Bernstein's socks, polynomial-time
provable coherence and entanglement

Alessio Benavoli   ALESSIO.BENAVOLI@UL.IE
Department of Computer Science and Information Systems (CSIS), University of Limerick, Ireland.
Alessandro Facchini   ALESSANDRO.FACCHINI@IDSIA.CH
Marco Zaffalon   ZAFFALON@IDSIA.CH
Dalle Molle Institute for Artificial Intelligence (IDSIA), Galleria 2, 6928 Manno Switzerland

Abstract

We recently introduced a bounded rationality approach for the theory of desirable gambles. It is based on the unique requirement that being nonnegative for a gamble has to be defined so that it can be provable in polynomial time. In this paper we continue to investigate properties of this class of models. In particular we verify that the space of Bernstein polynomials in which nonnegativity is specified by the Krivine-Vasilescu certificate is yet another instance of this theory. As a consequence, we show how it is possible to construct in it a thought experiment uncovering entanglement with classical (hence non quantum) coins.

1. Introduction

In a recent paper (Benavoli et al., 2019), we have shown that Quantum Theory (QT) is a theory of bounded rationality (Simon, 1957) based on a different notion of nonnegativity. This is tantamount to changing the class of gambles that should always be desired in such a way that the consistency problem becomes *provable in polynomial time* (we have called it *P-coherence*). Conversely, in the same settings, classical probability (standard "almost desirability" (Walley, 1991)) is NP-hard. As a consequence, we have thence proved that the only physics' axiom in QT is computational tractability, yielding all its weirdness (different logic of events, negative probabilities, and entanglement).

Interestingly, it turns out that entanglement is not peculiar to QT but an inherent characteristic of bounded rationality for desirable gambles based on P-coherence, a model first introduced in Benavoli et al. (2017a) and implemented using sum-of-squares polynomials in the real numbers.

A first goal of the present paper is to better understand the structural properties of P-coherence. An elegant way of doing this is to look at the linear space of gambles ($\mathscr{L}$) as an algebra of formulas, and thus define a logic on it. In doing so, we verify that a suf-

ficient condition for the reduction of P-coherence to classical logical consistency (that is the existence of a non-derivable formula from the considered set of assessments) is for the set of tautologies to satisfy a certain "pullup" property.

As a second goal of our work, we provide yet another instance of a P-coherence model by using the so-called Krivine-Vasilescu nonnegativity certificate for polynomials. We show that, by focusing on polynomials defined on the *simplex of probability*, this notion of nonnegativity defines the so-called *Bernstein nonnegative polynomials*, a class whose cone, when polynomials are of any degree d, has been introduced in De Cooman et al. (2009) to generalise de Finetti's representation result for exchangeable events (see also (de Cooman and Quaeghebeur, 2012; De Bock and De Cooman, 2012; De Cooman et al., 2015)).

Finally, we restrict our attention to a finite degree d, compute the dual of a (Bernstein) P-coherent set of desirable gambles and illustrate how P-coherence can provide an example of *entanglement with classical coins*: two coins that always land HH or TT, but for which it is not possible to find a "common cause" (Einstein et al., 1935) (a classical correlation model) that explains these results. The latter is done by deriving a Bell's type inequality in the Bernstein's world.

The title of this paper is freely inspired by Bell's work "Bertlmann's socks and the nature of reality" (Bell, 1981).

2. The logic of desirability

Let Ω denote the possibility space of an experiment (e.g., $\{Head, Tail\}$ or $\mathbb{R}^n$). A *gamble* g on Ω is a bounded real-valued function of Ω, interpreted as an uncertain reward. Accepting a gamble g by an agent, Alice, is regarded as a commitment to receive, or pay (depending on the sign), $g(\omega)$ *utiles*[1] whenever $\omega \in \Omega$

---

1. Abstract units of utility, we can approximately identify it with money provided we deal with small amounts of it (de Finetti, 1974, Sec. 3.2.5)

occurs. We denote by $\mathscr{L}$ the set of all the gambles on Ω, the subset of all *nonnegative* gambles, that is, of gambles for which Alice is never expected to lose utiles, is denoted as $\mathscr{L}^{\geq} := \{g \in \mathscr{L} : \inf g \geq 0\}$ (analogously negative gambles are denoted as $\mathscr{L}^{<} := \{g \in \mathscr{L} : \sup g < 0\}$). In the following, with $\mathscr{G} \subset \mathscr{L}$ we denote a finite set of gambles that Alice finds desirable (we will comment on the case when $\mathscr{G}$ may not be finite): these are the gambles that she is willing to accept and thus commits herself to the corresponding transactions.

The crucial question is now to provide a criterion for a set $\mathscr{G}$ of gambles representing assessments of desirability to be called *rational*. Intuitively Alice is rational if she avoids sure losses (also called Dutch books or arbitrages): that is, if, by considering the implications of what she finds desirable, she is not forced to find desirable a negative gamble. An elegant way to formalise this intuition is to see $\mathscr{L}$ as an algebra of formulas, and thus define a logic on it. Based on it, we can thus formulate rationality as *logical consistency*.

In the theory of (almost) desirable gambles, we may proceed as described in the next subsection.

2.1. A Gentzen system

First of all we introduce some basic notions from logic.

A *sequent* is a pair $(\mathscr{G}, g)$, written $\mathscr{G} \triangleright g$, where $\mathscr{G}$ is a set (possibly empty) of gambles, and g is a gamble.

One can read a sequent $\mathscr{G} \triangleright g$ as saying "whenever the gambles in $\mathscr{G}$ are desirable for (by) Alice, the gamble g is also desirable for (by) Alice".

