

Sum-Of-Squares for bounded rationality

Alessio Benavoli, Alessandro Facchini, Dario Piga, Marco Zaffalon

Istituto Dalle Molle di Studi Sull'Intelligenza Artificiale (IDSIA), Lugano (Switzerland)

Abstract

In the gambling foundation of probability theory, rationality requires that a subject should always (never) find desirable all nonnegative (negative) gambles, because no matter the result of the experiment the subject never (always) decreases her money. Evaluating the nonnegativity of a gamble in infinite spaces is a difficult task. In fact, even if we restrict the gambles to be polynomials in \mathbb{R}^n , the problem of determining nonnegativity is NP-hard. The aim of this paper is to develop a *computable theory of desirable gambles*. Instead of requiring the subject to desire all nonnegative gambles, we only require her to desire gambles for which she can efficiently determine the nonnegativity (in particular sum-of-squares polynomials). We refer to this new criterion as *bounded rationality*.

Keywords: Bounded rationality, theory of desirable gambles, sum-of-squares polynomials, polynomial gambles, updating.

1. Introduction

The subjective foundation of probability by [de Finetti \(1937\)](#) is based on the notion of rationality (coherence or equiv. self-consistency). A subject is considered rational if she chooses her odds so that there is no bet that leads her to a sure loss (no Dutch books are possible). In this way, since odds are the inverse of probabilities, de Finetti provided a justification of Kolmogorov's axiomatisation of probability as a rationality criterion on a gambling system.¹

Later [Williams \(1975\)](#) and [Walley \(1991\)](#) showed that it is possible to justify probability in a simpler and more elegant way. This approach is nowadays known as the *theory of desirable gambles*.² To understand this gambling framework, we introduce a subject, Alice, and an experiment whose result ω belongs to a possibility space Ω (e.g., the experiment may be tossing a coin or determining the future value of a derivative instrument). When Alice is uncertain about the result ω of the experiment, we can model her beliefs about this value by asking her whether she accepts engaging in certain risky transactions, called *gambles*, whose outcome depends on the actual outcome of the experiment ω . Mathematically, a gamble is a bounded real-valued function on Ω , $g : \Omega \rightarrow \mathbb{R}$, and if Alice accepts a gamble g , this means that she commits herself to receive $g(\omega)$ utiles³ if the experiment is performed and if the outcome of the experiment eventually happens to be the event $\omega \in \Omega$. Since $g(\omega)$ can be negative, Alice can also lose utiles and hence the desirability of a gamble depends on Alice's beliefs about Ω . Denote by \mathcal{L} the set of all the gambles on Ω . Alice examines gambles in \mathcal{L} and comes up with the subset \mathcal{K} of the gambles that she finds desirable. How can we characterise the rationality of the assessments represented by \mathcal{K} ?

Two obvious rationality criteria are: Alice should always accept (do not accept) gambles that are nonnegative (negative), because no matter the result of the experiment she never (always) decreases her utiles. But there is a world of difference between saying and doing. For instance, let us consider an infinite space of possibilities like $\Omega = \mathbb{R}^2$ and the gamble: $g(x_1, x_2) = 4x_1^4 + 4x_1^3x_2 - 3x_1^2x_2^2 + 5x_2^4$. Should Alice accept this gamble? In practice the answer to this question does not only depend on Alice's beliefs about the value of x_1 and x_2 . We can in fact verify that the

^{*}Fully documented templates are available in the elsarticle package on [CTAN](#).

¹De Finetti actually considered only finitely additive probabilities, while σ -additivity is assumed in Kolmogorov's axiomatisation.

²In this paper, we will refer in particular to the theory of almost desirable gambles.

³A theoretical unit of measure of utility, for indicating a supposed quantity of satisfaction derived from an economic transaction. It is expressed in some linear utility scale

above polynomial can be rewritten as $(2x_1^2 - 2x_2^2 + x_1x_2)^2 + (x_2^2 + 2x_1x_2)^2$ and, thus, is always nonnegative. Hence, rationality implies that Alice should always accept it. However, in these cases, we must also take into account the inherent difficulty of the problem faced by Alice when she wants to determine whether a given gamble is nonnegative or not. In other words, we need to quantify the computational complexity needed to address rationality.

The aim of this paper is to develop a *computable theory of desirable gambles* by relaxing the rationality criteria discussed above. In particular, instead of requiring Alice to accept all nonnegative gambles, we only require Alice to accept gambles for which she can efficiently determine the nonnegativity. We call this new criterion *bounded rationality*. The term bounded rationality was proposed by Herbert A. [Simon \(1957\)](#) – it is the idea that when individuals make decisions, their rationality is limited by the tractability of the decision problem, the cognitive limitations of their minds, and the time available to make the decision. Decision-makers in this view act as “satisficers”, seeking a satisfactory solution rather than an optimal one. We do not propose our model as a realistic psychological model of Alice’s behaviour, but we embrace the idea that the actual rationality of an agent is determined by its computational intelligence.

In this paper, we exploit the results on SOS polynomials and theory-of-moments relaxation to make numerical inferences in our theory of bounded rationality and to show that the theory of bounded rationality can be used to approximate the theory of desirable gambles. At the same time, we provide a gambling interpretation of SOS optimization. Some applications of the theoretical ideas presented in this paper can be found in [Lasserre \(2009\)](#); [Benavoli & Piga \(2016\)](#); [Piga & Benavoli \(2017\)](#). For instance, [Benavoli & Piga \(2016\)](#) use this approach to derive a novel set-membership filtering algorithm for nonlinear polynomial dynamical systems. Although their approach is not directly formulated as a theory of bounded rationality, SOS polynomials are used to propagate a set of probability measures in a computationally efficient way through the dynamics of a nonlinear system. It is worth mentioning that a relaxation of the rationality criteria for desirability has also been investigated by [Schervish et al. \(2000\)](#); [Pelessoni & Vicig \(2016\)](#). The first work focuses on relaxations of the “avoiding sure loss” axiom, while the second one focuses on two different criteria (additivity and positive scaling).

A preliminary version of this work appeared in ([Benavoli et al., 2017a](#)), but it includes an incorrect statement of duality. This has led us to re-evaluate the whole theory, resulting in a new definition of bounded rationality that we will present in the current manuscript.

2. Theory of desirable gambles

In this section, we briefly introduce the theory of desirable gambles. Let us denote by $\mathcal{L}^+ = \{g \in \mathcal{L} : g \geq 0\}$ the subset of the *nonnegative gambles* and with $\mathcal{K} \subset \mathcal{L}$ the subset of the gambles that Alice finds desirable. How can we characterise the rationality of the assessments in \mathcal{K} ?

Definition 1. We say that \mathcal{K} is a coherent set of (almost) desirable gambles (ADG) when it satisfies the following rationality criteria:

- A.1** If $g \in \mathcal{L}^+$ then $g \in \mathcal{K}$ (Accepting Partial Gains);
- A.2** If $g \in \mathcal{K}$ then $\sup g \geq 0$ (Avoiding Sure Loss);
- A.3** If $g \in \mathcal{K}$ then $\lambda g \in \mathcal{K}$ for every $\lambda > 0$ (Positive Scaling);
- A.4** If $g, h \in \mathcal{K}$ then $g + h \in \mathcal{K}$ (Additivity);
- A.5** If $g + \delta \in \mathcal{K}$ for every $\delta > 0$ then $g \in \mathcal{K}$ (Closure).

The criterion A.5 does not actually follow from rationality and can be omitted ([Seidenfeld et al., 1990](#); [Walley, 1991](#); [Miranda & Zaffalon, 2010](#)). However, it is useful to derive a connection between the theory of desirable gambles and probability theory and for this reason we consider it in this paper. This connection will be briefly discussed in Section 3.

To explain these rationality criteria, let us introduce a simple example: the toss of a fair coin $\Omega = \{Head, Tail\}$. A gamble g in this case has two components $g(Head) = g_1$ and $g(Tail) = g_2$. If Alice accepts g then she commits

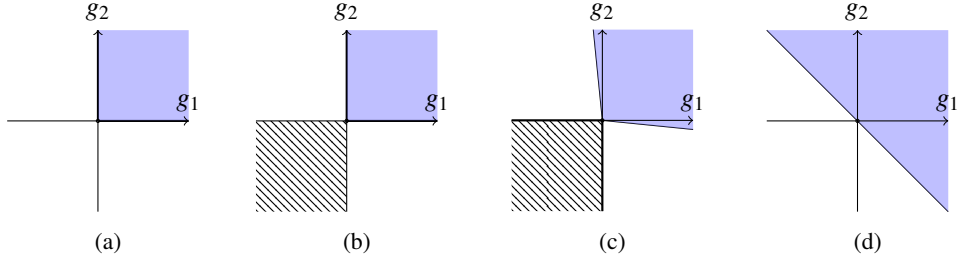


Figure 1: Alice’s sets of coherent almost desirable gambles for the experiment of tossing a fair coin.

herself to receive/pay g_1 if the outcome is Heads and g_2 if Tails. Since a gamble is in this case an element of \mathbb{R}^2 , $g = (g_1, g_2)$, we can plot the gambles Alice accepts in a 2D coordinate system with coordinate g_1 and g_2 .

A.1 says that Alice is willing to accept any gamble $g = (g_1, g_2)$ that, no matter the result of the experiment, may increase her wealth without ever decreasing it, that is with $g_i \geq 0$ – Alice always accepts the first quadrant, Figure 1(a). Similarly, Alice does not accept any gamble $g = (g_1, g_2)$ that will surely decrease her wealth, that is with $g_i < 0$. In other words, Alice always does not accept the interior of the third quadrant, Figure 1(b). This is the meaning of A.2. Then we ask Alice about $g = (-0.1, 1)$ – she loses 0.1 if Heads and wins 1 if Tails. Since Alice knows that the coin is fair, she accepts this gamble as well as all the gambles of the form λg with $\lambda > 0$, because this is just a “change of currency” (this is A.3). Similarly, she accepts all the gambles $g + h$ for any $h \in \mathcal{L}^+$, since these gambles are even more favourable for her (this is basically A.4). Now, we can ask Alice about $g = (1, -0.1)$ and the argument is symmetric to the above case. We therefore obtain the following set of desirable gambles (see Figure 1(c)): $\mathcal{K}_2 = \{g \in \mathbb{R}^2 \mid 10g_1 + g_2 \geq 0 \text{ and } g_1 + 10g_2 \geq 0\}$. Finally, we can ask Alice about $g = (-1, 1)$ – she loses 1 if Heads and wins 1 if Tails. Since the coin is fair, Alice may accept or not accept this gamble. A.5 implies that she must accept it (closure). A similar conclusion can be derived for the symmetric gamble $g = (1, -1)$. Figure 1(d) is her final set of desirable gambles about the experiment concerned with the toss of a fair coin, which in a formula becomes $\mathcal{K}_3 = \{g \in \mathbb{R}^2 \mid g_1 + g_2 \geq 0\}$. Alice does not accept any other gamble. In fact, if Alice would also accept for instance $h = (-2, 0.5)$ then, since she has also accepted $g = (1.5, -1)$, i.e., $g \in \mathcal{K}_3$, she must also accept $g + h$ (because this gamble is also favourable to her). However, $g + h = (-0.5, -0.5)$ is always negative, Alice always loses utiles in this case. In other words, by accepting $h = (-2, 0.5)$ Alice incurs a sure loss – she is irrational (A.2 is violated).

In this example, we can see that Alice’s set of desirable gambles is a closed half-space, but this does not have to be always the case. For instance, if Alice does not know anything about the coin, she should only accept nonnegative gambles: $\mathcal{K} = \mathcal{L}^+$. This corresponds to a state of complete ignorance, but all intermediate cases from complete beliefs on the probability of the coin to complete ignorance are possible. In general, \mathcal{K} is a pointed (whose vertex is the origin) closed convex cone that includes \mathcal{L}^+ and excludes the interior of the negative orthant (this follows by A.1–A.5).

For the coin, the space of possibilities is finite and in this case Alice can check if a gamble g is nonnegative by simply examining the elements of the vector g . In this paper, we are interested in infinite spaces, in particular $\Omega = \mathbb{R}^n$, where applying the above rationality criteria is far from easy. We aim at developing a theory of *bounded rationality* for this case. Before doing that, we briefly recall the connection between ADG and probability theory.

3. Duality for ADG

Duality can be defined for general spaces of possibilities Ω (Walley, 1991). However, for the purpose of the present paper, we consider gambles that are bounded real-valued function on \mathbb{R}^n , i.e., $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Let \mathcal{A} be an algebra of subsets of \mathbb{R}^n and $\mu : \mathcal{A} \rightarrow [-\infty, \infty]$ denotes a charge: that is μ is a finitely additive set function of \mathcal{A} (Aliprantis & Border, 2007, Ch.11). Let $\mathcal{A}_{\mathbb{R}}$ denote the algebra generated in \mathbb{R} by the collection of all half open intervals (Aliprantis & Border, 2007, Th.11.8):

Theorem 1. *Every bounded $(\mathcal{A}, \mathcal{A}_{\mathbb{R}})$ -measurable function is integrable w.r.t. any finite charge.*

Therefore, for any bounded $(\mathcal{A}, \mathcal{A}_{\mathbb{R}})$ -measurable function g and finite charge μ we can define $\int g d\mu$, which we can interpret as a linear functional $L(\cdot) := \langle \cdot, \mu \rangle$ on bounded $(\mathcal{A}, \mathcal{A}_{\mathbb{R}})$ -measurable gambles g (with some abuse of notation, we denote the set of bounded $(\mathcal{A}, \mathcal{A}_{\mathbb{R}})$ -measurable functions with \mathcal{L}). We denote by \mathcal{M} the set of all finite charges on \mathcal{L} and by \mathcal{M}^+ the set of nonnegative charges. A linear functional of gambles is said to be *nonnegative* whenever it satisfies $L(g) \geq 0$, for $g \in \mathcal{L}^+$. A nonnegative linear functional is called a *state* if moreover it preserves the unitary constant gamble. Hence, in this context, since $\langle 1, \mu \rangle = \int 1 d\mu = 1$, the set of states \mathcal{S} corresponds to the set of all probability charges. We define the *dual* of a subset \mathcal{D} of \mathcal{L} as:

$$\mathcal{D}^\bullet = \left\{ \mu \in \mathcal{M} : \int g d\mu \geq 0, \forall g \in \mathcal{D} \right\}. \quad (1)$$

Similarly, the dual of a subset \mathcal{R} of \mathcal{M} is the set:

$$\mathcal{R}^\bullet = \left\{ g \in \mathcal{L} : \int g d\mu \geq 0, \forall \mu \in \mathcal{R} \right\}. \quad (2)$$

Note that in both cases $(\cdot)^\bullet$ is always a closed convex cone (Aliprantis & Border, 2007, Lem.5.102(4)). Furthermore, one has that $(\cdot)^{\bullet\bullet} = (\cdot)$, whenever (\cdot) is a closed convex cone (Aliprantis & Border, 2007, Th.5.103), and $(\cdot)_1 \subseteq (\cdot)_2$ if and only if $(\cdot)_2^\bullet \subseteq (\cdot)_1^\bullet$ (Aliprantis & Border, 2007, Lem.5.102(1)). In particular, whenever $(\cdot)_1$ and $(\cdot)_2$ are closed convex cones, $(\cdot)_1 \subsetneq (\cdot)_2$ if and only if $(\cdot)_2^\bullet \subsetneq (\cdot)_1^\bullet$.

