum_{k=1}^{r} \gamma_k = 1$. Let us indicate with $Ind_{A \setminus T} := \{k \in \{1, \ldots, r\}$ such that : $\boldsymbol{f}_k \in A \setminus T\}$ and $Ind_T := \{k \in \{1, \ldots, r\}$ such that : $\boldsymbol{f}_k \in T\}$. Then, we have:
$$\boldsymbol{f} \geq \sum_{k \in Ind_{A \setminus T}} \gamma_k \boldsymbol{f}_k + \sum_{k \in Ind_T} \gamma_k \mathbf{0},$$
hence $\boldsymbol{f} \in \mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$.

- $\mathrm{ch}^{+}(A \cup \{\mathbf{0}\}) \subseteq \overline{\mathrm{ch}(A \cup T)}$. By definition, $\overline{\mathrm{ch}(A \cup T)}$ is a closed convex set that contains T. Therefore, from Proposition 17 in Appendix D and Lemma 26 in Appendix D, we have:
$$\mathrm{ch}(A \cup \{\mathbf{0}\}) \subseteq \overline{\mathrm{ch}(A \cup T)} \Rightarrow$$
$$\mathrm{ch}^{+}(A \cup \{\mathbf{0}\}) \subseteq \overline{\mathrm{ch}(A \cup T)}.$$

$\square$

Proof of Proposition 10. Consider a pair of finite sets of gambles (A, R) for which there exists a convex coherent set of almost desirable gambles $\overline{\mathcal{D}}$, such that $\overline{\mathcal{D}} \supseteq A$ and $\overline{\mathcal{D}} \cap R = \emptyset$. The minimal convex coherent set of almost desirable gambles that satisfies these conditions is $\overline{\mathrm{ch}(A \cup T)}$, see Lemma 27 in Appendix D. Moreover, thanks to Lemma 28 in Appendix D, we know that it can be rewritten as:

$$\overline{\mathrm{ch}(A \cup T)} = \mathrm{ch}^{+}(A \cup \{\mathbf{0}\}), \tag{D.3}$$

where $\mathrm{ch}^{+}(A \cup \{\mathbf{0}\})$ is a convex polyhedron. Any convex polyhedron can be written as an intersection of hyper-spaces, whose border is a piecewise affine function, see [Schrijver, 1998, Section 7]:

$$\overline{\mathrm{ch}(A \cup T)} = \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f}^\top \boldsymbol{\beta}_j + \alpha_j \geq 0,\ j = 1, \ldots, N\} \tag{D.4}$$

with $\boldsymbol{\beta}_j \in \mathbb{R}^n$, $\alpha_j \in \mathbb{R}$ for every $j \in \{1, \ldots, N\}$, $N \geq 1$. There exists therefore a binary piecewise affine classifier PAC, such that $\overline{\mathrm{ch}(A \cup T)} = \overline{\mathrm{ch}^+(A \cup \{\mathbf{0}\})} = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\}$. Note moreover that $\overline{\mathrm{ch}(A \cup T)} = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\} \supseteq (A \cup T)$ and $\overline{\mathrm{ch}(A \cup T)} \cap (R \cup F) = \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\} \cap (R \cup F) = \emptyset$ by hypothesis. This concludes the first part of the proof.

Notice that a similar reasoning can be repeated for every finitely generated convex coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, not necessarily equal to $\overline{\mathrm{ch}(A \cup T)}$.

Vice versa, consider a piecewise affine separable pair $(A \cup T, R \cup F)$, where A and R are finite sets. Let us consider a binary piecewise affine classifier $PAC \in \mathrm{PAC}(A \cup T, R \cup F)$. Now, the set:

$$\begin{aligned}\overline{\mathscr{D}} :=& \{\boldsymbol{f} \in \mathbb{R}^n : PAC(\boldsymbol{f}) = 1\} = \\ & \{\boldsymbol{f} \in \mathbb{R}^n : \boldsymbol{f}^\top \boldsymbol{\beta}_j + \alpha_j \geq 0, \text{ for all } j = 1, \ldots, N\},\end{aligned}$$

for some $\boldsymbol{\beta}_j \in \mathbb{R}^n$ and $\alpha_j \in \mathbb{R}$ for all $j \in \{1, \ldots N\}$ is a convex coherent set of gambles such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Indeed:

- $T \subseteq \overline{\mathscr{D}}$ and $\overline{\mathscr{D}} \cap F = \emptyset$ by definition, hence it satisfies D1' and D2';
- $\overline{\mathscr{D}}$ satisfies (CNV). Consider $\boldsymbol{f}, \boldsymbol{g} \in \overline{\mathscr{D}}$. Then $\gamma \boldsymbol{f} + (1-\gamma)\boldsymbol{g} \in \overline{\mathscr{D}}$, for all $\gamma \in [0,1]$. Indeed,

$$\begin{aligned}(\gamma \boldsymbol{f} + (1-\gamma)\boldsymbol{g})^\top \boldsymbol{\beta}_j + \alpha_j &= \\ (\gamma \boldsymbol{f})^\top \boldsymbol{\beta}_j + ((1-\gamma)\boldsymbol{g})^\top \boldsymbol{\beta}_j + \gamma \alpha_j + (1-\gamma)\alpha_j &= \\ \gamma(\boldsymbol{f}^\top \boldsymbol{\beta}_j + \alpha_j) + (1-\gamma)(\boldsymbol{g}^\top \boldsymbol{\beta}_j + \alpha_j) &\geq 0\end{aligned}$$