A *Gentzen-style rule* is a pair which consists of a set $\{\mathscr{G}_i \triangleright g_i \mid i \leq \alpha\}$ of sequents, called the premisses of the rule, and a sequent $\mathscr{G} \triangleright g$, called the consequence of the rule and therefore which follows from the set according to the rule. We let $\alpha \in \mathbb{N} \cup \{\infty\}$, and thus do not rule out the fact that a rule may be infinitary. A rule r is written symbolically in the form

$$(\text{r}) \quad \frac{\{\mathscr{G}_i \triangleright g_i \mid i < \alpha\}}{\mathscr{G} \triangleright g} \ ,$$

which can be read as "if Alice finds the gamble g_i desirable, given the fact that she desires the gambles in $\mathscr{G}_i$ (for $i < \alpha$), then necessarily whenever the gambles in $\mathscr{G}$ are desirable, the gamble g is also desirable".

An *axiom* is a rule in which the set of premisses is empty.

A *system* $\mathfrak{S}$ is a set of Gentzen-style rules. We say that a sequent $\mathscr{G} \triangleright g$ is *provable* in a system $\mathfrak{S}$ from a set of sequents $\{\mathscr{G}_i \triangleright g_i \mid i \leq \alpha\}$ if there is a well-founded tree whose leaves are labelled either with axioms or with members of $\{\mathscr{G}_i \triangleright g_i \mid i \leq \alpha\}$, whose root is labelled with $\mathscr{G} \triangleright g$ and the labelling of all nodes is consistent with the rules of $\mathfrak{S}$. A sequent is

provable in $\mathfrak{S}$ if it is provable from the empty set. The set of all gambles g such that $\mathscr{G} \triangleright g$, for some set $\mathscr{G}$, will sometimes be denoted by $\mathsf{Cn}_{\mathfrak{S}}(\mathscr{G})$. Finally, we say that a set $\mathscr{G}$ is *consistent in* $\mathfrak{S}$ if there is a gamble g such that $\mathscr{G} \triangleright g$ is not provable in $\mathfrak{S}$, that is

$$\exists g \in \mathscr{L} \quad \text{such that} \quad g \notin \mathsf{Cn}_{\mathfrak{S}}(\mathscr{G}). \tag{1}$$

A set $\mathscr{G}$ is *closed* in $\mathfrak{S}$ whenever $\mathsf{Cn}_{\mathfrak{S}}(\mathscr{G}) = \mathscr{G}$, and is called a *theory* of $\mathfrak{S}$. In general, theories are denoted by $\mathscr{K}$. When $\mathsf{Cn}_{\mathfrak{S}}$ is a consequence operator, that is it is reflexive, monotone and transitive, the consistent theories of $\mathfrak{S}$ completely characterise it, in the sense that $\mathsf{Cn}_{\mathfrak{S}}(\mathscr{G})$ coincides with the intersection of all consistent theories of $\mathfrak{S}$ extending $\mathscr{G}$.

The system $\mathfrak{T}$ for the theory of (almost) desirable gambles is thus defined as follows:

Structural axiom:

Reflexivity

$$(\text{R}) \quad \frac{}{\mathscr{G} \triangleright g} \ , \text{ for } g \in \mathscr{G}$$

Logic axiom:

Accept. nonneg.

$$(\text{ANN}) \quad \frac{}{\mathscr{G} \triangleright g} \ , \text{ for } g \in \mathscr{L}^{\geq}$$

Logic rules:

Positive hull

$$(\text{L}) \quad \frac{\mathscr{G} \triangleright g \qquad \mathscr{G}' \triangleright f}{\mathscr{G}, \mathscr{G}' \triangleright \mu g + \lambda f} \ , \text{ for } \mu, \lambda > 0$$

Closure

$$(\text{C}) \quad \frac{\{\mathscr{G} \triangleright g + \varepsilon^{\ell} \mid \ell > 0\}}{\mathscr{G} \triangleright g} \ , \text{ for } \varepsilon \in (0, 1)$$

It is easy to verify that $\mathsf{Cn}_{\mathfrak{T}}$ is indeed a consequence operator on $\mathscr{L}$, and thus in particular that the usual structural rules of Gentzen systems such as *weakening* and *cut* are derivable in $\mathfrak{T}$. It is immediate to verify that logical consistency in $\mathfrak{T}$ is indeed tantamount to rationality (coherence, no arbitrage):

Theorem 1 *Let $\mathscr{G}$ be a set of assessments. It holds that $\mathsf{Cn}_{\mathfrak{T}}(\mathscr{G}) = \mathrm{cl}(\mathrm{posi}(\mathscr{G} \cup \mathscr{L}^{\geq}))$, where posi denotes the positive hull operator and cl the topological closure operator. Moreover the following conditions are equivalent to each other*

1. *$\mathscr{G}$ is logically consistent in $\mathfrak{T}$*

2. *$\mathscr{G}$ avoids negativity (sure loss), that is $\mathscr{L}^{<} \cap \mathsf{Cn}_{\mathfrak{T}}(\mathscr{G}) = \emptyset$*

3. *$-1 \notin \mathsf{Cn}_{\mathfrak{T}}(\mathscr{G})$.*

Therefore, -1 can be regarded as playing the role of the (classical) *Falsum*.

Clearly when $\mathscr{G}$ is finite, $\mathsf{Cn}_{\mathfrak{T}}(\mathscr{G})$ simply coincides with the conic hull closure of $\mathscr{G}$. From Theorem 1 we thence obtain that the following *principle of explosion* is derivable in $\mathfrak{T}$:

$$\text{(Explosion)} \quad \frac{\mathscr{G} \rhd -1}{\mathscr{G} \rhd g} \text{ , for } g \in \mathscr{L}.$$

By observing that the mathematical dual of $\mathscr{K}$ is a closed convex set of probabilities, we can then provide a *semantic* (probabilistic interpretation) to $\mathfrak{T}$:

$$\mathscr{P}(\mathscr{G}) = \left\{ \mu \in \mathsf{S} \middle| \int_{\Omega} g(\Omega) d\mu(\Omega) \geq 0, \ \forall g \in \mathscr{G} \right\}, \quad (2)$$

where $\mathsf{S} = \{\mu \in \mathscr{M} \mid \inf \mu \geq 0, \ \int_{\Omega} d\mu(\Omega) = 1\}$ is the set of all probabilities on Ω, also-called (belief) *states*, and $\mathscr{M}$ the set of all charges (a charge is a finitely additive signed-measure (Aliprantis and Border, 2007, Ch.11)) on Ω. The duality actually provides us immediately with a sound and completeness results. Indeed, say that a state $\mu \in \mathscr{S}$ is a model of a set of gambles $\mathscr{G}$ whenever $\mu \in \mathscr{P}(\mathscr{G})$. Then the following is an immediate consequence of Theorem 1 and (Benavoli et al., 2017b, Theorem 4).