Based on those facts, it is thus possible to verify that the dual of a coherent set of desirable gambles can actually be completely described in terms of a (closed convex) set of states (probability charges). In this aim, we start by the following observations.

Proposition 1. *It holds that*

1. $(\mathcal{L})^\bullet = \{0\}$ and $\mathcal{L} = (\{0\})^\bullet$;
2. $(\mathcal{L}^+)^\bullet = \mathcal{M}^+$ and $\mathcal{L}^+ = (\mathcal{M}^+)^\bullet$;

Proof. Since $(\cdot)^{\bullet\bullet} = (\cdot)$, whenever (\cdot) is a closed convex cone, in both cases it is enough to verify only one of the claims. For the first item, the second claim is immediate. For the second item, we verify the first claim. The inclusion from right to left being clear, for the other direction observe that: (i) $g = I_{\{B\}}$ (with I_B being the indicator function on $B \in \mathcal{A}$), is a nonnegative gamble and, therefore, is in \mathcal{L}^+ ; (ii) if μ is negative in $B \in \mathcal{A}$, i.e., then $\int I_B d\mu$ is negative too and, thus, μ cannot be in $(\mathcal{L}^+)^\bullet$. \square

Proposition 2. *Let \mathcal{K} be a closed convex cone. The following claims are equivalent*

1. \mathcal{K} is coherent;
2. $\mathcal{K} \supseteq \mathcal{L}^+$ and $\mathcal{K} \neq \mathcal{L}$;
3. $(\mathcal{K})^\bullet \subseteq \mathcal{M}^+$ and $(\mathcal{K})^\bullet \neq \{0\}$.

Proof. (2) \Leftrightarrow (3): From Proposition 1, $(\mathcal{K})^\bullet = \{0\}$ if and only if $\mathcal{K} = \mathcal{L}$, and $\mathcal{K} \supseteq \mathcal{L}^+$ if and only if $(\mathcal{K})^\bullet \subseteq \mathcal{M}^+$. (1) \Rightarrow (2): Assume \mathcal{K} is coherent. By A.1 $\mathcal{K} \supseteq \mathcal{L}^+$ and by A.2 there is $g \in \mathcal{L}$ such that $\sup g < 0$ and $g \notin \mathcal{K}$. (2) \Rightarrow (1): Let $\mathcal{L}^+ \subseteq \mathcal{K} \subsetneq \mathcal{L}$. First of all, notice that, by Proposition 1, $(\mathcal{K})^\bullet \subseteq (\mathcal{L}^+)^\bullet = \mathcal{M}^+$. Now, assume that \mathcal{K} is not coherent. This means A.2 fails, i.e. there is $g \in \mathcal{L}$ such that $\sup g < 0$ and $g \in \mathcal{K}$. Hence, consider $\mu \in \mathcal{M}^+$ and pick $g \in \mathcal{K}$ such that $\sup g < 0$. It holds that $\langle g, \mu \rangle \geq 0$ if and only if $\mu = 0$, meaning that $(\mathcal{K})^\bullet = \{0\}$ and therefore, by Proposition 1 again, $\mathcal{K} = \mathcal{L}$, a contradiction. \square

Hence, the following theorem holds.

Theorem 2. *The map*

$$\mathcal{K} \mapsto \mathcal{P} := \mathcal{K}^\bullet \cap \mathcal{S}$$

establishes a bijection between coherent sets of desirable gambles and closed convex sets of states.

Proof. The proof is analogous to that by [Benavoli et al. \(2017b, Th.4\)](#). Let \mathcal{K} be a coherent set of desirable gambles. By Proposition 2, we get that \mathcal{K}^\bullet is a closed convex cone included in \mathcal{M}^+ that does not reduce to the origin. Thus, after normalisation, \mathcal{P} is nonempty. Preservation of closedness and convexity by finite intersections yields that \mathcal{P} is closed and convex. Furthermore $\mathbb{R}_+\mathcal{P} = \mathcal{K}^\bullet$, and therefore $\mathcal{K} = (\mathbb{R}_+\mathcal{P})^\bullet$, where $\mathbb{R}_+\mathcal{P} := \{\lambda\mu : \lambda \geq 0, \mu \in \mathcal{P}\}$, meaning that the map is an injection. We finally verify that the map is also a surjection. To do this, let \mathcal{P} be a non empty closed convex set of probability charges. It holds that $\mathbb{R}_+\mathcal{P}$ is a closed convex cone included in \mathcal{M}^+ different from $\{0\}$. Again by Proposition 2, we conclude that the dual $(\mathbb{R}_+\mathcal{P})^\bullet$ of $\mathbb{R}_+\mathcal{P}$ is a coherent set of desirable gambles and $\mathcal{P} = \mathbb{R}_+\mathcal{P} \cap \mathcal{S} = (\mathbb{R}_+\mathcal{P})^{\bullet\bullet} \cap \mathcal{S}$. \square

We therefore identify the dual of \mathcal{K} with

$$\mathcal{P} = \left\{ \mu \in \mathcal{M}^+ : \int g d\mu \geq 0, \int d\mu = 1, \forall g \in \mathcal{K} \right\}, \quad (3)$$

which is a closed convex-set of probability charges.

3.1. Unbounded gambles

In this paper, we will also consider unbounded real-valued functions in \mathbb{R}^n and, therefore, we need to introduce another definition of measurability. Let \mathcal{M} denote the space of finite signed Borel measures on \mathbb{R}^n , whose positive cone \mathcal{M}^+ is the space of finite Borel measures μ on \mathbb{R}^n . Let \mathcal{L} be the set of real-valued functions on \mathbb{R}^n that are integrable with respect to every measure $\mu \in \mathcal{M}^+$ and \mathcal{L}^+ be the cone of non-negative integrable functions. With an abuse of terminology, we call these functions gambles.⁴ Also in this case, we can define a cone of desirable gambles in \mathcal{L} satisfying the properties A.1–A.5. By defining the set of states as $\mathcal{S} := \{\mu \in \mathcal{M} : \int d\mu = 1\}$, we can prove the following.

Theorem 3. *The map*

$$\mathcal{K} \mapsto \mathcal{P} := \mathcal{K}^\bullet \cap \mathcal{S}$$

establishes a bijection between coherent sets of desirable gambles and closed convex sets of states.

The proof is identical to that for Theorem 2. Here, \mathcal{P} is the set of all probability measures on \mathbb{R}^n . In the sequel, we will refer to this duality when we will consider unbounded gambles.

4. Finite assessments

The goal of this and next sections is to define a practical notion of desirability. To this end, we first assume that the set of gambles that Alice finds desirable is finitely generated. By this, we mean that there is a finite set of gambles $G = \{g_1, \dots, g_{|G|}\}$ such that $\mathcal{K} = \text{posi}(G \cup \mathcal{L}^+)$, where the posi of a set $A \subset \mathcal{L}$ is defined as

$$\text{posi}(A) := \left\{ \sum_{j=1}^{|A|} \lambda_j g_j : g_j \in A, \lambda_j \geq 0 \right\}, \quad (4)$$

and where by $|A|$ we denote the cardinality of the set A . By using this definition, it is clear that whenever \mathcal{K} is finitely generated, it includes all nonnegative gambles and satisfies A.3, A.4 and A.5. Once Alice has defined G and so \mathcal{K} via posi, ADG assumes that she is able to perform the following operations: to check that \mathcal{K} avoids sure loss (A.2 is also satisfied); to determine the implication of desirability. It is easy to show that all above operations in ADG imply the assessment of the nonnegativity of a gamble.

Proposition 3. *Given a finite set $G \subset \mathcal{L}$ of desirable gambles, the set $\text{posi}(G \cup \mathcal{L}^+)$ includes the gamble f if and only if there exist $\lambda_j \geq 0$ for $j = 1, \dots, |G|$ such that*

$$f - \sum_{j=1}^{|G|} \lambda_j g_j \geq 0. \quad (5)$$

⁴For a more general extension of the theory of desirable gambles to unbounded gambles see [Troffaes & de Cooman \(2003\)](#).

There are two subcases of (5) that are particularly interesting. The first is when $f = h - \lambda_0$ for some $\lambda_0 \in \mathbb{R}$. That allows us to define the concept of lower prevision (Walley, 1991; Miranda, 2008).

Definition 2. Assume that $\mathcal{K} = \text{posi}(G \cup \mathcal{L}^+)$ is an ADG, then the solution of the following problem

$$\sup_{\lambda_0 \in \mathbb{R}, \lambda_j \geq 0} \lambda_0, \quad \text{s.t.} \quad h - \lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j \geq 0, \quad (6)$$

is called the lower prevision of h and denoted as $\underline{E}[h]$.

From a behavioural point of view, we can reinterpret this by saying that Alice is willing to buy gamble h at price λ_0 , since she is giving away λ_0 utiles while gaining h . The lower prevision is the supremum buying price for h . We can equivalently define the upper prevision of h as $\bar{E}[h] = -\underline{E}[-h]$. From Section 3, it can be easily shown that $\underline{E}[h]$ is the lower expectation of h computed w.r.t. the probability charges (or measures if we consider the case in Section 3.1) in \mathcal{P} . As a matter of fact, the dual of (6) is the moment problem:

$$\inf_{\mu \in \mathcal{M}^+} \int h d\mu \quad \text{s.t.} \quad \int d\mu = 1, \quad \int g_j d\mu \geq 0 \text{ for } j = 1, \dots, |G|. \quad (7)$$

Example 1 (Markov's inequality). Consider the nonnegative function x on $\Omega = [0, x_{\max}]$ with $x_{\max} \in \mathbb{R}^+$ and assume that Alice finds the gambles $g = x - m$ and $-g = -x + m$ to be desirable for some $m \in \Omega$, i.e.,

$$G = \{g, -g\}.$$

Consider the event $x \geq u$ for some $u \in \mathbb{R}^+$. Our goal is to determine Alice's upper prevision (infimum selling price) for this event, equivalently for the gamble $I_{\{[u, \infty)\}}$. We need to apply (6) which in this case can be written as:

$$\begin{aligned} \bar{E}(I_{\{[u, \infty)\}}) &= \inf_{\lambda_{1j} \leq 0, \lambda_0 \in \mathbb{R}} \lambda_0 \\ \text{s.t.} \\ \lambda_0 + \lambda_{11}(x - m) - \lambda_{12}(x - m) &\geq I_{\{[u, \infty)\}}(x), \quad \forall x \in \Omega. \end{aligned} \quad (8)$$

By defining $\lambda_1 = \lambda_{11} - \lambda_{12}$, which now spans \mathbb{R} , we can rewrite (8) as:

$$\begin{aligned} \bar{E}(I_{\{[u, \infty)\}}) &= \inf_{\lambda_j \in \mathbb{R}} \lambda_0 \\ \text{s.t.} \\ \lambda_0 + \lambda_1(x - m) &\geq 1, \quad \forall x \in \Omega : x \geq u, \\ \lambda_0 + \lambda_1(x - m) &\geq 0, \quad \forall x \in \Omega : x < u. \end{aligned} \quad (9)$$

It can be seen that $\bar{E}(I_{\{[u, \infty)\}}) = 1$ whenever $u \leq 0$ and $\bar{E}(I_{\{[u, \infty)\}}) = 0$ whenever $u \geq x_{\max}$. In the other cases, the above problem must satisfy:

$$\begin{aligned} \bar{E}(I_{\{[u, \infty)\}}) &\leq \inf_{\lambda_i \in \mathbb{R}} \lambda_0 \\ \text{s.t.} \\ \lambda_0 + \lambda_1(u - m) &\geq 1, \\ \lambda_0 + \lambda_1(-m) &\geq 0, \end{aligned} \quad (10)$$

where we have considered the worst cases for the inequalities. If $m < u$, the infimum is obtained when the inequalities in (10) are equalities and is equal to $\frac{m}{u}$. When $m \geq u$, the infimum is obtained for $\lambda_0 = 1, \lambda_1 = 0$. We have therefore derived Markov's inequality:

$$P(x \geq u) \leq \min\left(1, \frac{m}{u}\right) = \min\left(1, \frac{E[x]}{u}\right).$$

The last equality follows from the fact that (8) is equivalent to (see (7)):

$$\sup_{\mu \in \mathcal{M}^+} \int I_{\{[u, \infty)\}} d\mu \quad \text{s.t.} \quad \int d\mu = 1, \quad \int x d\mu = m, \quad (11)$$

and so m is just the expectation of x , i.e., $m = E[x] = \int x d\mu$.

The second subcase allows us to formulate sure loss as nonnegativity of a gamble (Walley et al., 2004, Alg. 2).

Proposition 4. *Let us consider $\mathcal{K} = \text{posi}(G \cup \mathcal{L}^+)$ and the following problem:*

$$\sup_{0 \leq \lambda_0 \leq 1, \lambda_j \geq 0} \lambda_0, \quad \text{s.t.} \quad -\lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j \geq 0. \quad (12)$$

\mathcal{K} incurs a sure loss iff the above problem has solution $\lambda_0^* = 1$ and avoids sure loss iff $\lambda_0^* = 0$.