  for all $j \in \{1, \ldots, N\}$ and $\gamma \in [0,1]$.
- $\overline{\mathscr{D}}$ is closed under the usual topology of $\mathbb{R}^n$ because it is the intersection of a finite number of closed half-spaces hence, thanks to Proposition 17 in Appendix D, it satisfies D5';
- clearly, by the fact that $PAC \in \mathrm{PAC}(A \cup T, R \cup F)$, it is also true that $A \subseteq \overline{\mathscr{D}}$ and $\overline{\mathscr{D}} \cap R = \emptyset$.

□

Example 21. *Let us consider* $A := \{[-1,2]^\top, [2,-1]^\top, [1,-0.5]^\top\}$ *and* $R := \{[-3,2]^\top, [1,-1]^\top\}$. *Then,* $(A \cup T, R \cup F)$ *is piecewise affine separable (see Example 15). However,* $(A \cup T, R \cup F)$ *is also piecewise linearly separable through the classifier* PLC_C

defined in Example 13. Notice, in particular, that $PLC_C \in PAC(\overline{A \cup T}, R \cup F)$*. Suppose, however, that* $\{\boldsymbol{f} \in \mathbb{R}^n : PLC_C(\boldsymbol{f}) = 1\} = \overline{\mathrm{ch}(A' \cup T)} = \mathrm{ch}^+(A' \cup \{\boldsymbol{0}\})$ *for some finite set* A'*. Since* $\{\boldsymbol{f} \in \mathbb{R}^n : PLC_C(\boldsymbol{f}) = 1\} \neq T$*, there exists at least a gamble* $\boldsymbol{f}' \in A'$ *such that* $\boldsymbol{f}' \notin T$*, i.e.,* $f_i' < 0$ *for some* $i \in \{1,2\}$*. Let us suppose moreover that* $f_i' = \min_{\boldsymbol{f} \in A'} f_i$*. We can choose* $\lambda > 1$ *sufficiently big such that* $\lambda \boldsymbol{f}' \notin \mathrm{ch}(A' \cup \{\boldsymbol{0}\})$*, but* $\lambda \boldsymbol{f} \in \{\boldsymbol{f} \in \mathbb{R}^n : PLC_C(\boldsymbol{f}) = 1\}$ *since the latter set is coherent. However,* $\lambda \boldsymbol{f}' \notin \mathrm{ch}^+(A' \cup \{\boldsymbol{0}\})$ *either, since* $\lambda f_i' < f_i' \leq g_i$*, for every* $\boldsymbol{g} \in \mathrm{ch}(A' \cup \{\boldsymbol{0}\})$.

Proof of Proposition 11. Consider a piecewise affine separable pair $(A \cup T, R \cup F)$, where A and R are finite sets of gambles. Consider also a classifier PAC in $\mathrm{PAC}(A \cup T, R \cup F)$ characterised by the parameters $\{\boldsymbol{\beta}_j, \alpha_j\}_{j=1}^N$ such that $\boldsymbol{\beta}_j \in \mathbb{R}^n$, $\alpha_j \in \mathbb{R}$ for every $j \in \{1, ..., N\}$. Let us assume without loss of generality that $\boldsymbol{\beta}_j \neq 0$ for every j.

$PAC \in \mathrm{PAC}(A, R)$. Moreover, let us suppose $\alpha_k < 0$ for some k. Then, $\boldsymbol{0}^\top \boldsymbol{\beta}_k + \alpha_k < 0$ and $PAC(\boldsymbol{0}) = -1$, a contradiction. Therefore, $\alpha_j \geq 0$ for every $j \in \{1, ..., N\}$.

Now, assume there is a vector $\boldsymbol{\beta}_k$ with strictly negative i-th component. Then, consider $\boldsymbol{t} \in T$ and $\epsilon > 0$ such that:

$$t_l = \begin{cases} \frac{\alpha_k + \epsilon}{|\boldsymbol{\beta}_k^T \boldsymbol{e}_i|} & \text{if } l = i, \\ 0 & \text{otherwise,} \end{cases}$$

where, as before, $\{\boldsymbol{e}_j\}_{j=1}^n$ is the canonical basis of $\mathbb{R}^n$. Then, $\boldsymbol{t}^\top \boldsymbol{\beta}_k + \alpha_k = -\epsilon < 0$, so $PAC(\boldsymbol{t}) = -1$, a contradiction. Therefore $\boldsymbol{\beta}_j \gneq 0$ for every $j \in \{1, ..., N\}$.

Finally, let us suppose $\alpha_j > 0$ for every $j \in \{1, ..., N\}$. Then, consider the gamble $\boldsymbol{f} \in F$ such that $f_i := -\frac{\min_k \alpha_k}{n * max_{i,j}(\beta_{ji})}$ for every $i \in \{1, .., n\}$. Then, $\boldsymbol{f}^\top \boldsymbol{\beta}_j + \alpha_j \geq -\min_k \alpha_k + \alpha_j \geq 0$ for every j. Hence, $PAC(\boldsymbol{f}) = 1$, a contradiction. So, there exists at least a $k \in \{1, ..., N\}$ such that $\alpha_k = 0$.