Theorem 2 *For every set of gambles $\mathscr{G} \cup \{g\} \subseteq \mathscr{L}$, it holds that*

$$g \in \mathsf{Cn}_{\mathfrak{T}}(\mathscr{G}) \iff \mathscr{P}(\{g\}) \subseteq \mathscr{P}(\mathscr{G}). \quad (3)$$

In particular $\mathscr{G}$ is inconsistent iff $\mathscr{P}(\mathscr{G}) = \emptyset$ (it has no probabilistic model).

Hence, whenever an agent is coherent (that is the rationality of her behaviour is represented by a logical consistent theory in $\mathfrak{T}$), Equation (2) states that desirability corresponds to nonnegative expectation (for all probabilities in $\mathscr{P}(\mathscr{G})$). When she is incoherent, $\mathscr{P}(\mathscr{G})$ turns out to be empty.

2.2. Coherence vs logical consistency

In (1) we have defined the notion of logical consistency for a theory as being non trivial (different from the whole language) and thus, from a semantic perspective, as having a model (that is $\mathscr{P}(\mathscr{G})$ is not empty). Theorem 1 attests that -1 is tantamount to the (classical) Falsum, that is an all implying formula that has no model.

The importance of being able to reduce incoherence to logical inconsistency can be appreciated by the following argument. Assume this is not the case, the explosion principle for -1 does not hold. In particular, this means that we may be able to find two different incoherent sets of gambles (formulas). But then such

sets cannot be separated in the dual space, meaning that duality would fail in providing us with a sound and complete probabilistic semantics for the system under consideration (that is satisfying the correspondence in Equation (3) of Theorem 2).

The accent in this work to the capability of reducing coherence (also when formulated as P-coherence) to logical consistency is therefore justified by the fact that we do not want a situation of incoherence (irrationality) to represent anything else than a situation of incoherence (irrationality), and thus for which, seen as a theory of a logic, there is no "natural" model (in the dual space).

We conclude this discussion with some remarks on the theory of desirable gambles when defined without the topological closure requirement and by stating that coherence means avoiding status quo (the constant 0 gamble). In such case, one would be tempted to define the corresponding Gentzen system by just dropping rule (C) and modifying the logic axiom with the stricter constraint $g \in \mathscr{L}^{\geq} \setminus \{0\}$. However, in such case coherence would not be equivalent to logical consistency: just consider the incoherent (but closed) theory $\mathscr{L}^{\geq}$. In particular, this means that it would not be possible to prove a soundness and completeness result with respect to lexicographic probabilities, as expected from the correspondence between coherent sets of desirable gambles as just defined, and sets of lexicographic probabilities (see e.g. (Benavoli et al., 2017b, Theorem 18)). In order to recover an analog of Theorem 2, we would need to add as an extra rule a principle of explosion stating that 0 is the Falsum.

3. The complexity of inference

In light of Theorem 1, when the theory is finitely generated (that is $\mathscr{G} = \{g_1, \ldots, g_{|\mathscr{G}|}\}$), the problem of checking whether or not $\mathscr{K}$ is consistent can be formulated as the following decision problem:

$$\exists \lambda_i \geq 0 : -1 - \sum_{i=1}^{|\mathscr{G}|} \lambda_i g_i \in \mathscr{L}^{\geq}. \quad (4)$$

If the answer is "yes", then the gamble -1 belongs to $\mathscr{K}$, proving $\mathscr{K}$'s inconsistency. Actually any inference task can ultimately be reduced to a problem of the form (4): the lower prevision (expectation) of a gamble q is

$$\underline{E}(q) = \sup_{\lambda_0 \in \mathbb{R}, \lambda_i \in \mathbb{R}^{\geq}} \lambda_0 : \ q - \lambda_0 - \sum_{i=1}^{|\mathscr{G}|} \lambda_i g_i \in \mathscr{L}^{\geq}.$$

Hence, the above decision problem unveils a crucial fact: the hardness of inference in classical probability

corresponds to the hardness of evaluating the nonnegativity of a function in the corresponding space (the "nonnegativity decision problem").

When Ω is infinite, and for generic functions, the nonnegativity decision problem is *undecidable*. To avoid such an issue, we may impose restrictions on the class of allowed gambles. For instance, instead of $\mathscr{L}$, we may consider $\mathscr{L}_R$: the class of multivariate polynomials of degree at most d (we denote by $\mathscr{L}_R^{\geq} \subset \mathscr{L}_R$ the subset of nonnegative polynomials and by $\mathscr{L}_R^{<} \subset \mathscr{L}_R$ the negative ones). In doing so, by Tarski-Seidenberg quantifier elimination theory (Tarski, 1951; Seidenberg, 1954), the decision problem becomes decidable, but still intractable, being in general NP-hard. If we accept that P$\neq$NP and we require that inference should be tractable (in P), we are stuck. What to do? A solution is to change the meaning of "being nonnegative" for a function by considering a subset $\Sigma^{\geq} \subsetneq \mathscr{L}_R^{\geq}$ for which the membership problem in (4) is in P.

In other words, a computationally efficient version of the theory of desirable gambles, which we denote by $\mathfrak{T}^{\star}$, should be based on a redefinition of the logical axiom scheme, i.e., by stating that

Accept. P-nonneg.

$$^{(\mathrm{P})} \ \overline{\mathscr{G} \triangleright g} \ , \text{ for } g \in \Sigma^{\geq}.$$

We thus denote by $\mathfrak{T}^{\star}$ the Gentzen system (on $\mathscr{L}_R$) obtaining from $\mathfrak{T}$ by substituting (P) to (ANN), and call it a *P-system*.