Proof. We briefly sketch the proof (see Walley et al. 2004). Assume that \mathcal{K} incurs a sure loss, then there exist $\lambda_j \geq 0$ for $j = 1, \dots, |G|$ such that $\sum_{j=1}^{|G|} \lambda_j g_j < 0$ and, thus, $-\sum_{j=1}^{|G|} \lambda_j g_j > 0$. Then we can increase λ_0 as much as we want. Similarly, we can prove the inverse implication. \square

Example 2 (Markov's inequality cont.). *Consider again the previous example, i.e.,*

$$G = \{x - m, -x + m\}.$$

We want to show when/if \mathcal{K} incurs a sure loss. Consider (12) and assume $m \notin \Omega$.

$$\sup_{0 \leq \lambda_0 \leq 1, \lambda_1 \in \mathbb{R}} \lambda_0, \quad \text{s.t.} \quad -\lambda_0 - \lambda_1(x - m) \geq 0 \quad x \in \Omega. \quad (13)$$

Note that if $m < 0$ then $x - m > 0$ in Ω and so $-\lambda_1(x - m) \geq 0$ provided that $\lambda_1 \leq 0$. Hence, the solution of the above optimisation problem is $\lambda_0 = 1$ (sure loss). Similarly, if $m > x_{\max}$ then $-\lambda_1(x - m) \geq 0$ provided that $\lambda_1 \geq 0$ and so the optimum is $\lambda_0 = 1$. Assume now that $m \in \Omega$, then the best solution is obtained for $\lambda_1 = 0$ and so the optimum is $\lambda_0 = 0$ and, therefore, \mathcal{K} avoids sure loss. We can conclude that, to avoid a sure loss, Alice should accept the gambles $x - m, -x + m$ only when

$$\inf_{x \in \Omega} x = 0 \leq m \leq \sup_{x \in \Omega} x = x_{\max}.$$

In the previous example we saw that $m = E[x] = \int x d\mu$, hence the above inequalities mean that

$$\inf_{x \in \Omega} x \leq E[x] \leq \sup_{x \in \Omega} x.$$

4.1. Complexity of inferences

When Ω is finite (e.g., coin toss), then a gamble g can also be seen as a vector in $\mathbb{R}^{|\Omega|}$ (where $|\Omega| = 2$ for the coin). Then (5) can be expressed as a linear programming problem, thus its complexity is polynomial: *Alice can determine the implication of her assessments of desirability in polynomial time.* In case $\Omega = \mathbb{R}^n$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, solving (5) means to check the existence of real parameters $\lambda_j \geq 0$ ($j = 1, \dots, |G|$) such that the function

$$F := f - \sum_{j=1}^{|G|} \lambda_j g_j \quad (14)$$

is nonnegative in \mathbb{R}^n . In order to study the problem from a computational viewpoint, and avoid undecidability results, it is clear that we must impose further restrictions on the class of functions F . At the same time we would like to keep the problem general enough, in order not to lose expressiveness of the model. A good compromise can be achieved by considering the case of multivariate polynomials. The decidability of $F \geq 0$ for multivariate polynomials can be proven by means of the Tarski-Seidenberg quantifier elimination theory (Tarski, 1951; Seidenberg, 1954).

Let $d \in \mathbb{N}$. By $\mathbb{R}_{2d}[x_1]$ we denote the set of all polynomials up to degree $2d$ in the indeterminate variable $x_1 \in \mathbb{R}$ with real-valued coefficients. With the usual definitions of addition and scalar multiplication, $\mathbb{R}_{2d}[x_1]$ becomes a vector space over the field \mathbb{R} of real numbers. We can introduce a basis for $\mathbb{R}_{2d}[x_1]$ that we denote as $v_{2d}(x_1)$ where $v_j(x_1) = [1, x_1, x_1^2, \dots, x_1^j]^\top$. We denote the dimension of $v_j(x_1)$ as $s_1(j)$ for $j = 0, 1, 2, \dots$, e.g., $s_1(2d) = 2d + 1$. Any polynomial in $\mathbb{R}_{2d}[x_1]$ can then be written as $p(x_1) = b^\top v_{2d}(x_1)$ being $b \in \mathbb{R}^{s_1(2d)}$ the vector of coefficients. One

can actually provide a square matrix representation of a polynomial, as for any polynomial in $\mathbb{R}_{2d}[x_1]$ there is a (non-unique) matrix $Q \in \mathbb{R}_s^{s_1(d) \times s_1(d)}$ such that $p(x_1) = v_d^\top(x_1)Qv_d(x_1)$, where $\mathbb{R}_s^{s_1(d) \times s_1(d)}$ is the space of $s_1(d) \times s_1(d)$ real-symmetric matrices. In virtue of these observations we may therefore also be interested in some subsets of $\mathbb{R}_{2d}[x_1]$ that are:

- (1) the subset of nonnegative polynomials, denoted as $\mathbb{R}_{2d}^+[x_1]$;
- (2) the subset of polynomials

$$\Sigma_{2d}[x_1] = \left\{ p(x_1) \in \mathbb{R}_{2d}[x_1] \mid p(x_1) = v_d^\top(x_1)Qv_d(x_1) \text{ with } Q \in \mathbb{R}_s^{s_1(d) \times s_1(d)}, Q \geq 0 \right\}. \quad (15)$$

The polynomials in $\Sigma_{2d}[x_1]$ are also called SOS polynomials. This is because any polynomial in $\mathbb{R}_{2d}[x_1]$ that is a sum of squares of polynomials belongs to $\Sigma_{2d}[x_1]$ and vice versa (Lasserre, 2009, Prop. 2.1). Clearly any polynomial in $\Sigma_{2d}[x_1]$ is necessarily nonnegative. It is therefore natural to ask whether the two sets, $\mathbb{R}_{2d}^+[x_1]$ and $\Sigma_{2d}[x_1]$, coincide, and therefore whether Equation (15) provides a representation theorem for nonnegative univariate polynomial. The answer to this question is affirmative (Lasserre, 2009, Prop. 2.3).

The previous framework can be extended to multivariate polynomials $\mathbb{R}_{2d}[x_1, \dots, x_n]$. Indeed, polynomials in $\mathbb{R}_{2d}[x_1, \dots, x_n]$ can be written as $p(x_1, \dots, x_n) = b^\top v_{2d}(x_1, \dots, x_n)$ with

$$v_j(x_1, \dots, x_n) = [1, x_1, \dots, x_n, x_1^2, x_1x_2, \dots, x_{n-1}x_n, x_n^2, \dots, x_1^{2d}, \dots, x_n^j]^\top, \quad (16)$$

$b \in \mathbb{R}^{s_n(2d)}$ with $s_n(j) = \binom{n+j}{j}$ for $j = 0, 1, 2, \dots$. Moreover, one can always find a real-symmetric matrix Q such that $p(x_1, \dots, x_n) = v_d^\top(x_1, \dots, x_n)Qv_d(x_1, \dots, x_n)$. Similarly to the univariate case, we can thence define the nonnegative polynomials $\mathbb{R}_{2d}^+[x_1, \dots, x_n]$ and the SOS polynomials $\Sigma_{2d}[x_1, \dots, x_n]$. In the multivariate case, however positive semi-definite real-symmetric matrices do not necessarily characterise being nonnegative. i.e., it is in general not true that every nonnegative polynomial is SOS. For instance $g(x_1, x_2) = x_1^2x_2^2(x_1^2 + x_2^2 - 1) + 1$ is a nonnegative polynomial that does not have a SOS representation (Lasserre, 2009, Sec.2.4). Hilbert (1888) showed the following.

Proposition 5. $\mathbb{R}_{2d}^+[x_1, \dots, x_n] = \Sigma_{2d}[x_1, \dots, x_n]$ holds iff either $n = 1$ or $d = 1$ or $(n, d) = (2, 2)$.

The problem of testing global nonnegativity of a polynomial function is in general *NP-hard*. If Alice wants to avoid the complexity associated with this problem, an alternative option is to consider a subset of polynomials for which a nonnegativity test is not *NP-hard*. The problem of testing if a given polynomial is SOS has polynomial complexity (we only need to check if the matrix of coefficients Q in (15) is positive semi-definite), see Lasserre (2009).

Example 3. Let us consider the polynomial $f = \frac{1}{4} - x(1-x)$, we want to show that f is SOS. Let us attempt to rewrite it as

$$f = \frac{1}{4} - x(1-x) = \sigma_0(x),$$

for $\sigma_0(x) \in \Sigma_{2d}$ and, therefore, $\sigma_0(x) = [1, x]Q[1, x]^T$ with Q being a 2×2 positive semi-definite matrix. By equating the coefficients of the polynomials in $\frac{1}{4} - x(1-x) = [1, x]Q[1, x]^T$, we find the matrix

$$Q = \begin{bmatrix} \frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

The matrix is positive semi-definite and, thus, the polynomial f is SOS.

5. Bounded rationality

In the bounded rationality theory we are going to present we will work with $\Omega = \mathbb{R}^n$ and make two important assumptions. We assume that \mathcal{L} is the set of multivariate polynomials of n variables and of degree less than or equal to $2d$, with $d \in \mathbb{N}$. We denote \mathcal{L} as \mathcal{L}_{2d} and the nonnegative polynomials as \mathcal{L}_{2d}^+ . Note that \mathcal{L}_{2d} is a vector space and A.1–A.5 are well-defined in \mathcal{L}_{2d} . This restriction is useful to define the computational complexity of our bounded rationality theory as a function of n and d . We now define our bounded rationality criteria, and point out the two assumptions.

Definition 3. We say that $\mathcal{C} \subset \mathcal{L}_{2d}$ is a **bounded-rationality coherent set of almost desirable gambles (BADG)** when it satisfies A.3–A.5 (i.e. it is a closed convex cone) and:

bA.1 If $g \in \Sigma_{2d}$ then $g \in \mathcal{C}$ (Bounded Accepting Partial Gain),

bA.2 If $g \in \Sigma_{2d}^-$ then $g \notin \mathcal{C}$ (Bounded Avoiding Sure Loss);

where $\Sigma_{2d} \subset \mathcal{L}_{2d}^+$ is the set of SOS of degree less than or equal to $2d$ and $\Sigma_{2d}^- := \{g \in \mathcal{L}_{2d} \mid \sup g < 0, -g \in \Sigma_{2d}\}$ is the set of negative SOS polynomials of degree less than or equal to $2d$ (or stated otherwise, it is the interior of $-\Sigma_{2d}$).

We have seen that A.1 implies that a coherent set of gambles must include all nonnegative gambles (and, therefore, \mathcal{L}_{2d}^+ that is the set of all nonnegative polynomials) and that, additionally, A.2 means that it should not include negative gambles. Here, we restrict A.1 and A.2 imposing bounded-rationality that implies that the set must only include SOS polynomials and avoid negative SOS polynomials, up to degree $2d$. In BADG theory, we ask Alice only to always accept gambles for which she can efficiently determine the nonnegativity (SOS polynomials) and to never accept gambles for which she can efficiently determine the negativity (negative SOS polynomials). Using the terminology from (de Cooman & Quaeghebeur, 2012, Definition 1)⁵, the set \mathcal{C} is said to be coherent *relative* to the pair constituted by the vector subspace of quadratic forms $v_{2d}(x_1, \dots, x_n)^T Q v_{2d}(x_1, \dots, x_n)$ defined by the symmetric real matrices Q and the closed⁶ convex cone of SOS polynomials (or equivalently the closed convex cone of polynomials defined by a positive semi-definite real-symmetric matrix).

In the multivariate case, we have seen that there are nonnegative polynomials that do not have a SOS representation. These polynomials should be in principle desirable for Alice in the ADG framework, but in BADG we do not enforce Alice to accept them. A similar reasoning holds for the difference between A.2 and bA.2: in the chosen framework, we cannot say that Alice is irrational if she chooses a negative gamble for which she cannot verify computationally the negativity. Despite the fact that in principle they should never be accepted by Alice in the ADG framework, in a BADG we thus do not enforce this property and we only ask Alice to avoid gambles in Σ_{2d}^- . For these reasons, BADG is a theory of bounded rationality.

Alice may not be able to prove that her set of desirable gambles satisfies A.2. However, by exploiting the fact that

$$\Sigma_{2d} \subseteq \mathcal{L}_{2d}^+ \subset \mathcal{L}^+,$$

a BADG set \mathcal{C} that satisfies A.2 but not A.1 can (theoretically) be turned to:

1. an ADG in \mathcal{L}_{2d} by considering its extension $\text{posi}(\mathcal{C} \cup \mathcal{L}_{2d}^+)$, and also to
2. an ADG in \mathcal{L} by considering its extension $\text{posi}(\mathcal{C} \cup \mathcal{L}^+)$.

The other way round is also true. Namely:

Proposition 6. Let $G \subseteq \mathcal{L}_{2d}$ be a finite set of assessments, and assume $\mathcal{H} = \text{posi}(G \cup \mathcal{L}_{2d}^+)$ (resp. $\mathcal{H} = \text{posi}(G \cup \mathcal{L}^+)$) is ADG. Then $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d}^+)$ is BADG.

Proof. We just verify the claim with \mathcal{L}_{2d}^+ , the other, mutatis mutandis, being verified the same way. Since by assumption \mathcal{H} is ADG, \mathcal{H} is a closed convex cone containing \mathcal{L}^+ and disjoint from the set of negative gambles. In particular \mathcal{H} is BADG. Therefore $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d}^+) \subseteq \mathcal{H}$ and it is a closed convex cone, meaning it is BADG too. \square

These remarks are important because, as it will be shown in the next sections, they will allow us to use BADG as a computable approximation of ADG.

In BADG theory, Proposition 3 is reformulated as follows.

⁵Notice that the authors use a different notion of coherence: they do not assume the closure condition (A.5), and they would require that both the zero gamble and gambles in $-\Sigma_{2d}$ are not desirable.

⁶Closedness of the convex cone of SOS was proved by Robinson (1969).