The converse immediately follows. □

Proof of Corollary 9. The proof follows from Proposition 10 and Proposition 11. □

Proof of Proposition 12. Consider a binary piecewise affine classifier PAC with parameters $\{\boldsymbol{\beta}_j, \alpha_j\}_{j=1}^N$ and a classifier LC_ψ of type (4.9) with parameters $\{\boldsymbol{\omega}_j', \boldsymbol{\beta}_j'\}_{j=1}^N$ such that $\boldsymbol{\beta}_j' = \boldsymbol{\omega}_j' = \begin{bmatrix} \boldsymbol{\beta}_j \\ \alpha_j \end{bmatrix}$ for all $j = 1, ..., N$. They classify gambles in the same way. Indeed, consider $\boldsymbol{f} \in \mathbb{R}^n$ and let us define $m := \min(\boldsymbol{f}^\top \boldsymbol{\beta}_1 + \alpha_1, \ldots, \boldsymbol{f}^\top \boldsymbol{\beta}_N +$

α_N). Then:

$$\sum_{j=1}^{N} \psi_j(f)^\top \begin{bmatrix} \beta_j \\ \alpha_j \end{bmatrix} := \sum_{j=1}^{N} \mathbb{I}_{\mathscr{B}'_j}(f)(f^\top \beta_j + \alpha_j) = Km$$

where $\{\mathscr{B}'_j\}_{j=1}^N$ is the partition whose elements are specified by Eq.(4.7) with $\omega'_j = \begin{bmatrix} \beta_j \\ \alpha_j \end{bmatrix}$ for every $j \in \{1,...,N\}$ and where $1 \leq K \leq N$. Hence, f is classified in the same way by the classifiers PAC and LC_ψ because $m \geq 0 \iff \sum_{j=1}^N \psi_j(f)^\top \begin{bmatrix} \beta_j \\ \alpha_j \end{bmatrix} \geq 0$.

□

Proof of Corollary 11. The proof follows from Corollary 9 and Proposition 12. □

Lemma 29. *Consider a pair of finite sets (A,R) for which there exists a positive additive coherent set of almost desirable gambles $\overline{\mathscr{D}}$, such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. The smallest such set is:*

$$\overline{\mathscr{D}} = \uparrow(A \cup \{\mathbf{0}\}) := \{f \in \mathbb{R}^n : (\exists g \in A \cup \{\mathbf{0}\})\, f \geq g\}.$$

Proof. $\uparrow(A \cup \{\mathbf{0}\})$ satisfies D1' and (PADD) by definition. Moreover, it is closed (it is a finite union of closed sets), hence it also satisfies D5' by Proposition 17 in Appendix D.

Let us indicate with $\mathrm{P}(A,R)$, the class of positive additive coherent sets of almost desirable gambles $\overline{\mathscr{D}}$, such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Clearly, every $\overline{\mathscr{D}} \in \mathrm{P}(A,R)$ satisfies $\overline{\mathscr{D}} \supseteq \uparrow(A \cup \{\mathbf{0}\})$ thanks to (PADD). However, $\overline{\mathscr{D}} \cap (R \cup F) = \emptyset$. Therefore, $\uparrow(A \cup \{\mathbf{0}\}) \cap (R \cup F) = \emptyset$ as well. Thus, $\uparrow(A \cup \{\mathbf{0}\})$ belongs to $\mathrm{P}(A,R)$ and it is the smallest set contained in it.

This concludes the proof. □

Proof of Proposition 13. Consider a pair of finite sets of gambles (A,R) for which there exists a positive additive coherent set of almost desirable gambles $\overline{\mathscr{D}}$, such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$. Then the smallest such set is $\uparrow(A \cup \{\mathbf{0}\})$, see Lemma 29 in Appendix D. In particular, the latter can be rewritten as:

$$\uparrow(A \cup \{\mathbf{0}\}) = \{f \in \mathbb{R}^n : PWPC(f) = 1\}$$

where $PWPC$ is a PWP classifier, defined as:

$$(\forall f \in \mathbb{R}^n)\; PWPC(f) := \begin{cases} 1 & \text{if } \exists g^j \in (A \cup \{\mathbf{0}\}) \text{ s.t. } f \geq g^j, \\ -1 & \text{otherwise.} \end{cases}$$

This concludes the proof since $A \cup T \subseteq \uparrow(A \cup \{0\}) = \{f \in \mathbb{R}^n : PWPC(f) = 1\}$ and $\uparrow(A \cup \{\mathbf{0}\}) \cap (R \cup F) = \{f \in \mathbb{R}^n : PWPC(f) = 1\} \cap (R \cup F) = \emptyset$. Similar reasoning can be repeated for every finitely generated positive additive coherent set of almost desirable gambles $\overline{\mathscr{D}}$ such that $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$, not necessarily equal to $\uparrow (A \cup \{\mathbf{0}\})$.