4. Coherence and semantics

4.1. P-coherence vs consistency

However, how can be sure that we have done things properly, that $\mathfrak{T}^{\star}$ is really just a computationally efficient version of $\mathfrak{T}$? In order to do so we would like to verify that, for finite sets $\mathscr{G} = \{g_1, \ldots, g_{|\mathscr{G}|}\}$, logical consistency can be checked in polynomial time by solving:

$$\exists \lambda_i \geq 0 \quad \text{such that} \quad -1 - \sum_{i=1}^{|\mathscr{G}|} \lambda_i g_i \in \Sigma^{\geq}. \quad (5)$$

Note that, the lower prevision of a gamble q in this case is

$$\underline{E}_B(q) = \sup_{\lambda_0 \in \mathbb{R}, \lambda_i \in \mathbb{R}^{\geq}} \lambda_0$$

$$s.t. \quad (6)$$

$$q - \lambda_0 - \sum_{i=1}^{|\mathscr{G}|} \lambda_i g_i \in \Sigma^{\geq}.$$

To verify that logical consistency can be checked in polynomial time, we need to find an analog of Theorem 1, but clearly we also need to assume that $\mathscr{L}_R$ contains all constant gambles.

In what follows, we provide some sufficient conditions for this to hold. First of all, it is reasonable to ask to the new variant of "being nonnegative" (that is to the set $\Sigma^{\geq}$) to be a closed convex cone. Avoiding nonnegativity (that is coherence, rationality, no arbitrage) can now be redefined as follows

Definition 3 (P-coherence) *A set $\mathscr{C} \subseteq \mathscr{L}_R$ is P-coherent if $\Sigma^{<} \cap \mathscr{C} = \emptyset$, where $\Sigma^{<}$ is the interior of $\{g \mid -g \in \Sigma^{\geq}\}$.*

From now on, we also always make another minimal reasonable assumption of $\Sigma^{<}$ being non empty. The next result states essentially that -1 represents P-incoherence.

Proposition 4 *Let $\mathscr{G} \subseteq \mathscr{L}_R$ a set of assessments. It holds that $\mathrm{cl\,posi}(\mathscr{G} \cup \Sigma^{\geq}) = \mathsf{Cn}_{\mathfrak{T}^{\star}}(\mathscr{G})$. Moreover the following are equivalent:*

1. $-1 \notin \mathrm{posi}(\mathscr{G} \cup \Sigma^{\geq})$

2. $\mathrm{posi}(\mathscr{G} \cup \Sigma^{\geq})$ is P-coherent

3. $\mathsf{Cn}_{\mathfrak{T}^{\star}}(\mathscr{G})$ is P-coherent

Analogously with $\mathfrak{T}$, one can ask whether the class of P-coherent theories characterises the system $\mathfrak{T}^{\star}$. It turns out that if we want to be sure this to be the case (see discussion in Subsection 2.2 for the reason why we want this) we need to add some structure to $\Sigma^{\geq}$:

Proposition 5 *Let $\mathscr{G} \subseteq \mathscr{L}_R$ a set of assessments and assume that*

(pullup) *for every $f \in \mathscr{L}_R$, there is $\varepsilon > 0$ such that $f + \varepsilon \in \Sigma^{\geq}$.*

The following are then equivalent

1. $\mathsf{Cn}_{\mathfrak{T}^{\star}}(\mathscr{G})$ is P-coherent

2. $\mathsf{Cn}_{\mathfrak{T}^{\star}}(\mathscr{G})$ is logically consistent.

Hence, whenever (pullup) holds and Proposition 5 can be applied, logical consistency in $\mathfrak{T}^{\star}$ for finitely generated theories can be checked efficiently. This is the case of QT and of the family of P-systems defined on Bernstein polynomials introduced in the following sections of this work.

4.2. Probabilistic interpretation of P-systems

Interestingly, we can associate a "probabilistic" interpretation as before to the system $\mathfrak{T}^{\star}$ by computing the dual of a theory. Since $\mathscr{L}_R$ is a topological vector space, we can consider its dual space $\mathscr{L}_R^*$ of all bounded linear

functionals $L : \mathscr{L}_R \to \mathbb{R}$. Hence, with the additional condition that linear functionals preserve the unitary gamble, the dual cone of a theory $\mathscr{C} \subset \mathscr{L}_R$ is given by

$$\mathscr{C}^\circ = \{L \in \mathscr{L}_R^* \mid L(g) \geq 0, \quad L(1) = 1, \; \forall g \in \mathscr{C}\}. \quad (7)$$

Based on Equation 7 and its properties, under the pullup assumption one then gets the analogous of Theorem 2 but for $\mathfrak{T}^\star$. The question now is whether we can "massage" this result and obtain a sound and complete classical probabilistic semantics. In this aim, first notice that to $\mathscr{C}^\circ$ we can associate its extension $\mathscr{C}^\bullet$ in $\mathscr{M}$, that is, the set of all charges on Ω extending an element in $\mathscr{C}^\circ$. However, as shown in (Benavoli et al., 2019), in doing this one cannot in general provide an adequate classical probabilistic interpretations to $\mathfrak{T}^\star$, except if one allows for instance the use of quasi-probabilities (probability distributions that admit negative values). This is essentially due to the fact that whenever $\Sigma^< \subsetneq \mathscr{L}_R^<$, there are negative gambles that cannot be proved to be negative in polynomial time. This observation (made mathematically precise in (Benavoli et al., 2019, Theorem 1)) provides for instance an explanation of all paradoxes of quantum mechanics, a special instance of a P-system.