Theorem 4. Given a finite set $G \subset \mathcal{L}_{2d}$ of desirable gambles, the set $\text{posi}(G \cup \Sigma_{2d})$ includes the gamble f if and only if there exist $\lambda_j \geq 0$ for $j = 1, \dots, |G|$ such that

$$f - \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d}. \quad (17)$$

Proof. Assume that $f \in \text{posi}(G \cup \Sigma_{2d})$ then there exist $\sigma_i \in \Sigma_{2d}$ and $\gamma_i \geq 0$ for $i = 1, \dots, m$ such that $f = \sum_{j=1}^{|G|} \lambda_j g_j + \sum_{i=1}^m \gamma_i \sigma_i$. This implies that $f - \sum_{j=1}^{|G|} \lambda_j g_j$ is SOS, proving one implication. The other implication can be proven similarly. \square

If we compare Proposition 3 and Theorem 4, then we see the difference between ADG and BADG:

$$\text{ADG: } f - \sum_{j=1}^{|G|} \lambda_j g_j \in \mathcal{L}^+, \quad \text{BADG: } f - \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d},$$

that consists in the definition of nonnegativity or, equivalently, nonnegative gambles, i.e., the gambles that Alice shall always accept.

Also for BADG we can consider the gamble $f = h - \lambda_0$ for some $\lambda_0 \in \mathbb{R}$ and define the concept of lower prevision.

Definition 4. Let $G \subset \mathcal{L}_{2d}$ be a finite set, and let $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d})$. Assume that \mathcal{C} is BADG, then the solution of the following problem

$$\sup_{\lambda_0 \in \mathbb{R}, \lambda_j \geq 0} \lambda_0, \quad \text{s.t.} \quad h - \lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d}, \quad (18)$$

is called the lower prevision of h and denoted as $\underline{E}^*[h]$.

We can prove that $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d})$ satisfies bounded avoiding sure loss exploiting the following result.

Proposition 7. Let us consider $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d})$ and the following problem:

$$\sup_{0 \leq \lambda_0 \leq 1, \lambda_j \geq 0} \lambda_0, \quad \text{s.t.} \quad -\lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d}. \quad (19)$$

\mathcal{C} does not satisfy b.A.2, and thus incurs in a sure loss iff the above problem has solution $\lambda_0^* = 1$, and it avoids bounded sure loss (b.A.2 is satisfied) iff $\lambda_0^* = 0$.

Proof. Assume that \mathcal{C} b.A.2 is false, and thus incurs in a sure loss. This means there exists $\lambda_j \geq 0$ for $j = 1, \dots, |G|$ such that $f := \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d}^-$. Hence, $-f$ is SOS (belongs to Σ_{2d}) and strictly positive, yielding that we can increase λ_0 as much as we want, because we can find $\sigma \in \Sigma_{2d}$ such that $-f = \sigma + \lambda_0$ is true (given $-f > 0$). In practice, we are exploiting the fact that for any positive scaling constant ρ the following equality still holds $-f\rho = \rho\sigma + \rho\lambda_0$ and so $\lambda_0 = 1$. Now assume that there is no $\lambda_j \geq 0$ for $j = 1, \dots, |G|$ such that $\sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_{2d}^-$. The only way for $-\lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j$ being SOS is that $\lambda_0 = 0$. \square

5.1. Duality for BADG

We can also define the dual of a BADG. In this case, the gambles g are polynomials, the nonnegative gambles that Alice accepts are SOS, and the negative gambles that she does not accept are the negative polynomials g such that $-g$ is SOS. Since we are dealing with a vector space, we can consider its dual space \mathcal{L}_{2d}^\bullet of all linear maps $L : \mathcal{L}_{2d} \rightarrow \mathbb{R}$ (linear functionals) and thus define the dual of $\mathcal{C} \subset \mathcal{L}_{2d}$ as the set $\{L \in \mathcal{L}_{2d}^\bullet : L(g) \geq 0, \forall g \in \mathcal{C}\}$. Since \mathcal{L}_{2d} has a basis, i.e., the monomials, if we introduce the scalars

$$y_{\alpha_1 \alpha_2 \dots \alpha_n} := L(x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}) \in \mathbb{R}, \quad (20)$$

where $\alpha_i \in \mathbb{N}$, we can rewrite $L(g)$ for any polynomial g as a function of the vector of variables $y \in \mathbb{R}^{s_n(2d)}$, whose components are the real variables $y_{\alpha_1 \alpha_2 \dots \alpha_m}$ defined above. This means that the dual space \mathcal{L}_{2d}^\bullet is isomorphic to $\mathbb{R}^{s_n(2d)}$, and we can thence define the dual maps $(\cdot)^\bullet$ between \mathcal{L}_{2d} and $\mathbb{R}^{s_n(2d)}$ as follows.

Definition 5. Let \mathcal{C} be a subset of \mathcal{L}_{2d} . Its dual is defined as

$$\mathcal{C}^\bullet = \left\{ y \in \mathbb{R}^{s_n(2d)} : L(g) \geq 0, \forall g \in \mathcal{C} \right\}, \quad (21)$$

where $L(g)$ is completely determined by y via the definition (20). Similarly, given a subset \mathcal{R} of $\mathbb{R}^{s_n(2d)}$, its dual is defined as

$$\mathcal{R}^\bullet = \{ g \in \mathcal{L}_{2d} : L(g) \geq 0, \forall y \in \mathcal{R} \}, \quad (22)$$

As before, $(\cdot)^\bullet$ is an anti-monotonic operation and its image is always a closed convex cone. Furthermore, one has that $(\cdot)_1^{\bullet\bullet} = (\cdot)_1$, and $(\cdot)_1 \subseteq (\cdot)_2$ if and only if $(\cdot)_2^\bullet \subseteq (\cdot)_1^\bullet$, whenever $(\cdot)_1$ and $(\cdot)_2$ are closed convex cones.

In what follows, we verify that, analogously to Section 3, the dual \mathcal{C}^\bullet is completely characterised by a closed convex set of states. But before doing that, we have to clarify what is a state in this context. In the previous section, we defined a nonnegative linear functional (operator) as a map that assigns nonnegative real numbers to gambles that are sure gains, that is to gambles satisfying the condition for rationality axiom A.1. In the actual bounded rationality theory, we have replaced sure gains with bounded sure gains. Hence, to define what is a state we cannot refer to nonnegative gambles but to gambles that are SOS. This means that, consistently with axiom bA.1, from the adopted bounded perspective on rationality, states are linear operators that assign nonnegative real numbers to SOS, and that additionally preserve the unit gamble. This latter condition is equivalent to:

$$y_0 = L(1) = 1.$$

In the aim of reducing the dual of a BADG to sets of states, we thus first provide a characterisation of nonnegative linear operators. In doing so, we define the matrix $M_{n,d}(y) := L(v_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)$, where the linear operator is applied component-wise. For instance, in the case $n = 1$ and $d = 2$, we have that

$$M_{1,2}(y) = L(v_2(x_1)v_2(x_1)^\top) = L \left(\begin{bmatrix} 1 & x_1 & x_1^2 \\ x_1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \end{bmatrix} \right) = \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix}.$$

In discussing properties of the dual space, we need the following well-known result from linear algebra:

Lemma 1. For any $M \in \mathbb{R}^{n \times n}$ and $v \in \mathbb{R}^n$, it holds that

$$\text{Tr}(M(vv^\top)) = \text{Tr}((vv^\top)M) = v^\top Mv. \quad (23)$$

Proposition 8. Let $g \in \mathbb{R}_{2d}[x_1, \dots, x_n]$ and Q a real symmetric-matrix such that $g(x_1, \dots, x_n) = v_d^\top(x_1, \dots, x_n)Qv_d(x_1, \dots, x_n)$. Then for every $y \in \mathbb{R}^{s_n(2d)}$, it holds that $L(g) = \text{Tr}(QM_{n,d}(y))$, where $M_{n,d}(y) = L(v_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)$ and $L(g)$ is completely determined by y via the definition (20).

Proof. By Lemma 1 and linearity of L and trace, we obtain that

$$\begin{aligned} L(g) &= L(v_d(x_1, \dots, x_n)^\top Qv_d(x_1, \dots, x_n)) \\ &= L(\text{Tr}(Qv_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)) \\ &= \text{Tr}(QL(v_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)) \\ &= \text{Tr}(QM_{n,d}(y)), \end{aligned}$$

where $M_{n,d}(y) = L(v_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)$. □

We then verify that

Proposition 9. Let $\mathcal{C} = \Sigma_{2d}$. Then its dual is

$$\mathcal{C}^\bullet = \left\{ y \in \mathbb{R}^{s_n(2d)} : M_{n,d}(y) \geq 0 \right\}, \quad (24)$$

Proof. Recall that, by definition, $M_{n,d}(y) = L(v_d(x_1, \dots, x_n)v_d(x_1, \dots, x_n)^\top)$, and that, by Equation (15), any $g \in \Sigma_{2d}$ can be written as $v_d(x_1, \dots, x_n)^\top Q v_d(x_1, \dots, x_n)$, with $Q \geq 0$. Fejér's trace theorem (Boyd & Vandenberghe, 2004, Ex.2.24) states that a matrix $M \in \mathbb{R}_s^{s_1(d) \times s_1(d)}$ is $M \geq 0$ if and only if $\text{Tr}(QM) \geq 0$, for any $Q \geq 0$. Hence Equation (24) is an immediate consequence of the following equivalences:

$$\begin{aligned} & M_{n,d}(y) \geq 0 \\ \iff & \text{Tr}(QM_{n,d}(y)), \forall Q \geq 0 \quad (\text{Fejér's trace theorem}) \\ \iff & L(g) \geq 0, \forall g \in \Sigma_{2d} \quad (\text{Proposition 8}) \end{aligned}$$

□

Hence a linear operator L is nonnegative if and only if $M_{n,d}(y) \geq 0$. Obviously, if $y = 0$, then $L(g) \geq 0$, for each $g \in \Sigma_{2d}$, and thus $\{0\}^\bullet = \mathcal{L}_{2d}$. From this observation, Proposition 9 and the properties of $(\cdot)^\bullet$ we therefore immediately get:

Proposition 10. It holds that

1. $(\mathcal{L}_{2d})^\bullet = \{0\}$ and $\mathcal{L}_{2d} = (\{0\})^\bullet$;
2. $(\Sigma_{2d})^\bullet = \mathcal{Y}^+$ and $\Sigma_{2d} = (\mathcal{Y}^+)^\bullet$,

where $\mathcal{Y}^+ := \left\{ y \in \mathbb{R}^{s_n(2d)} : M_{n,d}(y) \geq 0 \right\}$

Everything is therefore ready to provide an analogous characterisation of coherence as done with Proposition 2 but for BADG.

Proposition 11. Let $\mathcal{C} \subseteq \mathcal{L}_{2d}$ be a closed convex cone. The following claims are equivalent

1. \mathcal{C} is coherent;
2. $\mathcal{C} \supseteq \Sigma_{2d}$ and $\mathcal{C} \neq \mathcal{L}_{2d}$;
3. $(\mathcal{C})^\bullet \subseteq \mathcal{Y}^+$ and $(\mathcal{C})^\bullet \neq \{0\}$.

Proof. (2) \Leftrightarrow (3): From Proposition 10, $(\mathcal{C})^\bullet = \{0\}$ if and only if $\mathcal{C} = \mathcal{L}_{2d}$, and $\mathcal{C} \supseteq \Sigma_{2d}$ if and only if $(\mathcal{C})^\bullet \subseteq \mathcal{Y}^+$. (1) \Rightarrow (2): Assume \mathcal{C} is coherent. By bA.1 $\mathcal{C} \supseteq \Sigma_{2d}$, and by bA.2 there is $-g \in \Sigma_{2d}$ such that $\sup g < 0$ and $g \notin \mathcal{C}$. (2) \Rightarrow (1): Let $\Sigma_{2d} \subseteq \mathcal{C} \subsetneq \mathcal{L}_{2d}$. First of all, notice that, by Proposition 10, $(\mathcal{C})^\bullet \subseteq (\Sigma_{2d})^\bullet = \mathcal{Y}^+$. Now, assume that \mathcal{C} is not coherent. This means bA.2 fails. Hence we can pick $-g \in \Sigma_{2d}$ such that $\sup g < 0$ and $g \in \mathcal{C}$. Consider $y \in \mathcal{Y}^+$. It holds that $L(g) \geq 0$ if and only if $y = 0$, meaning that $(\mathcal{C})^\bullet = \{0\}$ and therefore, by Proposition 10 again, $\mathcal{C} = \mathcal{L}_{2d}$, a contradiction. □

As before, we denote by \mathcal{S} the set of states (here seen as a subset of $y \in \mathbb{R}^{s_n(2d)}$). By Proposition 11 and reasoning exactly as for Theorem 3, we then have the following result (see for instance Lasserre 2009).

Theorem 5. The map

$$\mathcal{C} \mapsto \mathcal{P} := \mathcal{C}^\bullet \cap \mathcal{S}$$

is a bijection between BADGs and closed convex subsets of \mathcal{S} . We can therefore identify the dual of a BADG \mathcal{C} with

$$\mathcal{C}^\bullet = \left\{ y \in \mathbb{R}^{s_n(2d)} : L(g) \geq 0, L(1) = 1, M_{n,d}(y) \geq 0, \forall g \in \mathcal{C} \right\}, \quad (25)$$

where $L(g)$ is completely determined by y via the definition (20).

Proof. Let \mathcal{C} be a coherent BADG. By Proposition 11, we get that \mathcal{C}^\bullet is a closed convex cone included in \mathcal{Y}^+ that does not reduced to the origin. Thus \mathcal{P} is nonempty. Preservation of closedness and convexity by finite intersections yields that \mathcal{P} is closed and convex. Furthermore $\mathbb{R}_+ \mathcal{P} = (\mathcal{C}^\bullet)^\bullet$, and therefore $\mathcal{C} = (\mathbb{R}_+ \mathcal{P})^\bullet$, where $\mathbb{R}_+ \mathcal{P} := \{\lambda y : \lambda \geq 0, y \in \mathcal{P}\}$, meaning that the map is an injection. We finally verify that the map is also a surjection. To do this, let $\mathcal{P} \subseteq \mathcal{S}$ be a non empty closed convex set of states. It holds that $\mathbb{R}_+ \mathcal{P}$ is a closed convex cone included in \mathcal{Y}^+ different from $\{0\}$. Again by Proposition 11, we conclude that the dual $(\mathbb{R}_+ \mathcal{P})^\bullet$ is a coherent BADG and $\mathcal{P} = \mathbb{R}_+ \mathcal{P} \cap \mathcal{S} = (\mathbb{R}_+ \mathcal{P})^{\bullet\bullet} \cap \mathcal{S}$. \square

In Section 3.1, by considering the space of all measurable gambles and identifying gambles representing sure gain with nonnegative gambles, states coincide with probability measures. Henceforth we have shown that the dual of an ADG is a closed convex set of probability measures. In (25) there is no reference to probability, and thus there is no guaranty that in the bounded rationality case states correspond indeed to probabilities. However, by considering the Borel sigma-algebra on \mathbb{R}^n , we can define the integral $\int x_1^{\alpha_1} x_2^{\alpha_2}, \dots, x_n^{\alpha_n} d\mu$ with respect to the finite signed measure μ . In this context, we can interpret $y_{\alpha_1 \alpha_2 \dots \alpha_n}$ as the expectation of $x_1^{\alpha_1} x_2^{\alpha_2}, \dots, x_n^{\alpha_n}$ w.r.t. the signed measure μ .