Vice versa, consider a PWP separable pair $(A \cup T, R \cup F)$, where A and R are finite sets of gambles. Consider also a classifier $PWPC \in \mathrm{PWPC}(A \cup T, R \cup F)$. Then:

$$\overline{\mathscr{D}} := \{\boldsymbol{f} \in \mathbb{R}^n : PWPC(\boldsymbol{f}) = 1\}$$

is, by construction, a positive additive coherent set. Indeed, it satisfies D1', D2', (PADD). Further, it is closed because it is a finite union of closed sets hence, by Proposition 17 in Appendix D, it also respects axiom D5'. It satisfies also $\overline{\mathscr{D}} \supseteq A$ and $\overline{\mathscr{D}} \cap R = \emptyset$ by hypothesis. Moreover, $\{\boldsymbol{f} \in \mathbb{R}^n : PWPC(\boldsymbol{f}) = 1\} = \uparrow (\mathscr{G} \cup \{\mathbf{0}\})$, where $\mathscr{G}$ is the set of generators of $PWPC$. Hence $\overline{\mathscr{D}}$ is also finitely generated. □

Proof of Proposition 14. Consider a PWP separable pair $(A \cup T, R \cup F)$, where A and R are finite sets of gambles. Consider also a classifier $PWPC \in \mathrm{PWPC}(A \cup T, R \cup F)$, characterised by the generators $\{\boldsymbol{g}^j\}_{j=1}^N$ such that $\boldsymbol{g}^j \in \mathbb{R}^n$ for every $j \in \{1, \ldots N\}$. Clearly, $PWPC \in \mathrm{PWPC}(A, R)$.

Moreover, let us suppose there is a $\boldsymbol{g}^k$ such that $\boldsymbol{g}^k < 0$. Consider $0 < \epsilon < \min_i |\boldsymbol{g}_i^k|$. Then $\boldsymbol{f} := -\epsilon > \boldsymbol{g}^k$, so $PWPC(\boldsymbol{f}) = 1$. But this is a contradiction because $\boldsymbol{f} \in F$. Therefore, for every $j \in \{1, .., N\}$, there exists $i \in \{1, \ldots, n\}$ such that $\boldsymbol{g}_i^j \geq 0$.

Let us suppose now that for every $j \in \{1, \ldots N\}$ there exists $l \in \{1, \ldots, n\}$ such that $g_l^j > 0$. Consider $t = \mathbf{0}$. Then $t \not\geq \boldsymbol{g}^j$ for every $j \in \{1, \ldots, N\}$. Hence, $PWPC(\mathbf{0}) = -1$ that is a contradiction. Therefore there exists at least a $\boldsymbol{g}^k$ with $k \in \{1, .., N\}$ such that $\boldsymbol{g}^k \leq 0$ and $\boldsymbol{g}^k \not< 0$.

The converse immediately follows. □

Proof of Corollary 12. The proof follows from Proposition 13 and Proposition 14. □

Proof of Proposition 15. Consider a PWP classifier $PWPC$ with generators $\mathscr{G} = \{\boldsymbol{g}^j\}_{j=1}^N$ and a classifier LC_ρ of type (4.14) with parameters $\{\boldsymbol{\omega}^j, \boldsymbol{\beta}_j'\}_{j=1}^N$ such that

$$\boldsymbol{\beta}_j' = \begin{bmatrix} 1 \\ \cdots \\ 1 \\ -\omega_1^j \\ \cdots \\ -\omega_n^j \end{bmatrix}$$ and $\boldsymbol{\omega}^j = \boldsymbol{g}^j$ for all $j = 1, \ldots, N$. They classify gambles in the same

way. Indeed, consider $\boldsymbol{f} \in \mathbb{R}^n$ and let us define $m := \max_{\{k\in\{1,..,N\}\}}(\min_{\{l\in\{1,...,n\}\}}(f_l - g_l^k))$. Then:

$$\sum_{j=1}^{N} \boldsymbol{\rho}_j(\boldsymbol{f})^\top \boldsymbol{\beta}_j' := \sum_{j=1}^{N}\sum_{i=1}^{n} \mathbb{I}_{\zeta_{i,j}}(\boldsymbol{f})(f_i - g_i^j) = KLm$$

where $\{\zeta_{i,j}\}_{i\in\{1,...,n\}\},\{j\in\{1,...,N\}}$ is the partition whose elements are specified by Eq.(4.12) with $\omega_i^j = g_i^j$ for every i, j and where $1 \leq L \leq n$, $1 \leq K \leq N$. Hence, $\boldsymbol{f}$ is classified in the same way by the classifiers $PWPC$ and LC_{ρ} because $(\exists \boldsymbol{g}^j \in \mathcal{G} : \boldsymbol{f} \geq \boldsymbol{g}^j) \iff m \geq 0 \iff \sum_{j=1}^{N} \boldsymbol{\rho}_j(\boldsymbol{f})^\top \boldsymbol{\beta}_j' \geq 0$.

□

Proof of Corollary 13. The proof follows from Corollary 12 and Proposition 15.

□

Proof of Proposition 16. Let us consider a set $\overline{\mathcal{D}} \subseteq \mathbb{R}^n$ satisfying D1', D2', (PADD), D5'. We proceed by addressing the points of the statement in turn.