5. Krivine-Vasilescu's nonnegativity

Let $\mathscr{L}_R$ be the space of all polynomials of n variables of degree R in Ω.[2] Let us assume that $\Omega \subset \mathbb{R}^n$ is a compact semi-algebraic set, i.e., a compact set described by polynomial inequalities

$$\Omega = \{x \in \mathbb{R}^n : c_j(x) \geq 0, \; j = 1, \ldots, m\}. \quad (8)$$

Let $\bar{c}_j$ be equal to $\sup_{x \in \Omega} c_j(x)$, define

$$\hat{c}_j(x) = \begin{cases} c_j(x)/\bar{c}_j & \text{if} \quad \bar{c}_j > 0, \\ c_j(x) & \text{if} \quad \bar{c}_j = 0. \end{cases} \quad (9)$$

Therefore, it results that $\hat{c}_j(x) \geq 0$ and $1 - \hat{c}_j(x) \geq 0$ for each $x \in \Omega$. Consider the closed convex cone

$$\Sigma_d^\geq = \left\{ \sum_{(\alpha,\beta) \in \mathbb{N}_d^{2m}} u_{\alpha\beta} \hat{c}_1^{\alpha_1} \cdots \hat{c}_m^{\alpha_m} \right.$$
$$(1-\hat{c}_1)^{\beta_1} \cdots (1-\hat{c}_m)^{\beta_m} : u_{\alpha\beta} \in \mathbb{R}^\geq \Bigg\}, \quad (10)$$

where $\mathbb{N}_d^{2|c|} = \{(\alpha,\beta) \in \mathbb{N}^{2|c|} : |\alpha + \beta| \leq d\}$ and $|\alpha| = \sum_{i=1}^m \alpha_i$. We denote with R the maximum degree of the polynomials in $\Sigma_d^\geq$, so that $\Sigma_d^\geq \subset \mathscr{L}_R$.

---

2. The R in $\mathscr{L}_R$ should stay for "Restricted", here we also use it to denote degree of the polynomial R.

Assumption 1 *We assume that $\Sigma_d^\geq$ satisfies the "pullup" property for every $d \in \mathbb{N}$.*

We can then define the Krivine-Vasilescu nonnegativity certificate (Krivine, 1964; Vasilescu, 2003; Lasserre, 2009).

Definition 6 (Krivine-Vasilescu) *A polynomial of $g \in \mathscr{L}_R$ is "nonnegative" in Ω when it belongs to $\Sigma_d^\geq$ defined in* (10).

That is, the function is "nonnegative" whenever the coefficients $u_{\alpha\beta}$ are nonnegative.

Proposition 7 *Given definition of $\Sigma_d^\geq$ in* (10), *the consistency problem* (5) *can be solved in polynomial time*

The proof is immediate and consists in showing that the membership problem $g \in \Sigma_d^\geq$ can be formulated as a *linear programming* problem. We will give an example of that in the next section. Let $\mathscr{G}$ be a finite set of assessments, and $\mathscr{C}$ its deductive closure $\text{posi}(\mathscr{G} \cup \Sigma_d^\geq)$, with the given definition $\Sigma_d^\geq$ in (10), satisfying $-1 \notin \mathscr{C}$. By Proposition 4 $\mathscr{C}$ is P-coherent, and therefore also logical consistent.

Moreover, it is not difficult to prove that the dual of $\mathscr{C}$ is

$$\mathscr{Q} = \left\{ L \in \mathsf{S} \mid L(g) \geq 0, \; \forall g \in \mathscr{G} \right\}, \quad (11)$$

with the set of (belief) *states* defined as:

$$\mathsf{S} = \Big\{ L \in \mathscr{L}_R^* \mid L(1) = 1,$$
$$L(\hat{c}_1^{\alpha_1} \cdots \hat{c}_m^{\alpha_m} (1-\hat{c}_1)^{\beta_1} \cdots (1-\hat{c}_m)^{\beta_m}) \geq 0, \quad (12)$$
$$\forall \, (\alpha,\beta) \in \mathbb{N}_d^{2|c|} \Big\}.$$

Note that, the linear operator acts on the monomials (this follows by linearity) and, therefore, the dual space is isomorphic to $\mathbb{R}^{s_n(d)}$, with $s_n(d)$ being the number of all monomials for a generic polynomial of n variables and degree d. That means we can define the real numbers

$$z_{\gamma_1 \gamma_2 \ldots \gamma_n} = L(x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}) \in \mathbb{R}, \quad (13)$$

with $\gamma_i \in \mathbb{N}$, and we can rewrite $L(f)$, for any polynomial $f \in \mathscr{L}_R$, as a function of the vector of variables $z \in \mathbb{R}^{s_n(d)}$, whose components are the real variables $z_{\gamma_1 \gamma_2 \ldots \gamma_n}$ defined above.

5.1. Simplex of probability

A particular case of the Krivine-Vasilescu's nonnegativity criterion is obtained when Ω is the probability

simplex:

$$\Omega = \left\{ \theta \in \mathbb{R}^n : \ \theta_j \geq 0, \ \ 1 - \sum_{j=1}^n \theta_j \geq 0 \right\}, \qquad (14)$$

where the changed notation, θ instead of x, reflects the fact that the variables are probabilities.[3] In this case, we can simplify the definition of nonnegativity in (10) as (Lasserre, 2009, Sec.5.4.1):

$$\Sigma_d^{\geq} = \left\{ \sum_{\alpha \in \mathbb{N}_d^{n+1}} u_\alpha \theta_1^{\alpha_1} \cdots \theta_n^{\alpha_n} \right.$$
$$\left. (1 - \theta_1 - \cdots - \theta_n)^{\alpha_{n+1}} : u_\alpha \in \mathbb{R}^{\geq} \right\}. \qquad (15)$$

Note that, the maximum degree of the polynomials in $\Sigma_d^{\geq}$ is d and, therefore, $R = d$. It is easy to prove the following.