Note that $y_0 = L(1) = 1$ implies that $\int 1 d\mu = 1$ under this interpretation (normalization). Therefore, we can interpret $M_{n,d}(y)$ as a truncated moment matrix. However, since \mathcal{C} does not include all nonnegative gambles, we cannot conclude that the signed measures are nonnegative or, in other words, that μ is a probability measure. The constraint $M_{n,d}(y) \geq 0$ is not strong enough to guarantee nonnegativity of μ (it is only a necessary condition).

In the standard theory, negative probabilities are considered a manifestation of incoherence. Here, they are a consequence of the assumption of bounded rationality. Finally, the dual of the lower prevision problem (18) is then given by the convex SemiDefinite Programming (SDP) problem:

$$\inf_{y \in \mathbb{R}^{sn(2d)}} L(h), \quad \text{s.t.} \quad L(g) \geq 0, \quad L(1) = 1, \quad M_{n,d}(y) \geq 0. \quad (26)$$

5.2. Non-SOS positive polynomials

What does it mean for our theory of bounded rationality that there are positive polynomials that are not SOS? First, notice that, by Proposition 11, $g \in \Sigma_{2d}$ if and only if for every $y \in \mathcal{Y}^+$, it holds that $L(g) \geq 0$ (where L is completely determined by y via equation (20)). This means that if g is positive but not SOS, then there is $y \in \mathcal{Y}^+$ such that $L(g) < 0$. Another, equivalent, way to see this goes as follows. Let us assume that the polynomial $g \in \mathcal{L}_{2d}$ is positive but not SOS, its lower prevision is:

$$\sup_{\lambda_0 \in \mathbb{R}} \lambda_0 \text{ s.t. } g - \lambda_0 \in \Sigma_{2d}.$$

The solution of the above problem is $\lambda_0 < 0$. Note in fact that, since $g(x) = v^T(x)Qv(x)$ is not SOS, this implies that the matrix Q is indefinite (it is not a PSD matrix). Hence, the only way to satisfy $g - \lambda_0 \in \Sigma_{2d}$ is for $\lambda_0 < 0$. By duality, we can then prove that the problem

$$\inf_{y \in \mathbb{R}^{sn(2d)}} L(g) \text{ s.t. } L(1) = 1, \quad M_{n,d}(y) \geq 0,$$

has a negative solution, i.e. $L(g) < 0$.

Example 4. Let us consider the positive non-SOS polynomial $f(x) = 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - x_1^2 x_2^2$ in $\mathbb{R}_6[x_1, x_2]$, the basis $v_3(x) = [1, x_1, x_2, x_1^2, x_1 x_2, x_2^2, x_1^3, x_1^2 x_2, x_1 x_2^2, x_2^3]^T$ and the following PSD matrix $M_{2,3}(y) = L(v_3(x)v_3(x)^T)$:

$$\begin{bmatrix} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} & y_{30} & y_{21} & y_{12} & y_{03} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} & y_{40} & y_{31} & y_{22} & y_{13} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} & y_{31} & y_{22} & y_{13} & y_{04} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} & y_{50} & y_{41} & y_{32} & y_{23} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} & y_{41} & y_{32} & y_{23} & y_{14} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} & y_{32} & y_{23} & y_{14} & y_{05} \\ y_{30} & y_{40} & y_{31} & y_{50} & y_{41} & y_{32} & y_{60} & y_{51} & y_{42} & y_{33} \\ y_{21} & y_{31} & y_{22} & y_{41} & y_{32} & y_{23} & y_{51} & y_{42} & y_{33} & y_{24} \\ y_{12} & y_{22} & y_{13} & y_{32} & y_{23} & y_{14} & y_{42} & y_{33} & y_{24} & y_{15} \\ y_{03} & y_{13} & y_{04} & y_{23} & y_{14} & y_{05} & y_{33} & y_{24} & y_{15} & y_{06} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 353 & 0 & 353 & 0 & 0 & 0 & 0 \\ 0 & 353 & 0 & 0 & 0 & 0 & 249572 & 0 & 66 & 0 \\ 0 & 0 & 353 & 0 & 0 & 0 & 0 & 0 & 66 & 0 & 249572 \\ 353 & 0 & 0 & 249572 & 0 & 66 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 66 & 0 & 0 & 0 & 0 & 0 & 0 \\ 353 & 0 & 0 & 66 & 0 & 249572 & 0 & 0 & 0 & 0 & 0 \\ 0 & 249572 & 0 & 0 & 0 & 0 & 706955894 & 0 & 17 & 0 & 0 \\ 0 & 0 & 66 & 0 & 0 & 0 & 0 & 0 & 17 & 0 & 17 \\ 0 & 66 & 0 & 0 & 0 & 0 & 0 & 17 & 0 & 17 & 0 \\ 0 & 0 & 249572 & 0 & 0 & 0 & 0 & 0 & 17 & 0 & 706955894 \end{bmatrix} \quad (27)$$

Since $L(f) = 1 + y_{42} + y_{24} - y_{22}$ and $y_{22} = 66, y_{24} = y_{42} = 17$ in (27), we have that $L(f) = -31 < 0$. The above matrix is PSD but it is not the truncated moment matrix of any probability measure (if it would be then $L(g) \not< 0$).

How is it possible?

The cone $\mathcal{Y}^+ = \left\{ y \in \mathbb{R}^{s_n(2d)} : M_{n,d}(y) \geq 0 \right\}$ includes the evaluation functionals of the polynomials.⁷ Evaluation functionals coincide with the rank one matrices M that give the evaluation of g at a point \bar{x} as $Tr(QM) = g(\bar{x})$, for any decomposition of $g(x) = v_d^T(x)Qv_d(x)$. Such matrices have the form $M = v_d(\bar{x})v_d(\bar{x})^T$. However, contrary to what happens in the standard theory of desirable gambles as described in Sections 2 and 3, these matrices do not exhaust all extremes of the closed convex set \mathcal{Y}^+ .

Said in another way, since in the space of Borel measures the evaluation functionals are the atomic measures and since $v(\bar{x})v(\bar{x})^T = \int v_d(x)v_d(x)^T \delta_{\{\bar{x}\}} dx$, there does not exist a mixture of atomic measures $\sum_{i=1}^m w_i \delta_{\{\bar{x}^{(i)}\}}$ such that

$$\int v_3(x)v_3(x)^T \left(\sum_{i=1}^m w_i \delta_{\{\bar{x}^{(i)}\}} \right) dx = \sum_{i=1}^m w_i v_3(\bar{x}^{(i)})v_3(\bar{x}^{(i)})^T = M_{2,3}(y).$$

In (27), the only way to satisfy the above equality is that some of weights w_i are negative.

Similarly, since $-L(-g) = L(g)$, we can also conclude that, for a negative gamble g whose inverse $-g$ is not SOS, we have $L(g) > 0$. Alice may accept a negative gamble not belonging to Σ_{2d}^- ! In BADG, this is allowed because evaluating the negativity of a non SOS gamble not in Σ_{2d}^- may be computationally intractable. In the next section, we will show that/when we can use BADG as a computable approximating theory for ADG.

6. BADG as an approximating theory for ADG

We are going to show that we can use BADG as a computable approximating theory for ADG. Since we are dealing with unbounded gambles, we will refer to the ADG formulation in Section 3.1. So let us consider the BADG set $\mathcal{C} = \text{posi}(G \cup \Sigma_{2d})$ and the corresponding ADG set $\mathcal{K} = \text{posi}(G \cup \mathcal{L}^+)$ (same G), where \mathcal{L}^+ is the set of measurable non-negative gambles. We have the following result.

Theorem 6. *Assume that \mathcal{K} avoids sure loss (i.e. satisfies A.2) and let $f \in \mathcal{L}_{2d}$, then BADG is a conservative approximation of ADG theory in the sense that $\underline{E}^*(f) \leq \underline{E}(f)$, where $\underline{E}(f)$ is the coherent lower prevision of f computed with respect to the set of probability measures compatible with Alice's assessments of desirability G .*

Proof. Notice that since \mathcal{K} satisfies A.2, \mathcal{C} also satisfies A.2, and hence bA.2. Now let λ_0^* be the supremum value of λ_0 such that $f - \lambda_0 - \sum_j^{|\mathcal{G}|} \lambda_j g_j \in \Sigma_{2d}$ and λ_0^{**} the supremum value such that $f - \lambda_0 - \sum_j^{|\mathcal{G}|} \lambda_j g_j \geq 0$. Since the constraint $f - \lambda_0 - \sum_j^{|\mathcal{G}|} \lambda_j g_j \in \Sigma_{2d}$ is more demanding than $f - \lambda_0 - \sum_j^{|\mathcal{G}|} \lambda_j g_j \geq 0$, it follows that $\lambda_0^* \leq \lambda_0^{**}$. \square

The fact that $\underline{E}[f]$ is equal to the minimum of f when G is empty, i.e., Alice is in a state of full ignorance, is one of the reasons why SOS polynomials are used in optimisation. In fact, $\underline{E}^*[f]$ provides a lower bound for the minimum of f (Lasserre, 2009).

Example 5 (Covariance inequality). *Let us consider the case $n = 2, d = 1$, the matrix $M_{2,1}(y)$ is in this case*

$$M_{2,1}(y) = L \left(\begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{bmatrix} \right) = \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix}.$$

Let us assume that Alice finds the following 8 polynomials to be desirable:

$$G = \{ \pm(x_1 - m_1), \pm(x_2 - m_2), \pm(x_1^2 - m_1^2 - s_1^2), \pm(x_2^2 - m_2^2 - s_2^2) \}.$$

Since $L(\pm(x_1 - m_1)) = \pm(y_{10} - m_1)$, $L(\pm(x_2 - m_2)) = \pm(y_{01} - m_2)$, $L(\pm(x_1^2 - m_1^2 - s_1^2)) = \pm(y_{20} - m_1^2 - s_1^2)$, $L(\pm(x_2^2 - m_2^2 - s_2^2)) = \pm(y_{02} - m_2^2 - s_2^2)$ and given that $y = [y_{10}, y_{01}, y_{20}, y_{11}, y_{02}]^T$, we have that the dual of the BADG \mathcal{C} is:

$$\mathcal{C}^\bullet = \left\{ y^T \in \mathbb{R}^5 : y_{10} = m_1, y_{01} = m_2, y_{20} = m_1^2 + s_1^2, y_{02} = m_2^2 + s_2^2, M_{2,1}(y) \geq 0 \right\}. \quad (28)$$

⁷An evaluation functional over \mathcal{L} is a linear functional that evaluates each gamble g at a point \bar{x} .

Hence, it follows that

$$M_{2,1}(y) = \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} = \begin{bmatrix} 1 & m_1 & m_2 \\ m_1 & m_1^2 + s_1^2 & y_{11} \\ m_2 & y_{11} & m_2^2 + s_2^2 \end{bmatrix}.$$

Assume that we aim at computing $E^*[(x_1 - m_1)(x_2 - m_2)]$, $\bar{E}^*[(x_1 - m_1)(x_2 - m_2)]$, i.e., the lower/upper prevision of the gamble $(x_1 - m_1)(x_2 - m_2)$. Note that $L((x_1 - m_1)(x_2 - m_2)) = y_{11} - m_1 m_2$. From $M_{2,1}(y) \geq 0$ we can derive that

$$-s_1 s_2 \leq y_{11} - m_1 m_2 \leq s_1 s_2.$$

These inequalities follow by $\det(M_{2,1}(y)) \geq 0$. From (26), we thus have that $E^*[(x_1 - m_1)(x_2 - m_2)] = -s_1 s_2$ and $\bar{E}^*[(x_1 - m_1)(x_2 - m_2)] = s_1 s_2$. By interpreting m_1, m_2, s_1^2, s_2^2 as the means and variances of x_1, x_2 and observing that the above two inequalities can be written as

$$(y_{11} - m_1 m_2)^2 \leq s_1^2 s_2^2,$$

from Theorem 6 (this theorem holds because \mathcal{C} satisfies A.2) we can derive that

$$E[(x_1 - m_1)(x_2 - m_2)]^2 \leq \text{Var}(x_1)^2 \text{Var}(x_2)^2. \quad (29)$$

This is the well-known covariance inequality in probability theory. Observe that there exists a probability measure in BADG for which the above inequality is tight: $\frac{1}{2}\delta_{\left(\begin{smallmatrix} m_1 - s_1 \\ m_2 - s_2 \end{smallmatrix}\right)}(x) + \frac{1}{2}\delta_{\left(\begin{smallmatrix} m_1 + s_1 \\ m_2 + s_2 \end{smallmatrix}\right)}(x)$, here $\delta_{(a)}$ denotes an atomic measure (Dirac's delta) at a . It can in fact be verified that this probability measure satisfies all the moment constraints:

$$E[x_1] = m_1, \quad E[x_2] = m_2, \quad E[x_1 x_2] = m_1 m_2 + s_1^2 s_2^2, \quad E[x_1^2] = m_1^2 + s_1^2, \quad E[x_2^2] = m_2^2 + s_2^2.$$

Hence, Theorem 6 is tight in this case. However, there are also signed measures that are compatible with Alice's assessments:

$$\delta_{\left(\begin{smallmatrix} m_1 - \frac{s_1}{\sqrt{2}} \\ m_2 - \frac{s_2}{\sqrt{2}} \end{smallmatrix}\right)}(x) - \delta_{\left(\begin{smallmatrix} m_1 \\ m_2 \end{smallmatrix}\right)}(x) + \delta_{\left(\begin{smallmatrix} m_1 + \frac{s_1}{\sqrt{2}} \\ m_2 + \frac{s_2}{\sqrt{2}} \end{smallmatrix}\right)}(x),$$

and that achieve the equality in (29) but that are not probabilities.