- Let us consider $\boldsymbol{f} \in \mathbb{R}^n$. The set $\{\mu \in \mathbb{R} : \boldsymbol{f} - \mu \in \overline{\mathcal{D}}\}$ is not empty. Indeed, by D1', $\boldsymbol{f} - \mu = t \in T \subseteq \overline{\mathcal{D}}$ for any $\mu \leq \min_i f_i$. Moreover, $\boldsymbol{f} - (\max_i f_i + \epsilon) \in F$ for every $\epsilon > 0$ in combination with D2' proves the set $\{\mu \in \mathbb{R} : \boldsymbol{f} - \mu \in \overline{\mathcal{D}}\}$ is also bounded from above. This permits to conclude that $dom(\underline{P}) = \mathbb{R}^n$ and $\underline{P}(\boldsymbol{f}) \in \mathbb{R}$ for any $\boldsymbol{f} \in \mathbb{R}^n$.

- The previous point also shows that $\min_i f_i \leq \underline{P}(\boldsymbol{f}) \leq \max_i f_i$ for every $\boldsymbol{f} \in \mathbb{R}^n$.

- $\boldsymbol{f} - \underline{P}(\boldsymbol{f}) \in \overline{\mathcal{D}}$ by D5'. Since, moreover, $\boldsymbol{f} - \underline{P}(\boldsymbol{f}) - \epsilon \notin \overline{\mathcal{D}}$ for every $\epsilon > 0$, we have $\underline{P}(\boldsymbol{f} - \underline{P}(\boldsymbol{f})) = 0$.

- Let us consider $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^n$ such that $\boldsymbol{f} \geq \boldsymbol{g}$. Then
$$\{\mu \in \mathbb{R} : \boldsymbol{g} - \mu \in \overline{\mathcal{D}}\} \subseteq \{\mu \in \mathbb{R} : \boldsymbol{f} - \mu \in \overline{\mathcal{D}}\},$$
because if $\boldsymbol{g} - \mu \in \overline{\mathcal{D}}$ then $\boldsymbol{f} - \mu \geq \boldsymbol{g} - \mu \in \overline{\mathcal{D}}$ by (PADD). Hence, $\underline{P}(\boldsymbol{f}) \geq \underline{P}(\boldsymbol{g})$.

- Let us consider $\boldsymbol{f} + r$ for some $\boldsymbol{f} \in \mathbb{R}^n$ and $r \in \mathbb{R}$. Then
$$\underline{P}(\boldsymbol{f}) + r \in \{\mu \in \mathbb{R} : \boldsymbol{f} + r - \mu \in \overline{\mathcal{D}}\}.$$
Indeed, $\boldsymbol{f} + r - (\underline{P}(\boldsymbol{f}) + r) = \boldsymbol{f} - \underline{P}(\boldsymbol{f}) \in \overline{\mathcal{D}}$. However, $\boldsymbol{f} + r - (\underline{P}(\boldsymbol{f}) + r + \epsilon) = \boldsymbol{f} - \underline{P}(\boldsymbol{f}) - \epsilon \notin \overline{\mathcal{D}}$ for every $\epsilon > 0$. Hence, $\underline{P}(\boldsymbol{f}) + r = \max\{\mu \in \mathbb{R} : \boldsymbol{f} + r - \mu \in \overline{\mathcal{D}}\} =: \underline{P}(\boldsymbol{f} + r)$.

- Let us suppose $\underline{P}(\mathbf{0}) = \epsilon > 0$, then $\underline{P}(-\epsilon) = \underline{P}(\mathbf{0}-\epsilon) = \underline{P}(\mathbf{0})-\epsilon = \epsilon-\epsilon = 0$, by translation invariance. Hence, by definition of $\underline{P}(-\epsilon)$, we can conclude that $-\epsilon - \underline{P}(-\epsilon) = -\epsilon \in F \cap \overline{\mathscr{D}}$, which is a contradiction since $\overline{\mathscr{D}}$ satisfies D2'. If instead $\underline{P}(\mathbf{0}) < 0$, then $\mathbf{0}-0 = \mathbf{0} \notin \overline{\mathscr{D}}$, which is a contradiction since $\overline{\mathscr{D}}$ satisfies D1'. Therefore, $\underline{P}(\mathbf{0}) = 0$.

Suppose now that $\overline{\mathscr{D}}$ also satisfies (CNV). Let us consider $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^n$ and $\gamma \in [0,1]$. Then

$$\gamma\underline{P}(\boldsymbol{f}) + (1-\gamma)\underline{P}(\boldsymbol{g}) \in \{\mu \in \mathbb{R} : \gamma\boldsymbol{f} + (1-\gamma)\boldsymbol{g} - \mu \in \overline{\mathscr{D}}\}.$$

Indeed, $\gamma\boldsymbol{f} + (1-\gamma)\boldsymbol{g} - \gamma\underline{P}(\boldsymbol{f}) - (1-\gamma)\underline{P}(\boldsymbol{g}) = \gamma(\boldsymbol{f} - \underline{P}(\boldsymbol{f})) + (1-\gamma)(\boldsymbol{g} - \underline{P}(\boldsymbol{g})) \in \overline{\mathscr{D}}$. Hence $\gamma\underline{P}(\boldsymbol{f}) + (1-\gamma)\underline{P}(\boldsymbol{g}) \leq \sup\{\mu \in \mathbb{R} : \gamma\boldsymbol{f} + (1-\gamma)\boldsymbol{g} - \mu \in \overline{\mathscr{D}}\} =: \underline{P}(\gamma\boldsymbol{f} + (1-\gamma)\boldsymbol{g})$.