Proposition 8 $\Sigma_d^{\geq}$ *satisfies the "pullup" property for every $d \in \mathbb{N}$.*

We have also that (12) in this case becomes:

$$\mathsf{S} = \left\{ L \in \mathscr{L}_R^* \mid L(1) = 1, \right.$$
$$L(\theta_1^{\alpha_1} \cdots \theta_n^{\alpha_n} (1 - \theta_1 - \cdots - \theta_n)^{\alpha_{n+1}}) \geq 0, \qquad (16)$$
$$\left. \forall \ \alpha \in \mathbb{N}_d^{n+1}, |\alpha| \leq d \right\}.$$

We recall that the Bernstein (multivariate) polynomials of degree d on an $n+1$-dimensional simplex are (Prautzsch et al., 2013, Ch.10):

$$B_{\gamma,d}(\theta) = \binom{d}{\gamma} \theta_1^{\gamma_1} \ldots \theta_n^{\gamma_n} (1 - \theta_1 - \cdots - \theta_n)^{d-|\gamma|},$$

where $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma_i \in \mathbb{N}$. Moreover, all Bernstein polynomials of fixed degree d form a basis for the linear space of all polynomials whose degree is at most d, and they form a partition of unity: $\sum_{\gamma : |\gamma| = d} B_{\gamma,d}(\theta) = 1$. By exploiting these properties, we can prove the following.

Proposition 9 *A polynomial $g : \Omega \to \mathbb{R}$ of degree d belongs to the cone in (15) iff it belongs to*

$$\widetilde{\Sigma}_d^{\geq} = \left\{ \sum_{\alpha \in \mathbb{N}_d^{n+1}, |\alpha| = d} u_\alpha \theta_1^{\alpha_1} \cdots \theta_n^{\alpha_n} \right.$$
$$\left. (1 - \theta_1 - \cdots - \theta_n)^{\alpha_{n+1}} : u_\alpha \in \mathbb{R}^{\geq} \right\}. \qquad (17)$$

---

3. Note that, in this case $m = n + 1$ in (8).

For this reason, we call $\widetilde{\Sigma}_d^{\geq}$ (equiv. $\Sigma_d^{\geq}$) the cone of *Bernstein nonnegative polynomials* of degree d.[4]

In general, it holds that

$$\widetilde{\Sigma}_d^{\geq} \subsetneq \mathscr{L}_R^{\geq}$$

that is, there exist nonnegative polynomials that are not included in $\widetilde{\Sigma}_d^{\geq}$.

Example 1 *We use this counter-example from (De Cooman et al., 2015)*

$$q(\theta) = \theta_1^2 - \theta_1 \theta_2 + \theta_2^2,$$

with $n = d = 2$, which is nonnegative in Ω. Now consider the cone

$$\widetilde{\Sigma}_2^{\geq} = \{ u_{002}(1 - \theta_1 - \theta_2)^2 + u_{011}\theta_2(1 - \theta_1 - \theta_2) + u_{020}\theta_2^2 +$$
$$u_{101}\theta_1(1 - \theta_1 - \theta_2) + u_{110}\theta_1\theta_2 + u_{200}\theta_1^2 : u_{ijk} \in \mathbb{R}^{\geq} \} \qquad (18)$$

and an empty $\mathscr{G}$. The lower prevision of q can be computed as follows:

$$\underline{E}_B(q) = \sup_{\lambda_0 \in \mathbb{R}, u_{ijk} \in \mathbb{R}^{\geq}} \lambda_0$$
$$-u_{002} + u_{101} - u_{200} + 1 = 0$$
$$-2u_{002} + u_{011} + u_{101} - u_{110} - 1 = 0$$
$$2u_{002} - u_{101} = 0 \qquad (19)$$
$$-u_{002} + u_{011} - u_{020} + 1 = 0$$
$$2u_{002} - u_{011} = 0$$
$$-\lambda_0 - u_{002} = 0$$

where the equality constraints have been obtained by equating the coefficients of the monomials in

$$q(\theta) - \lambda_0 = u_{002}(1 - \theta_1 - \theta_2)^2 + u_{011}\theta_2(1 - \theta_1 - \theta_2)$$
$$+ u_{020}\theta_2^2 + u_{101}\theta_1(1 - \theta_1 - \theta_2) + u_{110}\theta_1\theta_2 + u_{200}\theta_1^2.$$

The solution of the above LP problem is

$$[\lambda_0, u_{002}, u_{011}, u_{020}, u_{101}, u_{110}, u_{200}]$$
$$= [-0.5, 0.5, 1, 1.5, 1, 0, 1.5],$$

which means that $\underline{E}_B(q) = -0.5 < 0$. It is easy to verify that $\underline{E}_B(q) < 0$ for any $d \geq 2$ and, therefore, q does not belong to $\Sigma_d^{\geq}$ for any d. However, Figure 1 shows that $\underline{E}_B(q)$ quickly tends to zero at the increase of the degree d of $\Sigma_d^{\geq}$. Therefore, we can build a hierarchy of LPs (Lasserre, 2009, Sec.5.4) of increasing size such that

$$\underline{E}_B(q) \xrightarrow{d \to \infty} \underline{E}(q).$$

This is true in general provided that $\mathscr{G} \cap \mathscr{L}_R^{<} = \emptyset$.

---

4. We could again simplify (16), but we leave the redundant formulation (16) because it requires to specify L on the monomials too.
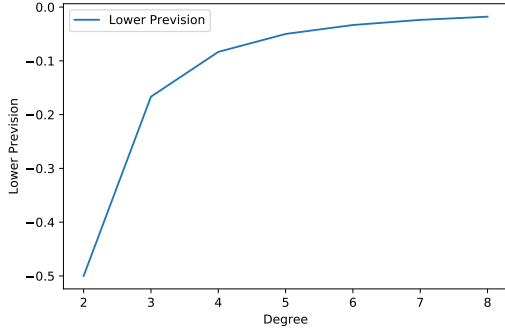
Figure 1: Convergence of $\underline{E}_B(q)$ to $\underline{E}(q)$ at the increase of the degree d.