Example 6. Consider the case $n = 1, d = 2$ and assume that

$$G = \{\pm(x_1 - 1), \pm(x_1^2 - 1)\}.$$

Therefore, we have that

$$M_{1,2}(y) = L\left(\begin{bmatrix} 1 & x_1 & x_1^2 \\ x_1 & x_1^2 & x_1^3 \\ x_1^2 & x_1^3 & x_1^4 \end{bmatrix}\right) = \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & y_3 \\ 1 & y_3 & y_4 \end{bmatrix}.$$

Assume we are interested in computing the upper prevision $\bar{E}^*[x_1^4]$. From (26), we have that this upper prevision is equal to

$$\sup_{y_3, y_4 \in \mathbb{R}^2} y_4 \quad \text{s.t.} \quad M_{1,2}(y) \geq 0. \quad (30)$$

Note that the supremum is unbounded, since all matrices of the form

$$M_{1,2}(y) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & y_4 \end{bmatrix}$$

are positive semi-definite for every $y_4 \geq 1$. $M_{1,2}(y)$ is positive semi-definite, but it is not the truncated moment matrix of any probability measure. Note in fact that $E[x_1] = E[x_1^2] = 1$ would imply the probability measure to be atomic on $x_1 = 1$ and so $\bar{E}[x_1^4] = 1$ and, therefore, it cannot be that $E[x_1^4] > 1$. But (for instance for $y_4 = 2$) we can find an atomic signed measure $1.014\delta_{1.043} + 1.182\delta_{3.952} + 0.004\delta_{-1.654} - 0.920\delta_{3.938} - 0.281\delta_{3.920}$ that has those moments, but it is not a probability measure (negative weights).

In the next section, we restrict Ω to avoid unbounded previsions.

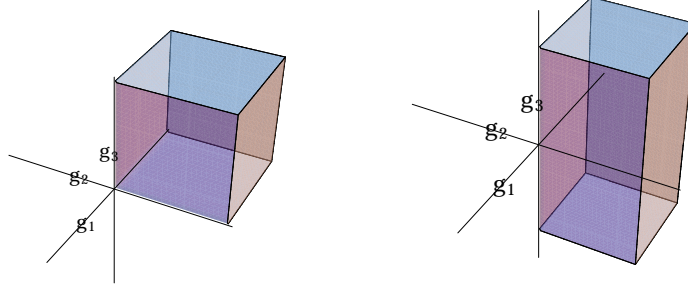


Figure 2: Cones of nonnegative gambles

7. BADG on semi-algebraic sets

Let us assume that Ω is a semi-algebraic set, i.e., a set described by polynomial inequalities

$$\Omega = \left\{ x = [x_1, \dots, x_n]^\top \in \mathbb{R}^n : c_0(x) = 1 \geq 0, c_j(x) \geq 0, j = 1, \dots, |C| \right\}, \quad (31)$$

where $C = \{c_1, \dots, c_{|C|}\}$ with $c_j(x) \in \mathbb{R}_{2n_{c_j}}[x]$ or $c_j(x) \in \mathbb{R}_{2n_{c_j}-1}[x]$ (depending if the polynomial has an even or odd degree). That means that Alice knows that x belongs to the set $\Omega \subseteq \mathbb{R}^n$. Note that when $C = \emptyset$, we have $\Omega = \mathbb{R}^n$. We have introduced the redundant constraint $c_0(x) = 1 \geq 0$ for convenience in the proofs to follow.

In ADG, the knowledge that x belongs to $\Omega \subset \mathbb{R}^n$ changes the cone of nonnegative gambles in \mathbb{R}^n from all gambles g such that $g \geq 0$ to all gambles g such that $gI_\Omega \geq 0$. In other words, the cone of nonnegative gambles is in this case:

$$\mathcal{L}_\Omega^+ = \{g : \mathbb{R}^n \rightarrow \mathbb{R} : gI_\Omega \geq 0\}. \quad (32)$$

Actually in ADG we do not need to change A.1 to take into account the information $x \in \Omega$, because we can define the cone of nonnegative gambles directly in Ω

$$\mathcal{L}^+ = \{g : \Omega \rightarrow \mathbb{R} : g \geq 0\}. \quad (33)$$

To explain that, let us go back for a moment to the coin toss example but considering the possibility space $\{Head, Tail, Side\}$. A gamble g in this case has three components $g(Head) = g_1$, $g(Tail) = g_2$ and $g(Side) = g_3$. If Alice is in a state of complete ignorance, according to A.1 she shall only accept gambles such that $g_i \geq 0$ for $i = 1, 2, 3$. Her set of desirable gambles is depicted in Figure 2 (left), that is the set of all nonnegative gambles in \mathbb{R}^3 . Assume she knows that the possibility space is actually $\Omega = \{Head, Tail\}$ (*Side* is impossible), according to A.1 she shall then accept all gambles $\{g = [g_1, g_2] \in \mathbb{R}^2 : [g_1, g_2] \geq 0\}$ (this is the meaning of (33)), which are all the nonnegative gambles in \mathbb{R}^2 . Equivalently, according to (32), we can see this last cone as the 2D projection of the cone $\{g = [g_1, g_2, g_3] \in \mathbb{R}^3 : gI_\Omega = [g_1, g_2, 0] \geq 0\}$, which is showed in Figure 2 (right). Hence, in \mathbb{R}^3 , the knowledge $\Omega = \{Head, Tail\}$ may be translated in a new definition of the cone of nonnegative gambles (Figure 2 (right)), although this is not necessary in ADG.

In BADG, to express the knowledge $x \in \Omega$, we cannot use (32) because indicator functions are not polynomials. Similarly, we cannot use (33). The reason is that SOS are the computable nonnegative polynomials in \mathbb{R}^n and if we restrict the domain to Ω , then (in general) we do not know an equivalent class of computable nonnegative polynomials in Ω . Hence, we need to find another way to model $x \in \Omega$.

Let us consider (32) and notice that, for every nonnegative gamble $g : \mathbb{R}^n \rightarrow \mathbb{R}^+$, the gamble $g c_j$ is in \mathcal{L}_Ω^+ for every c_j . Similarly, we have that $g c_i c_j \in \mathcal{L}_\Omega^+$, $g c_i c_j c_k \in \mathcal{L}_\Omega^+$ and so on. The set of gambles generated in this way forms a convex cone,

$$\tilde{\mathcal{L}}_\Omega^+ := \left\{ h : h = \sum_{J \subseteq \{1, \dots, |C|\}} g c_J, g \in \mathcal{L}^+ \right\},$$

with $c_J = \prod_{j \in J} c_j$, that is included in \mathcal{L}_Ω^+ .

Since c_i are polynomials, so are σc_i and $\sigma c_i c_j$ and so on for any SOS σ . Moreover, since σ is SOS and so nonnegative, we also know that σc_i , $\sigma c_i c_j$, $\sigma c_i c_j c_k$, etc., are nonnegative in Ω . This set forms a convex sub-cone of $\tilde{\mathcal{L}}_\Omega^+$,

$$\tilde{\mathcal{L}}_\Omega^+ := \left\{ h : h = \sum_{J \subseteq \{1, \dots, |C|\}} \sigma_J c_J, \sigma_J \in \Sigma_{2d} \right\},$$

and the nonnegativity of its elements can be efficiently evaluated (it reduces to verify that σ_J is SOS) (Schmüdgen, 1991).

It is then natural in our theory of bounded rationality to translate the constraint $x \in \Omega$ in a computable sub-cone of the previous form.

We therefore give the following more general definition of BADG.

Definition 6. We say that $\mathcal{C} \subset \mathcal{L}_{2d}$ is a **bounded-rationality coherent set of almost desirable gambles (BADG)** on the semi-algebraic set Ω in (31), when $d \geq \max_j n_{c_j}$ and \mathcal{C} satisfies A.3–A.5 (i.e. it is a closed convex cone) and:

ba.1 If $g \in \Xi_{2d}$ then $g \in \mathcal{C}$ (bounded accepting partial gain);

ba.2 If $g \in \Xi_{2d}^-$ then $g \notin \mathcal{C}$ (bounded avoiding sure loss);

where Ξ_{2d} is defined as

$$\begin{aligned} \Xi_{2d} &= \left\{ \sigma_0 c_0 + \sum_{j=1}^{|C|} \sigma_j c_j : \sigma_j \in \Sigma_{2d-2n_{c_j}} \right\} \\ &= \left\{ \sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j : \sigma_0 \in \Sigma_{2d}, \sigma_j \in \Sigma_{2d-2n_{c_j}} \right\}. \end{aligned}$$

and Ξ_{2d}^- is the interior of $-\Xi_{2d}$.

Some remarks:

1. This is our bounded rationality approximation of \mathcal{L}_Ω^+ . It can be noticed that the set Ξ_{2d} does not include the terms $\sigma c_i c_j c_k$ that are also nonnegative in Ω . The number of these terms is exponential in the number of polynomials that define the set Ω and, therefore, in general not suitable for computational complexity reasons.
2. In Sections 7.2, we will show that, under certain assumptions on Ω , this definition of BADG is not conservative (Putinar, 1993).
3. Definition 6 reduces to Definition 3 when $C = \emptyset$ (so that $\Xi_{2d} = \Sigma_{2d}$).
4. Results and Definitions in Section 6 can be generalised accordingly by simply taking into account that the new set of nonnegative gambles is now Ξ_{2d} (before it was Σ_{2d}).

From now on we will use Definition 6 as definition of BADG. It means that Alice shall accept all polynomials of the form $\sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j$ because they are nonnegative in Ω . Again this is only a sufficient condition, since in general there exist nonnegative polynomials in Ω that cannot be expressed as $\sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j$.

Example 7. Let us consider the set

$$\Omega = \{x \in \mathbb{R} : 1 - x \geq 0, x \geq 0\}, \quad (34)$$

and let us assume that Alice is in a state of full ignorance. According to Definition 6, she shall only accept gambles f such that

$$f = \sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j,$$

for $\sigma_0 \in \Sigma_{2d}$, $\sigma_j \in \Sigma_{2d-2n_{c_j}}$. Assume $f = 2x + x^2$ and $d = 1$, to prove that f is always desirable in Ω we must show that

$$f = 2x + x^2 = [1, x]Q_0[1, x]^T + q_1x + q_2(1 - x),$$

with $Q_0, q_i \geq 0$ (Q_0 is a matrix, q_i are scalars). By equating the coefficients of the polynomials we find the solution $q_1 = 2, q_2 = 0$ and

$$Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since $Q_0 \geq 0$ is positive semi-definite and $q_1, q_2 \geq 0$, this shows that f is nonnegative in Ω . Instead, the polynomial $f = \frac{1}{8} - x(1 - x)$ cannot be written as $[1, x]Q_0[1, x]^T + q_1x + q_2(1 - x)$ with $Q_0, q_i \geq 0$. This polynomial is negative for $x = 1/2$.

Example 8 (Markov's inequality again). In Example 1 we have shown how to derive Markov's inequality from ADG:

$$\begin{aligned} \overline{E}(I_{\{[u, \infty)\}}) &= \inf_{\lambda_i \in \mathbb{R}} \lambda_0 \\ \text{s.t.} & \\ \lambda_0 + \lambda_1(x - m) &\geq I_{\{[u, \infty)\}}(x), \quad \forall x \in \Omega. \end{aligned} \quad (35)$$

Note the presence of the indicator function that is not a polynomial. However, the indicator is a piecewise polynomial and, therefore, the above problem can be rewritten as

$$\begin{aligned} \overline{E}(I_{\{[u, \infty)\}}) &= \inf_{\lambda_i \in \mathbb{R}} \lambda_0 \\ \text{s.t.} & \\ \lambda_0 + \lambda_1(x - m) - 1 &\geq 0, \quad \forall x \in [u, x_{\max}], \\ \lambda_0 + \lambda_1(x - m) &\geq 0, \quad \forall x \in [0, u]. \end{aligned} \quad (36)$$

Assume that $u \in [0, x_{\max}]$, we can exploit the results of this section and rewrite the BADG formulation of the above problem as

$$\begin{aligned} \inf_{\lambda_i \in \mathbb{R}, \sigma_j} \lambda_0 \\ \text{s.t.} \\ \lambda_0 + (x - m)\lambda_1 - 1 &= \sigma_0(x) + \sigma_1(x)(x - u) + \sigma_2(x)(x_{\max} - x), \quad \forall x \in \mathbb{R}, \\ \lambda_0 + (x - m)\lambda_1 &= \sigma_3(x) + \sigma_4(x)x + \sigma_5(x)(u - x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

where $\sigma_i(x) \in \Sigma_{2(d-1)}$ for $i = 1, 2, 4, 5$ and $\sigma_i(x) \in \Sigma_{2d}$ for $i = 0, 3$. It can be verified numerically that for $d \geq 2$ and $m < u$, the solution of the above problem is equal to m/u and, therefore, it coincides with that of ADG. For $m \geq u$ ($\lambda_0 = 1, \lambda_1 = 0, \sigma_j = 0$), the infimum is 1 same as ADG.

7.1. Duality

We now extend the duality established in Section 5.1 to the case of BADG defined on semi-algebraic sets. As before, the first crucial step consists in establishing the following result.

Proposition 12. Let $\mathcal{C} = \Xi_{2d}$. Then its dual is

$$\mathcal{C}^\bullet = \left\{ y \in \mathbb{R}^{S_n(2d)} : M_{n,d}(y) \geq 0, M_{n,d-n_{c_j}}(c_j y) \geq 0, \forall c_j \in C \right\}, \quad (37)$$

where $M_{n,r}(cy) := L(c(x_1, \dots, x_n)v_r(x_1, \dots, x_n)v_r(x_1, \dots, x_n)^\top)$.