The other properties follow from [Walley, 1991, Section 2.6.1]. □

Bibliography

Aczél, J. and Wagner, C. [1980]. A characterization of weighted arithmetic means, *SIAM Journal on Algebraic Discrete Methods* **1**(3): 259–260.

Allais, M. [1953]. Le comportment de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école americaine, *Econometrica* **22**.

Anscombe, F. J. and Aumann, R. J. [1963]. A definition of subjective probability, *The Annals of Mathematical Statistics* **34**: 199–2005.

Arrow, K. [1951]. *Social Choice and Individual Values*, John Wiley & Sons.

Beeri, C., Fagin, R., Maier, D. and Yannakakis, M. [1983]. On the desirability of acyclic database schemes, *Journal of the ACM (JACM)* **30**(3): 479–513.

Benavoli, A., Facchini, A., Piga, D. and Zaffalon, M. [2019]. Sum-of-squares for bounded rationality, *International Journal of Approximate Reasoning* **105**: 130 – 152.

Bewley, T. F. [2002]. Knightian decision theory. Part I, *Decisions in Economics and Finance* **25**(2): 79–110.

Boole, G. [1854]. *An Investigation of the Laws of Thought: on which are Founded the Mathematical Theories of Logic and Probabilities*, Dover.

Buss, S. R. [1998]. *Handbook of Proof Theory*, Elsevier.

Casanova, A., Benavoli, A. and Zaffalon, M. [2021]. Nonlinear desirability as a linear classification problem, *ISIPTA 2021: Proceedings of International Symposium on Imprecise Probabilities: Theories and Applications*, Vol. 147, PMLR, pp. 61–71.

Casanova, A., Benavoli, A. and Zaffalon, M. [2023]. Nonlinear desirability as a linear classification problem, *International Journal of Approximate Reasoning* **152**: 1–32.

Casanova, A., Kohlas, J. and Zaffalon, M. [2021]. Algebras of sets and coherent sets of gambles, *ECSQARU 2021: Proceedings of European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Vol. 12897, Lecture Notes in Computer Science, Springer, pp. 603–615.

Casanova, A., Kohlas, J. and Zaffalon, M. [2022a]. Information algebras in the theory of imprecise probabilities, *International Journal of Approximate Reasoning* **142**: 383–416.

Casanova, A., Kohlas, J. and Zaffalon, M. [2022b]. Information algebras in the theory of imprecise probabilities, an extension, *International Journal of Approximate Reasoning* **150**: 311–336.

Casanova, A., Miranda, E. and Zaffalon, M. [2020]. Social pooling of beliefs and values with desirability, *The Thirty-Third International Flairs Conference*, AAAI, pp. 569–574.

Casanova, A., Miranda, E. and Zaffalon, M. [2021]. Joint desirability foundations of social choice and opinion pooling, *Annals of Mathematics and Artificial Intelligence* **89**: 965–1011.

Chapman, B. [2002]. Rational aggregation, *Politics, Philosophy & Economics* **1**(3): 337–354.

Chevaleyre, Y., Endriss, U., Lang, J. and Maudet, M. [2007]. A short introduction to computational social choice, *in* J. van Leeuwen, G. F. Italiano, v. W., C. Meinel, H. Sack and F. Plasil (eds), *SOFSEM 2007: Theory and Practice of Computer Science, 33rd Conference on Current Trends in Theory and Practice of Computer Science, Proceedings*, Vol. 4362 of *Lecture Notes in Computer Science*, Springer, pp. 51–69.

Chew, S. H. and MacCrimmon, K. [1979a]. Alpha utility theory: a generalization of expected utility theory, *Faculty of Commerce and Business Administration, Working Paper* **669**.

Chew, S. H. and MacCrimmon, K. [1979b]. Alpha utility theory, lottery composition, and the Allais paradox, *Faculty of Commerce and Business Administration, Working Paper* **686**.

De Cooman, G. [2005]. Belief models: an order-theoretic investigation, *Annals of Mathematics and Artificial Intelligence* **45**(1): 5–34.

De Cooman, G., Miranda, E. and Zaffalon, M. [2011]. Independent natural extension, *Artificial Intelligence* **175**(12-13): 1911–1950.

De Cooman, G. and Quaeghebeur, E. [2012]. Exchangeability and sets of desirable gambles, *International Journal of Approximate Reasoning* **53**: 363–395.

De Cooman, G., van Camp, A. and de Bock, J. [2023]. The logic behind desirable sets of things, and its filter representation, *arXiv preprint arXiv:2302.08176* .

Davey, B. A. and Priestley, H. A. [2002]. *Introduction to Lattices and Order*, Cambridge University Press.

Dawid, A. P. [1979]. Conditional independence in statistical theory, *Journal of the Royal Statistical Society: Series B (Methodological)* **41**(1): 1–15.

Dawid, A. P. [2001]. Separoids: a mathematical framework for conditional independence and irrelevance, *Annals of Mathematics and Artificial Intelligence* **32**(1–4): 335–372.

Dietrich, F. [2006]. Judgment aggregation:(im) possibility theorems, *Journal of Economic Theory* **126**(1): 286–298.