5.2. Updating via partition of unity

$\mathscr{L}_R$ is a space of polynomials and, therefore, it does not include indicator functions. That means we cannot define conditioning. However, we can still update our beliefs in a weaker way using a partition of unity. We have seen that the Bernstein polynomials form a partition of unity, e.g.,

$$n = 3, d = 1: \quad \{\theta_1, \theta_2, \theta_3, 1 - \theta_1 - \theta_2 - \theta_3\}, \quad (20)$$

they are nonnegative functions and sum up to one. We can then compute an updated lower prevision for a gamble q as:

$$\underline{E}_B(q|\pi) = \sup_{\lambda_j \geq 0, \lambda_0} \lambda_0$$
$$s.t. \quad (21)$$
$$(q - \lambda_0)\pi - \sum_{j=1}^{|G|} \lambda_j g_j \in \Sigma_d^{\geq},$$

where d must be large enough to guarantee that the membership problem is well-posed and π denotes any *subset sum* of the partition of unity (in the example in (20) $\pi \in \{\theta_1, \theta_2, \theta_3, 1 - \theta_1 - \theta_2 - \theta_3, \theta_1 + \theta_2, \theta_1 + \theta_3, \theta_2 + \theta_3, \dots\}$). To alternatively justify this rule, we point out that π can be interpreted as a multinomial likelihood[5] and the result of (21) as a bounded rationality version of a *regular* posterior (Walley, 1991, Appendix J5).

5.3. A Bell inequality in the Bernstein world

In this section, we derive a Bell's type inequality in the Bernstein world: a probabilistic inequality that holds in $\mathfrak{T}$ but that is violated in $\mathfrak{T}^*$ (Bernstein world). We will derive it by building a negative polynomial that has positive prevision in $\mathfrak{T}^*$. In the next section, we

will show that the *state* assigning a positive prevision to such polynomial is *entangled*! For this purpose, we consider two coins, that we denote as l (left) and, respectively, r (right), and define

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ 1 - \theta_1 - \theta_2 - \theta_3 \end{bmatrix} = \text{Prob} \begin{bmatrix} H_l H_r \\ T_l H_r \\ H_l T_r \\ T_l, T_r \end{bmatrix},$$

where H_i, T_i denote the outcome Heads and, respectively, Tails for the left or right coin. In this case, the possibility space is

$$\Omega = \left\{ \theta \in \mathbb{R}^3 : \ \theta_1, \theta_2, \theta_3 \geq 0, \ 1 - \theta_1 - \theta_2 - \theta_3 \geq 0 \right\}. \quad (22)$$

Note that, the following *marginal* relationships hold:

$$\theta_{H_l} = \text{Prob}(H_l) = \theta_1 + \theta_3, \quad \theta_{H_r} = \text{Prob}(H_r) = \theta_1 + \theta_2.$$

Let d be equal to 2 and consider the *state*:

$$\begin{aligned} L(\theta_1) &= z_{100} = 1/3 & L(\theta_1^2) &= z_{200} = 1/3 \\ L(\theta_2) &= z_{010} = 1/6 & L(\theta_2^2) &= z_{020} = 0 \\ L(\theta_3) &= z_{001} = 1/6 & L(\theta_3^2) &= z_{002} = 0 \\ L(\theta_1\theta_2) &= z_{110} = 0 & L(\theta_1\theta_3) &= z_{101} = 0 \\ L(\theta_2\theta_3) &= z_{011} = 1/6 & L(1) &= z_{000} = 1, \end{aligned} \quad (23)$$

which belongs to (16).

Now consider the polynomial:

$$q(\theta) = -(\theta_1 + \theta_2)^2 - (\theta_1 + \theta_3)(-2\theta_1 - 2\theta_2 + 1) - \varepsilon,$$

with $\varepsilon > 0$ and observe that $q(\theta) \leq -\varepsilon$. The polynomial is negative! However, its prevision[6] $E_B(q)$ w.r.t. the state (23) is equal to

$$L(q) = -z_{001} + 2z_{011} - z_{020} - z_{100} + 2z_{101} + z_{200} - \varepsilon$$
$$= \frac{1}{6} - \varepsilon \geq 0. \quad (24)$$

Therefore, we have violated an inequality that holds in classical probability ($E(q) \leq -\varepsilon$ in $\mathfrak{T}$), although the set of desirable gambles

$$\mathscr{C} = \{g \in \mathscr{L}_R \mid L(g) \geq 0\},$$

with L defined in (23), is logical consistent in $\mathfrak{T}^*$ (P-coherent). This is the essence of Bell's type inequalities: the quantum weirdness which is also present in Bernstein's world.
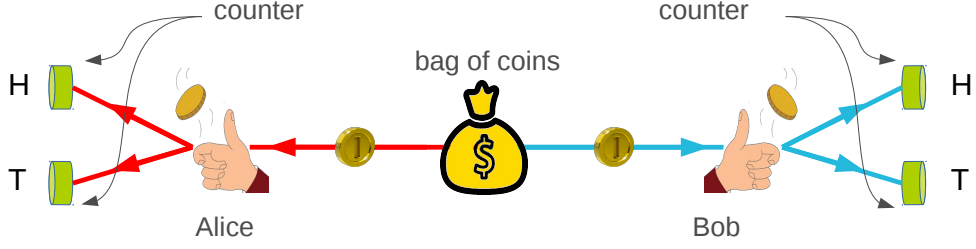

Figure 2: Coin toss experiment in Bernstein's world.

5.4. Entanglement

We continue the previous example and we set up a thought experiment that uncovers the entanglement of the two coins.

Assume two coins are drawn from a bag in the state (23). We give the left coin to Alice and the right coin to Bob as depicted in Figure 2.

We will now show that after the coins move apart, there are "matching" correlations between the output of their toss. That is a measurement (though a toss) of the bias of one coin will allow the prediction, with certainty, of the outcome of the measurement (toss) on the other coin.