Proof. The proof is structurally the same as the one for Proposition 9. The inclusion from right to left being easy, for the other inclusion we reason as follows. First of all, notice that elements of Ξ_{2d} are combinations of polynomials of the form $\sigma_j(x_1, \dots, x_n)c_j(x_1, \dots, x_n)$ with $\sigma_j \in \Sigma_{2(d-n_{c_j})}$. Any $\sigma_j \in \Sigma_{2(d-n_{c_j})}$ can be written as $v_{d-n_{c_j}}(x_1, \dots, x_n)^\top Q v_{d-n_{c_j}}(x_1, \dots, x_n)$ (see Eq. (15)). From Equation (23), $c_j(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)^\top Q v_{d-n_{c_j}}(x_1, \dots, x_n)$ is equal to $\text{Tr}(Qc(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n))$ with $Q \geq 0$. Because of linearity of L and trace

$$\begin{aligned} & L(\text{Tr}(Qc(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)^\top)) \\ &= \text{Tr}(QL(c(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)^\top)) \\ &= \text{Tr}(QM_{n,d-n_{c_j}}(c_j y)), \end{aligned}$$

where $M_{n,d-n_{c_j}}(c_j y) = L(c(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)v_{d-n_{c_j}}(x_1, \dots, x_n)^\top)$. This means that $\text{Tr}(QM_{n,d-n_{c_j}}(c_j y)) \geq 0 \ \forall Q \geq 0$, and therefore $M_{n,d-n_{c_j}}(c_j y) \geq 0$ for every c_j . We conclude by considering that $c_0(x) = 1$. \square

The matrix $M_{n,r}(c_j y)$ is called localizing matrix by Lasserre (2009).

As an immediate consequence of Proposition 12 and the properties of $(\cdot)^\bullet$, it is then possible to verify an analogous of Proposition 10. Since Proposition 11 also holds for BADG defined on semi-algebraic sets, by reasoning exactly as in Theorem 5, we therefore can prove the following.

Theorem 7. *The map*

$$\mathcal{C} \mapsto \mathcal{C}^\bullet \cap \mathcal{S}$$

is a bijection between BADGs in the semi-algebraic set Ω and closed convex subsets of \mathcal{S} . We can therefore identify the dual of a BADG \mathcal{C} in the semi-algebraic set Ω with

$$\mathcal{C}^\bullet = \left\{ y \in \mathbb{R}^{s_n(2d)} : L(g) \geq 0, L(1) = 1, M_{n,d-n_{c_j}}(c_j y) \geq 0, \forall c_j \in \mathcal{C}, M_{n,d}(y) \geq 0, \forall g \in \mathcal{C} \right\}, \quad (38)$$

where $L(g)$ is completely determined by y via the definition (20).

To understand the above dual set, let us consider again the following example.

Example 9. *Let us consider the set*

$$\Omega = \{x \in \mathbb{R} : 1 - x \geq 0, x \geq 0\}. \quad (39)$$

Assume that Alice is in a state of complete ignorance and that $d = 1$. Then the dual (38) is

$$\mathcal{C}^\bullet = \left\{ y_0, y_1, y_2, y_3 : \right. \quad (40)$$

$$\left. y_0 = 1, M_{1,1}(y) = \begin{bmatrix} y_0 & y_1 \\ y_1 & y_2 \end{bmatrix} \geq 0, M_{1,0}(c_1 y) = y_1 \geq 0, M_{1,0}(c_2 y) = 1 - y_1 \geq 0 \right\}. \quad (41)$$

By interpreting $M_{1,0}(c_1 y), M_{1,0}(c_2 y)$ as truncated moment matrices, we can see

$$M_{1,0}(c_1 y) = L(x) = \int x d\mu \geq 0, M_{1,0}(c_2 y) = L(1-x) = \int (1-x) d\mu \geq 0.$$

Hence, the assessment $x \in \Omega = [0, 1]$ has been relaxed in BADG to $E[x] \in [0, 1]$.

7.2. Convergence of BADG to ADG

If we consider Theorem 4 then we can notice that, for fixed G , the set $\text{posi}(G \cup \Sigma_{2d})$ depends on the degree d of the SOS polynomials Σ_{2d} . By increasing d we add more nonnegative gambles and, therefore, enlarge the cone \mathcal{C} . We can then ask: what happens if we increase $d \rightarrow \infty$?

Let us assume that the semi-algebraic set Ω in (31) is compact. The compactness implies that polynomial gambles defined on Ω are now bounded.

Proposition 13 (Schmüdgen 1991). *Let Ω be as in (31) and compact. If f is strictly positive on Ω then there exist $\sigma_j \in \Sigma[x_1, \dots, x_n]$ such that*

$$f = \sum_{J \subseteq \{1, \dots, |C|\}} \sigma_J c_J, \quad (42)$$

where $c_J = \prod_{j \in J} c_j$.

Since σ_J is SOS and so nonnegative, we know that $\sigma_J c_i$, $\sigma_J c_i c_j$, $\sigma_J c_i c_j c_k$ etc. are nonnegative in Ω . This set is the convex sub-cone of \mathcal{L}_Ω^+ we introduced previously in Section 7. Schmüdgen (1991) proved that any strictly non-negative polynomial f on Ω can be written as $\sum_{J \subseteq \{1, \dots, |C|\}} \sigma_J c_J$ for some SOS σ_J . The problem with this result is that the sum on the right hand side has an exponential number of terms. By imposing a further assumption on Ω , we can make a major simplification.

We first define the following convex cone generated by the family of polynomials $c_j(x)$ (Lasserre, 2009, Sec 2.5):

$$\Xi = \left\{ \sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j : \sigma_j \in \Sigma[x_1, \dots, x_n] \right\},$$

where this time we are not restricting the degree of the SOS polynomials.

Proposition 14 (Putinar 1993). *Assume that there exists a polynomial $u \in \Xi$ such that the level set $\{x : \mathbb{R}^n : u(x) \geq 0\}$ is compact. Let Ω be as in (31). If f is a strictly positive polynomial in Ω then $f \in \Xi$, i.e., there exist $\sigma_j \in \Sigma[x_1, \dots, x_n]$ such that*

$$f = \sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j. \quad (43)$$

This is a very general and powerful proposition and shows that our Definition 6 of BADG in Ω is not restrictive: for any strictly positive polynomial f on Ω there exist SOS polynomials such that $f = \sigma_0 + \sum_{j=1}^{|C|} \sigma_j c_j$. Note that, since Ω is compact, if we know a scalar $d > 0$ such that $\Omega \subset \{x : \mathbb{R}^n : \|x\| \leq d\}$ then we can add the constraint $\|x\| \leq d$ to Ω without changing Ω . With this new representation, Ξ satisfies the assumption in Proposition 14 (Lasserre, 2009, Sec 2.5).

Proposition 15. *Given the set G of gambles Alice finds desirable, a semi-algebraic set Ω satisfying the assumption in Proposition 14. Assume that \mathcal{H} avoids sure loss. Then for every polynomial f , BADG converges to ADG for $d \rightarrow \infty$ in the sense that when $\underline{E}^*(f)$ is finite then $\underline{E}^*(f) \rightarrow \underline{E}(f)$ (from below) (Lasserre, 2009, Th. 4.1).*

However, we have already shown, for instance in the Covariance Inequality example, that it often happens that $\underline{E}^*(f) = \underline{E}(f)$ even for finite d (Lasserre, 2009, Sec. 4.1).

8. Updating

We assume that Alice considers an event “indicated” by a certain polynomial $h(x) \geq 0$, meaning that Alice knows that x belongs to the set $A = \{x \in \mathbb{R}^n : h(x) \geq 0\}$. In ADG we will use this information to update (condition) her set of desirable gambles based on $A \subseteq \Omega$ (Walley, 1991; Couso & Moral, 2011). Let $G \subseteq \mathcal{L}$ be finite, and $\mathcal{H} = \text{posi}(G \cup \mathcal{L}^+)$. Then $\mathcal{H}_A = \{g \in \mathcal{L} : g I_A \in \mathcal{H}\}$, where I_A is the indicator function on A . From (6), it then follows that the conditional lower prevision of a gamble f is

$$\begin{aligned} & \sup_{\lambda_j \geq 0, \lambda_0} \lambda_0 \\ & \text{s.t.} \\ & (f - \lambda_0) I_A - \sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0, \quad \forall x \in \Omega, \end{aligned}$$

which is equivalent to

$$\begin{aligned}
& \sup_{\lambda_j \geq 0, \lambda_0} \lambda_0 \\
& \text{s.t.} \\
& f - \lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0, \quad \forall x \in A, \\
& \quad - \sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0, \quad \forall x \notin A.
\end{aligned} \tag{44}$$

By writing $\neg A := \Omega \setminus A$, the dual of \mathcal{K}_A coincides with

$$\begin{aligned}
& \inf_{\mu_1 \in \mathcal{M}(A)^+, \mu_2 \in \mathcal{M}(\neg A)^+} \int f d\mu_1 \\
& \text{s.t.} \\
& \int_A d\mu_1 = 1 \\
& \int_A g_j d\mu_1 + \int_{\neg A} g_j d\mu_2 \geq 0, \quad \forall j = 1, \dots, |G|.
\end{aligned} \tag{45}$$

Proposition 16. Assume that $\bar{E}(I_A) > 0$ then the above optimisation problem is equivalent to

$$\begin{aligned}
\sup_{v \in \mathbb{R}} v : \quad & \inf_{\mu \in \mathcal{M}(\Omega)^+} \int (f - v) I_A d\mu \geq 0 \\
& \text{s.t.} \\
& \int d\mu = 1 \\
& \int g_j d\mu \geq 0, \quad \forall j = 1, \dots, |G|.
\end{aligned} \tag{46}$$

This is also called regular extension (Walley, 1991, Appendix J).

How do we do that in the BADG framework? In BADG we cannot completely use this information because again Σ_{2d} does not include indicator functions. However, we can still exploit the information in A in a weaker way as shown in the previous section. In fact, if we know that $h(x) \geq 0$, then we also know:

$$\begin{aligned}
& \sigma_1(x)h(x) \geq 0 \quad \forall x \in A, \\
& -\sigma_2(x)h(x) \geq 0 \quad \forall x \in \neg A,
\end{aligned}$$

for $\sigma_i \in \Sigma_{2(d-n_h)}$, where the degree of $h(x)$ is $2n_h$ if even or $2n_h - 1$ if odd (so that the degree of $\sigma_i(x)h(x)$ is less than $2d$). Hence, a possible way to define updating in BADG is as follows.

Definition 7. Let G be a finite subset of \mathcal{L}_{2d} , and $\mathcal{C} = \text{posi}(G \cup \Xi_{2d})$ be a set of BADG in Ω . Given the event $A = \{x \in \Omega : h(x) \geq 0\}$ for some polynomial $h(x)$, of degree $2n_h$ if even or $2n_h - 1$ if odd, then, the set \mathcal{C}_A that includes all the gambles $f \in \mathcal{L}_{2d}$ such that there exist $\lambda_i \geq 0$, with $i = 1, \dots, |G|$, and $\sigma_{i0} \in \Sigma_{2d}$, $\sigma_{ij} \in \Sigma_{2(d-n_c_j)}$, $\sigma_a, \sigma_b \in \Sigma_{2(d-n_h)}$:

$$f - \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h \quad \text{and} \quad - \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{20} + \sum_{j=1}^{|C|} \sigma_{2j} c_j - \sigma_b h \tag{47}$$

is called the **updated set of desirable gambles based on A** .

The above Definition is consistent with that in (44), since the condition $f - \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h$ is sufficient for

$$f - \sum_{i=1}^{|G|} \lambda_i g_i \geq 0, \quad \forall x \in A \subseteq \Omega.$$

In fact, given that $\sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h$ is nonnegative in $A \subseteq \Omega$, if we can write $f - \sum_{i=1}^{|G|} \lambda_i g_i$ as $\sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h$ then this implies that $f - \sum_{i=1}^{|G|} \lambda_i g_i$ is also nonnegative in $A \subseteq \Omega$. Similarly, the condition $- \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{20} +$

$\sum_{j=1}^{|C|} \sigma_{2j} c_j - \sigma_b h$ is sufficient for

$$-\sum_{i=1}^{|G|} \lambda_i g_i \geq 0, \forall x \in \neg A.$$

Observe that, in the state of full ignorance, since G is empty, there is only one constraint $f = \sigma_0 + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h$.

Theorem 8. *Assume that $\mathcal{C}|_A$ is BADG in Ω . Then it holds that*

$$\begin{aligned} \mathcal{C}|_A^\bullet = & \left\{ y \in \mathbb{R}^{s_n(d)} : \exists z \in \mathbb{R}^{s_n(d)} \text{ such that} \right. \\ & M_{n,d}(y), M_{n,d-n_{c_j}}(c_j y), M_{n,d}(z), M_{n,d-n_{c_j}}(c_j z) \geq 0 \forall c_j, \\ & M_{n,d-n_h}(hy), M_{n,d-n_h}(-hz) \geq 0, L_y(1) = 1, \\ & \left. L_y(g) + L_z(g) \geq 0, \forall g = 1, \dots, |G| \right\}. \end{aligned}$$

Proof. The argument of the proof is similar to that of Proposition 12. Note in fact that to define the dual of $f - \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h$ we can exploit Proposition 12 and account for the presence of the additional constraint $h \geq 0$. The variable z is introduced to define the dual of $-\sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{20} + \sum_{j=1}^{|C|} \sigma_{2j} c_j - \sigma_b h$. The constraint $L_y(g) + L_z(g) \geq 0$ connects the two duals and arises due to the presence of the term $-\sum_{i=1}^{|G|} \lambda_i g_i$ in both equalities (47). \square

To understand the above dual set, we can compare it with (45). The vector y has the same role of μ_1 and z that of μ_2 . The constraints $M_{n,d}(y), M_{n,d}(z)$ are the bounded rationality analogous of $\mu_1, \mu_2 \in \mathcal{M}^+(\mathbb{R}^n)$. The constraints $M_{n,d-n_{c_j}}(c_j y), M_{n,d-n_{c_j}}(c_j z) \geq 0$ are the bounded rationality analogous of the support constraints $\mu_1 \in \mathcal{M}^+(\Omega)$ and $\mu_2 \in \mathcal{M}^+(\Omega)$. The constraints $M_{n,d-n_h}(hy), M_{n,d-n_h}(-hz) \geq 0$ are the bounded rationality representation of the constraints $\mu_1 \in \mathcal{M}^+(A)$ and $\mu_2 \in \mathcal{M}^+(\neg A)$. Finally, $L_y(g) + L_z(g) \geq 0$ is equivalent to $\int_A g_j d\mu_1 + \int_{\neg A} g_j d\mu_2 \geq 0$.