Dietrich, F. [2010]. Bayesian group belief, *Social Choice and Welfare* **35**(4): 595–626.

Feldman, A. M. and Serrano, R. [2006]. *Welfare Economics and Social Choice Theory*, Springer Science & Business Media.

De Finetti, B. [1937]. La prévision: ses lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré*, Vol. 7, pp. 1–68.

Föllmer, H. and Schied, A. [2002]. Robust preferences and convex measures of risk, *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*, Springer, pp. 39–56.

Fréchet, M. [1951]. Sur les tableaux de corrélation dont les marges sont données, *Ann. Univ. Lyon, Section A, Series 3* **14**: 53–77.

Galaabaatar, T. and Karni, E. [2013]. Subjective expected utility with incomplete preferences, *Econometrica* **81**(1): 255–284.

Grätzer, G. [2003]. *General Lattice Theory*, Birkhäuser Verlag.

Greer, R. [1984]. Chapter 2: A tutorial on polyhedral convex cones, *in* R. Greer (ed.), *Trees and Hills: Methodology for Maximizing Functions of Systems of Linear Relations*, Vol. 96 of *North-Holland Mathematics Studies*, North-Holland, pp. 15 – 81.

Grossi, D. and Pigozzi, G. [2014]. Judgment aggregation: a primer, *Synthesis Lectures on Artificial Intelligence and Machine Learning* **8**(2): 1–151.

Hodges, W. [1993]. *Model Theory*, Cambridge University Press.

Hodges, W. [1997]. *A shorter model theory*, Cambridge University Press.

Hodges, W. [2001]. *Logic*, Penguin UK.

Jansana, R. [2022]. Algebraic Propositional Logic, *in* E. N. Zalta and U. Nodelman (eds), *The Stanford Encyclopedia of Philosophy*, Winter 2022 edn, Metaphysics Research Lab, Stanford University.

Kahneman, D. and Tversky, A. [2013]. Prospect theory: an analysis of decision under risk, *Handbook of the Fundamentals of Financial Decision Making: Part I*, World Scientific, pp. 99–127.

Kohlas, J. [2003]. *Information Algebras: Generic Structures for Inference*, Springer-Verlag.

Kohlas, J. [2017]. Algebras of information. A new and extended axiomatic foundation, *arXiv preprint arXiv:1701.02658* .

Kohlas, J., Casanova, A. and Zaffalon, M. [2021]. Information algebras of coherent sets of gambles in general possibility spaces, *ISIPTA 2021: Proceedings of International Symposium on Imprecise Probabilities: Theories and Applications*, Vol. 147, PMLR, pp. 191–200.

Kohlas, J. and Schmid, J. [2014]. An algebraic theory of information: An introduction and survey, *Information* **5**(2): 219–254.

Kohlas, J. and Schmid, J. [2020]. Commutative information algebras: representation and duality theory, *arXiv preprint arXiv:2012.15090.* .

Konieczny, S. and Pérez, R. [1999]. Merging with integrity constraints, *Symbolic and Quantitative Approaches to Reasoning and Uncertainty 5*, Springer, pp. 233–244.

Konieczny, S. and Pérez, R. [2002]. Merging information under constraints: a logical framework, *Journal of Logic and Computation* **12**(5): 773–808.

Kornhauser, L. A. and Sager, L. G. [1986]. Unpacking the court, *Yale LJ* **96**: 82.

Lauritzen, S. L. and Spiegelhalter, D. J. [1988]. Local computations with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society: Series B (Methodological)* **50**(2): 157–194.

Lindley, D. V., Tversky, A. and Brown, R. V. [1979]. On the reconciliation of probability assessments, *Journal of the Royal Statistical Society, Series A* **142**(2): 146–162.

List, C. and Pettit, P. [2002]. Aggregating sets of judgments: An impossibility result, *Economics & Philosophy* **18**(1): 89–110.

Miranda, E. and Zaffalon, M. [2010]. Notes on desirability and conditional lower previsions, *Annals of Mathematics and Artificial Intelligence* **60**(3): 251–309.

Miranda, E. and Zaffalon, M. [2020]. Compatibility, desirability, and the running intersection property, *Artificial Intelligence* **283**: 103274.

Miranda, E. and Zaffalon, M. [2023]. Nonlinear desirability theory, *International Journal of Approximate Reasoning* **154**: 176–199.

Miranda, E., Zaffalon, M. and de Cooman, G. [2012]. Conglomerable natural extension, *International Journal of Approximate Reasoning* **53**(8): 1200–1227.

Nau, R. F. [1992]. Indeterminate probabilities on finite sets, *The Annals of Statistics* **20**: 1737–1767.

Nau, R. F. [2003]. A generalization of Pratt-Arrow measure to nonexpected-utility preferences and inseparable probability and utility, *Management Science* **49**(8): 1089–1104.

Nau, R. F. [2006]. The shape of incomplete preferences, *The Annals of Statistics* **34**(5): 2430–2448.

Von Neumann, J. and Morgestern, O. [1947]. *Theory of Games and Economic Behaviour*, Princeton University Press.

Oikonomou, K. N. and Grünwald, P. D. [2015]. Explicit bounds for entropy concentration under linear constraints, *IEEE Transactions on Information Theory* **62**(3): 1206–1230.