Assume that Alice tosses her coin first and that it lands Heads, then she can compute her updated prevision (through (21) with $\pi = \theta_1 + \theta_3$) for the gamble $q(\theta) = \theta_1 + \theta_2$ (Heads on Bob's coin). We can easily do this computation in the dual space.[7] Note that

$$0 = L((q-\lambda_0)\pi) = -\lambda_0 z_{001} - \lambda_0 z_{100} + z_{011} \\ + z_{101} + z_{110} + z_{200}$$

has solution $\lambda_0 = 1$. Alice instantaneously knows that the result of the toss of Bob's coin will be Heads. Similarly, we can consider all the other cases

$$\begin{array}{lll} q = 1 - \theta_1 - \theta_2, & \pi = \theta_1 + \theta_3, & \lambda_0 = 0 \\ q = \theta_1 + \theta_2, & \pi = 1 - \theta_1 - \theta_3, & \lambda_0 = 0 \\ q = 1 - \theta_1 - \theta_2, & \pi = 1 - \theta_1 - \theta_3, & \lambda_0 = 1. \end{array}$$

This means that as soon as Alice sees the result of the toss of the left coin, she immediately knows that the result of the toss of Bob's coin will be the same. The two coins are totally "correlated". Classical correlations can be explained by a common cause, or correlated "elements of reality" (Einstein et al., 1935). This is not the case in Bernstein's world.

---

7. Since the state (23) is precisely specified, we do not need to solve any optimisation to compute the posterior prevision.

In fact, although the marginal operators satisfy

$$\begin{aligned} L(\theta_{H_r}) = L(\theta_1 + \theta_2) & = z_{100} + z_{010} = \frac{1}{2}, \\ L(\theta_{H_l}) = L(\theta_1 + \theta_3) & = z_{100} + z_{001} = \frac{1}{2}, \\ L(\theta_{H_r}^2) = L((\theta_1 + \theta_2)^2) & = z_{200} + 2z_{110} + z_{020} = \frac{1}{3}, \\ L(\theta_{H_l}^2) := L((\theta_1 + \theta_3)^2) & = z_{200} + 2z_{101} + z_{002} = \frac{1}{3}, \end{aligned} \tag{25}$$

these are the same moments we would get if the marginal distribution of the two coins is uniform. That means that if we send an ensemble of coins to Alice (Bob), and she (he) tosses them, she (he) will experience Heads half of the times. A classical correlation model that is compatible with these marginal moments is given by this probabilistic mixture of atomic charges (Dirac's deltas):

$$p(\theta) = \frac{1}{2}\delta_{\left\{\begin{bmatrix} \frac{1}{6}(3-\sqrt{3}) \\ 0 \\ 0 \\ \frac{1}{6}(3+\sqrt{3}) \end{bmatrix}\right\}}(\theta) + \frac{1}{2}\delta_{\left\{\begin{bmatrix} \frac{1}{6}(3+\sqrt{3}) \\ 0 \\ 0 \\ \frac{1}{6}(3-\sqrt{3}) \end{bmatrix}\right\}}(\theta).$$

However, this probabilistic model (or any other) can never satisfy the moment constraints (23) or, equivalently, can never violate the Bell's type inequality presented in the previous section. We have entanglement!

6. Conclusions

In this work, after a brief description and analysis of the structural properties of P-coherent models, we have shown that the space of Bernstein polynomials in which nonnegativity is specified by the Krivine-Vasilescu certificate is yet another instance of this theory and that, therefore, it is possible to construct

in it a thought experiment uncovering entanglement with classical (hence non quantum) coins.

As a final side remark, we believe that formulating the theory of desirable gambles directly as a logic system provides an elegant way for extending the framework to the accept-reject one but also for merging the latter with the theory of choice functions.

References

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer, 2007.

John S Bell. Bertlmann's socks and the nature of reality. *Le Journal de Physique Colloques*, 42(C2): C2–41, 1981.

Alessio Benavoli, Alessandro Facchini, Dario Piga, and Marco Zaffalon. SOS for bounded rationality. In *Proc. ISIPTA'17 Int. Symposium on Imprecise Probability: Theories and Applications,*, pages 1–12. PJMLR, 2017a.

Alessio Benavoli, Alessandro Facchini, Marco Zaffalon, and José Vicente-Pérez. A polarity theory for sets of desirable gambles. In Alessandro Antonucci, Giorgio Corani, Inés Couso, and Sébastien Destercke, editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 37–48. PMLR, 10–14 Jul 2017b.

Alessio Benavoli, Alessandro Facchini, and Marco Zaffalon. Computational Complexity and the Nature of Quantum Mechanics. *ArXiv e-prints 1606.03615*, February 2019.

Jasper De Bock and Gert De Cooman. Imprecise bernoulli processes. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 400–409. Springer, 2012.

Gert de Cooman and Erik Quaeghebeur. Exchangeability and sets of desirable gambles. *International Journal of Approximate Reasoning*, 53:363–395, 2012.

Gert De Cooman, Erik Quaeghebeur, Enrique Miranda, et al. Exchangeable lower previsions. *Bernoulli*, 15(3):721–735, 2009.

Gert De Cooman, Jasper De Bock, and Marcio Alves Diniz. Coherent predictive inference under exchangeability with imprecise probabilities. *Journal of Artificial Intelligence Research*, 52:1–95, January 2015. doi: http://dx.doi.org/10.1613/jair.4490.

Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*, volume 1. John Wiley & Sons, Chichester, 1974.

Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47 (10):777, 1935.

Jean-Louis Krivine. Anneaux préordonnés. *Journal d'analyse mathématique*, 12(1):307–326, 1964.

Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2009.

Hartmut Prautzsch, Wolfgang Boehm, and Marco Paluszny. *Bézier and B-spline techniques.* Springer Science & Business Media, 2013.

Abraham Seidenberg. A new decision method for elementary algebra. *Annals of Mathematics*, pages 365–374, 1954.

Herbert Alexander Simon. *Models of man: social and rational; mathematical essays on rational human behavior in society setting.* Wiley, 1957.

Alfred Tarski. A decision method for elementary algebra and geometry. 1951.

Florian-Horia Vasilescu. Spectral measures and moment problems. *Spectral Theory and Its Applications, Theta*, pages 173–215, 2003.

Peter Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1991. ISBN 9780412286605.