Theorem 9. *Let G be a finite subset of \mathcal{L}_{2d} , and $A = \{x \in \Omega : h(x) \geq 0\}$. Assume that $\mathcal{K} = \text{posi}(G \cup \mathcal{L}^+)$ avoids sure loss and let $f \in \mathcal{L}_{2d}$. Then we have that $\underline{E}_{\mathcal{C}|_A}(f) \leq \underline{E}_{\mathcal{K}|_A}(f)$ where $\mathcal{C} = \text{posi}(G \cup \Xi_{2d})$.*

Proof. From the definition of conditioning for ADG we aim to find the supremum λ_0 such that $(f - \lambda_0)I_A - \sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0 \forall x \in \mathbb{R}^n$. It can be rewritten as the two constraints on the left and relaxed to the constraints on the right:

$$\begin{aligned} -\sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0 \forall x \in \neg A, & \quad -\sum_{j=1}^{|G|} \lambda_j g_j(x) = \sigma_{20} + \sum_{j=1}^{|C|} \sigma_{2j} c_j - \sigma_b h, \\ f - \lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j(x) \geq 0 \forall x \in A, & \quad f - \lambda_0 - \sum_{j=1}^{|G|} \lambda_j g_j(x) = \sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sigma_a h. \end{aligned}$$

where the equalities on the right must hold $\forall x \in \mathbb{R}^n$. \square

In case the set A is defined by several polynomial constraints $A = \{h_1(x) \geq 0, \dots, h_{|A|}(x) \geq 0\}$, we cannot use (47) because in general we cannot write $\neg A$ as a single polynomial constraint. However, we can relax (47) to:

$$f - \sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{10} + \sum_{j=1}^{|C|} \sigma_{1j} c_j + \sum_{i=1}^{|A|} \sigma_{ai} h_i \quad \text{and} \quad -\sum_{i=1}^{|G|} \lambda_i g_i = \sigma_{20} + \sum_{j=1}^{|C|} \sigma_{2j} c_j, \quad (48)$$

which is a conservative approximation. It can actually be proven that the complement of a semi-algebraic set is the union of semi-algebraic sets (Tarski, 1951; Seidenberg, 1954). Hence, the exact way to consider the constraints $\neg A$ is to translate them in a bunch of SDP problems.

ask	53.8	49.5	45.2	41.1	29.3	25.7	22.3	19.1	16.2	13.6	11.3	9.2	7.5	6.1	4.9	3.9	3.2	2.15	1.55	1.15	0.90	0.3
bid	53.3	49	44.8	40.6	28.9	25.3	21.9	18.7	15.9	13.3	11	8.9	7.2	5.8	4.6	3.7	3	2	1.40	1	0.75	0.2
strike	2490	2495	2500	2505	2520	2525	2530	2535	2540	2545	2550	2555	2560	2565	2570	2575	2580	2590	2600	2610	2620	2675

Table 1: Ask and bid price for a call option on the S&P500 index: maturity 30days, quote day 2017-10-03.

9. Case study: European Options

As an example of application of BADG, we consider a problem from finance. An European call option on an underlying security with strike k and maturity T gives the holder the option of buying the underlying security at price k at time T . If the price S_T is more than k , then the holder will exercise the option and make a profit of $S_T - k$. Conversely, if it is less than k , the holder will not exercise and does not make a profit. Thus, the payoff of this option is $\max(S_T - k, 0)$. Since options are traded, a key problem in financial economics is to determine the belief of the market about the future value of S_T from the ask and bid⁸ prices of these options. Table 1 shows the ask and bid price for 22 call options on the S&P500 index. What does the first column of the table mean? It means that “the market” believes that the gambles $\max(S_T - 2490, 0) - 53.3$ and $53.8 - \max(S_T - 2490, 0)$ are desirable, since there exists someone that is willing to sell the option $\max(S_T - 2490, 0)$ at price 53.8 and to buy it at price 53.3. As inference, we aim to compute the market’s selling and buying price for the gamble $f = I_{[c, \infty)}(S_T)$ for some $c \in \mathbb{R}$. In this case, the set of desirable gambles includes 44 gambles:

$$G = \{\max(S_T - 2490, 0) - 53.3, 53.8 - \max(S_T - 2490, 0), \dots, \max(S_T - 2675, 0) - 0.2, 0.3 - \max(S_T - 2675, 0)\}.$$

Note that, for simplicity, we have assumed that the discount factor is one.⁹ Moreover, observe that the gambles in G and f are piecewise polynomials. We aim to apply BADG to solve this problem by exploiting the same trick used in Example 8. Consider for instance the case G only includes $\max(S_T - 2490, 0) - 53.3, 53.8 - \max(S_T - 2490, 0), \max(S_T - 2495, 0) - 49, 49.5 - \max(S_T - 2495, 0)$ and $c = 2490$, then the lower prevision of f can be computed in BADG as:

$$\begin{aligned} & \sup_{\lambda_0 \in \mathbb{R}, \lambda_j \geq 0} \lambda_0 \\ & \text{s.t.} \\ & f - \lambda_0 - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 = \sigma_0(S_T) + (2490 - S_T)\sigma_3(S_T), \\ & f - \lambda_0 + (S_T - 2490)\lambda_1 - (S_T - 2490)\lambda_2 - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 \\ & = \sigma_1(S_T) + (S_T - 2490)\sigma_4(S_T) + (2495 - S_T)\sigma_5(S_T), \\ & f - \lambda_0 + (S_T - 2490)\lambda_1 - (S_T - 2490)\lambda_2 + (S_T - 2495)\lambda_3 - (S_T - 2495)\lambda_4 \\ & - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 = \sigma_2(S_T) + (S_T - 2495)\sigma_6(S_T), \end{aligned} \tag{49}$$

which, exploiting the definition of f , is equal to

$$\begin{aligned} & \sup_{\lambda_0 \in \mathbb{R}, \lambda_j \geq 0} \lambda_0 \\ & \text{s.t.} \\ & -\lambda_0 - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 = \sigma_0(S_T) + (2490 - S_T)\sigma_3(S_T), \\ & 1 - \lambda_0 + (S_T - 2490)\lambda_1 - (S_T - 2490)\lambda_2 - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 \\ & = \sigma_1(S_T) + (S_T - 2490)\sigma_4(S_T) + (2495 - S_T)\sigma_5(S_T), \\ & 1 - \lambda_0 + (S_T - 2490)\lambda_1 - (S_T - 2490)\lambda_2 + (S_T - 2495)\lambda_3 - (S_T - 2495)\lambda_4 \\ & - 53.8\lambda_1 + 53.3\lambda_2 - 49.5\lambda_3 + 49\lambda_4 = \sigma_2(S_T) + (S_T - 2495)\sigma_6(S_T), \end{aligned} \tag{50}$$

⁸The bid price is the max price that a buyer is willing to pay for a security. The ask price is the min price that a seller is willing to receive.

⁹The discount factor is the factor by which a future cash flow must be multiplied in order to obtain the present value.

with $\sigma_i(S_T) \in \Sigma_0$ for $i = 3, 4, 5, 6$ and $\sigma_i(S_T) \in \Sigma_2$ for $i = 0, 1, 2$. This approach can be generalised to all 44 gambles and allows us to deal with piecewise polynomials.

The application of SOS polynomials to European option pricing was first proposed by Lasserre et al. (2006). The authors consider the problem of pricing an option given information (moments) on the probability density function of S_T . Here, we are considering the inverse problem and we are interested in studying it from a desirable gambles point of view and, in particular, to investigate the effect of the updating in the inference.

In particular, for this example, the BADG lower and upper previsions of f are shown in Figure 3. It is worth noticing that they coincide with those computed using ADG – we have verified it numerically by discretising S_T and solving a linear programming problem. Note that the discretisation approach can only be used when the number of variables is small and, in any case, provides only an inner approximation of the lower and upper previsions. However, since in this case BADG and ADG coincide, we can refer to these lower and upper previsions as the lower and upper probabilities of the event $S_T > c$.

Assume that we aim to update our inference given the information “ $S_T \geq 2540$ is true”, meaning that Alice knows that x belongs to the set $A = \{S_T \in \mathbb{R} : S_T - 2540 \geq 0\}$. We can apply the approach discussed in Section 8 and compute an updated set of desirable gambles. The corresponding lower and upper probabilities for the event $S_T > c$ are shown in Figure 4 (right, blue) together with the previous lower and upper probability for comparison.

Options’ data includes other information apart from bid and ask prices, such as trading volume for the day. We can use such information for updating, for example by using the trading volume to build a weighting function across the strikes. An example of weighting function $W(S_T)$ is shown in Figure 4 (left). We can then compute an updated BADG by replacing $f(S_T) - \lambda_0$ in (49) with $(f(S_T) - \lambda_0)W(S_T)$. This is another way of defining an updating rule in BADG that is similar to updating with probability density functions in standard probability. The updated lower probability is shown in Figure 4 (right, green).

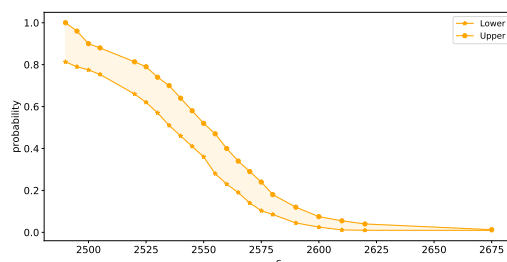


Figure 3: Lower and upper probability that $S_T > c$ for BADG.

Acknowledgement

We thank the anonymous reviewers for their careful reading of the manuscript. Their insightful comments and suggestions have greatly helped improve and clarify this work.

This work was partially supported by the Swiss NRP 75 Big Data grant no. 407540-167199.

Aliprantis, C., & Border, K. (2007). *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer.

Benavoli, A., Facchini, A., Piga, D., & Zaffalon, M. (2017a). Sos for bounded rationality. In *Proc. ISIPTA’17 Int. Symposium on Imprecise Probability: Theories and Applications*, (pp. 1–12). PMLR.

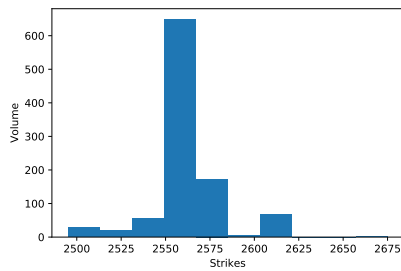
Benavoli, A., Facchini, A., Zaffalon, M., & Vicente-Prez, J. (2017b). A polarity theory for sets of desirable gambles. In A. Antonucci, G. Corani, I. Couso, & S. Destercke (Eds.), *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications* (pp. 37–48). PMLR volume 62 of *Proceedings of Machine Learning Research*.

Benavoli, A., & Piga, D. (2016). A probabilistic interpretation of set-membership filtering: Application to polynomial systems through polytopic bounding. *Automatica*, 70, 158 – 172.

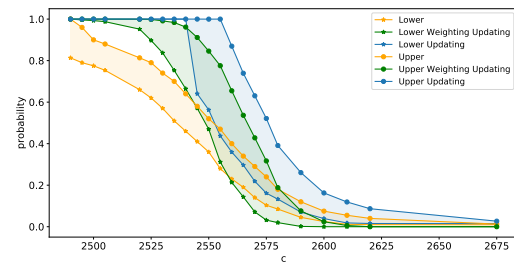
Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

de Cooman, G., & Quaeghebeur, E. (2012). Exchangeability and sets of desirable gambles. *International Journal of Approximate Reasoning*, 53, 363–395.

Couso, I., & Moral, S. (2011). Sets of desirable gambles: Conditioning, representation, and precise probabilities. *International Journal of Approximate Reasoning*, 52, 1034–1055.



(a) Weighting function



(b) Lower/Upper and updated lower/upper probabilities that $S_T > c$ for BADG.

Figure 4: Updating

- de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7, 1–68.
- Hilbert, D. (1888). Über die darstellung definiten formen als summe von formenquadraten. *Mathematische Annalen*, 32, 342–350.
- Lasserre, J. B. (2009). *Moments, positive polynomials and their applications* volume 1. World Scientific.
- Lasserre, J.-B., Prieto-Rumeau, T., & Zervos, M. (2006). Pricing a class of exotic options via moments and sdp relaxations. *Mathematical Finance*, 16, 469–494.
- Miranda, E. (2008). A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48, 628–658.
- Miranda, E., & Zaffalon, M. (2010). Notes on desirability and conditional lower previsions. *Annals of Mathematics and Artificial Intelligence*, 60, 251–309.
- Pelessoni, R., & Vicig, P. (2016). 2-coherent and 2-convex conditional lower previsions. *International Journal of Approximate Reasoning*, 77, 66–86.
- Piga, D., & Benavoli, A. (2017). A unified framework for deterministic and probabilistic d-stability analysis of uncertain polynomial matrices. *Automatic Control, IEEE Transactions on*, 62, 5437–5444.
- Putinar, M. (1993). Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42, 969–984.
- Robinson, R. M. (1969). Some definite polynomials which are not sums of squares of real polynomials. In *Notices of the American Mathematical Society* (p. 554). volume 16.
- Schervish, M. J., Seidenfeld, T., & Kadane, J. B. (2000). How sets of coherent probabilities may serve as models for degrees of incoherence. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8, 347–355.
- Schmüdgen, K. (1991). The k-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289, 203–206.
- Seidenberg, A. (1954). A new decision method for elementary algebra. *Annals of Mathematics*, (pp. 365–374).
- Seidenfeld, T., Schervish, M. J., & Kadane, J. B. (1990). Decisions without ordering. In W. Sieg (Ed.), *Acting and reflecting* (pp. 143–170). Dordrecht: Kluwer volume 211 of *Synthese Library*.
- Simon, H. A. (1957). *Models of man: social and rational; mathematical essays on rational human behavior in society setting*. Wiley.
- Tarski, A. (1951). A decision method for elementary algebra and geometry, .
- Troffaes, M. C. M., & de Cooman, G. (2003). Extension of coherent lower previsions to unbounded random variables. *Intelligent systems for information processing: from representation to applications*, (pp. 277–288).
- Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. New York: Chapman and Hall.
- Walley, P., Pelessoni, R., & Vicig, P. (2004). Direct algorithms for checking consistency and making inferences from conditional probability assessments. *Journal of Statistical Planning and Inference*, 126, 119–151.
- Williams, P. M. (1975). *Notes on conditional previsions*. Technical Report School of Mathematical and Physical Science University of Sussex, UK.