Pelessoni, R. and Vicig, P. [2003]. Imprecise previsions for risk measurement, *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* **11**(04): 393–412.

Pelessoni, R. and Vicig, P. [2005]. Uncertainty modelling and conditioning with convex imprecise previsions, *International Journal of Approximate Reasoning* **39**(2–3): 297–319.

Pelessoni, R. and Vicig, P. [2016]. 2-coherent and 2-convex conditional lower previsions, *International Journal of Approximate Reasoning* **77**: 66–86.

Piga, D., Bemporad, A. and Benavoli, A. [2020]. Rao-Blackwellized sampling for batch and recursive Bayesian inference of Piecewise Affine models, *Automatica* **117**: 109002.

Pini, M. S., Rossi, F., Venable, K. B. and Walsh, T. [2008]. Aggregating partially ordered preferences, *Journal of Logic and Computation* **19**(3): 475–502.

Quaeghebeur, E. [2014]. *in* T. Augustin, F. Coolen, G. de Cooman and M. Troffaes (eds), *Introduction to Imprecise Probabilities*, Wiley, pp. 1–27.

Quaeghebeur, E., de Cooman, G. and Hermans, F. [2015]. Accept & reject statement-based uncertainty models, *International Journal of Approximate Reasoning* **57**: 69–102.

Quiggin, J. [1982]. A theory of anticipated utility, *Journal of Economic Behavior & Organization* **3**(4): 323–343.

Rockafellar, R. T. [1970]. *Convex Analysis*, Vol. 36, Princeton University Press.

Savage, L. J. [1972]. *The Foundations of Statistics*, Dover, New York. Second revised edition, first published 1954.

Schrijver, A. [1998]. *Theory of Linear and Integer Programming*, John Wiley & Sons.

Seidenfeld, T., Kadane, J. B. and Schervish, M. J. [1989]. On the shared preferences of two Bayesian decision makers, *The Journal of Philosophy* **86**(5): 225–244.

Seidenfeld, T., Schervish, M. J. and Kadane, J. B. [1995]. A representation of partially ordered preferences, *The Annals of Statistics* **23**: 2168–2217. Reprinted in Seidenfeld et al. [1999], pp. 69–129.

Seidenfeld, T., Schervish, M. J. and Kadane, J. B. [1999]. *Rethinking the Foundations of Statistics*, Cambridge University Press, Cambridge.

Sen, A. K. [1970]. *Collective Choice and Social Welfare*, Elsevier.

Shafer, G., Shenoy, P. P. and Mellouli, K. [1987]. Propagating belief functions in qualitative Markov trees, *International Journal of Approximate Reasoning* **1**(4): 349–400.

Shenoy, P. P. and Shafer, G. [1990]. Axioms for probability and belief-function propagation, *in* R. D. Shachter, T. S. Levitt, L. N. Kanal and J. F. Lemmer (eds), *Uncertainty in Artificial Intelligence*, Vol. 9 of *Machine Intelligence and Pattern Recognition*, North-Holland, pp. 169–198.

Spohn, W. [1980]. Stochastic independence, causal independence, and shieldability, *Journal of Philosophical Logic* **9**: 73–99.

Steen, L. A., Seebach, J. A. and Steen, L. A. [1978]. *Counterexamples in Topology*, Vol. 18, Springer.

Stewart, R. and Quintana, I. O. [2018]. Probabilistic opinion pooling with imprecise probabilities, *Journal of Philosophical Logic* **47**(1): 17–45.

Studenỳ, M. [1995]. Marginal problem in different calculi of AI, *Advances in Intelligent Computing—IPMU'94: 5th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems Paris, France, July 4–8, 1994 Selected Papers 5*, Springer, pp. 348–359.

Tarski, A. [1983]. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, Hackett Publishing.

Troffaes, M. C. and de Cooman, G. [2014]. *Lower Previsions*, John Wiley & Sons.

Vorob'ev, N. N. [1962]. Consistent families of measures and their extensions, *Theory of Probability & its Applications* **7**(2): 147–163.

Walley, P. [1982]. The elicitation and aggregation of beliefs, *Technical report*, University of Warwick, Coventry. Statistics Research Report 23.

Walley, P. [1991]. *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London.

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

Weymark, J. A. [1984]. Arrow's theorem with social quasi-orderings, *Public Choice* **42**(3): 235–246.

Williams, P. M. [1975]. Notes on conditional previsions, *Technical report*, School of Mathematical and Physical Science, University of Sussex, UK. Reprinted in Williams [2007].

Williams, P. M. [2007]. Notes on conditional previsions, *International Journal of Approximate Reasoning* **44**: 366–383. Revised journal version of Williams [1975].

Wilson, N. and Mengin, J. [1999]. Logical deduction using the local computation framework, *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Springer, pp. 386–396.

Wilson, N. and Moral, S. [1994]. A logical view of probability, *ECAI*, Vol. 94, Citeseer, pp. 386–390.

Zaffalon, M. and Miranda, E. [2017]. Axiomatising incomplete preferences through sets of desirable gambles, *Journal of Artificial Intelligence Research* **60**: 1057–1126.

Zaffalon, M. and Miranda, E. [2021]. Desirability foundations of robust rational decision making, *Synthese* **198**(27): 6529–6570.