Bell inequalities from variable-elimination methods

Costantino Budroni$^1$ and Adán Cabello$^{1,2}$

$^1$ Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain
$^2$ Department of Physics, Stockholm University, S-10691 Stockholm, Sweden

E-mail: cbudroni@us.es and adan@us.es

Received 28 December 2011, in final form 17 July 2012
Published 5 September 2012
Online at stacks.iop.org/JPhysA/45/385304

Abstract

Tight Bell inequalities are facets of Pitowsky’s correlation polytope and are usually obtained from its extreme points by solving the hull problem. Here, we present an alternative method based on a combination of algebraic results on extensions of measures and variable-elimination methods, e.g., the Fourier–Motzkin method. Our method is shown to overcome some of the computational difficulties associated with the hull problem in some non-trivial cases. Moreover, it provides an explanation for the arising of only a finite number of families of Bell inequalities in measurement scenarios where one experimenter can choose between an arbitrary number of different measurements.

PACS number: 03.65.Ud

1. Introduction

Bell inequalities [1] play a fundamental role in the investigation of non-classical features of quantum mechanics (QM), from the foundational problems raised by the impossibility of a local hidden variable description of quantum systems, to applications for quantum information processing like entanglement-based reduction of classical communication complexity [2, 3], device-independent secure communication [4, 5] and randomness expansion [6]. Similar inequalities, called non-contextuality inequalities, arise in the discussion of quantum contextuality [7–9].

In fact, the origin of all these inequalities is rooted in logic and probability [10–12] and can be traced back [13] to Boole’s notion of conditions of possible experience [14]; they are necessary conditions for the interpretation of a set of experimentally observed relative frequencies as probabilities in a single probability space.

From an algebraic point of view [15], Bell and non-contextuality inequalities can be seen as conditions for the extension of a function, defined on a subset of a Boolean algebra, to a normalized measure on the entire Boolean algebra, a problem already investigated by Horn and Tarski [16].
A general method for the derivation of Bell and non-contextuality inequalities is based on Pitowsky’s notion of correlation polytope [11]. Such a polytope appears in many fields and with different names (see [17] for an extensive survey of the topic). Its geometrical structure is completely defined by a set together with a family of its subsets.

In the framework of QM, the above set is a set of observables and the subsets correspond to subsets of compatible or jointly measurable observables. The operational definition of such notions is still under discussion [18]. They have been defined in different ways according to the experimental scenario and the physical assumptions involved (e.g., space-like separated measurements, sequential measurements). Such distinctions are irrelevant for our discussion. In contrast, the use of an abstract notion of compatibility allows us for a unified presentation of our mathematical results which are valid in all the above-mentioned scenarios.

Given a set of observables together with the subsets of compatible observables, the corresponding correlation polytope is defined as the convex hull of a set of vectors, the vertices or extreme points of the polytope, representing possible truth assignments for observables and logical conjunctions between compatible pairs, triples, etc.

Every convex polytope has two representations: one as the convex hull of its vertices (V-representation) and the other as the intersection of a finite number of half-spaces (H-representation), each given by a linear inequality. Complete sets of tight (i.e. facet supporting) Bell inequalities associated with a given set of observables precisely amount to the H-representations of the corresponding correlation polytopes.

The H-representation of a convex polytope can be computed starting from its V-representation by solving the hull problem. For high-dimensional polytopes, this is a very difficult problem. In fact, this method has been used only to compute simple cases [19, 20], by means of programs, such as cdd [21], lrs [22] and porta [23].

More recently, an alternative approach has been proposed by Avis et al [24, 25], which is based on the relation between correlation polytopes and a family of convex polytopes called cut polytopes, well-studied objects in polyhedral combinatorics, and a variable-elimination method derived from the Fourier–Motzkin method of elimination of variables (see, e.g., [26]) and called triangular elimination. Their method provided more than 200 million new tight Bell inequalities for cases in which the corresponding hull problem is computationally intractable. However, since such a method is based on cut polytopes, it can be applied only to the bipartite case.

Here, we present an alternative general method based on the application of Fourier–Motzkin variable-elimination method to conditions derived in [15] as consistency conditions for putting together partial extensions of quantum probabilities in order to obtain a classical probability description. More precisely, such conditions are expressed in terms of a system of linear inequalities where also correlations between incompatible observables appear as variables: a classical probability space representation exists for a given set of QM predictions if and only if the corresponding system of linear inequalities admits a solution; Bell, or non-contextuality, inequalities are obtained by eliminating the variables associated with correlations between incompatible observables. Our approach can be seen as a generalization of Fine’s derivation of the Clauser–Horne–Shimony–Holt (CHSH) polytope [10].

The above method also provides a generalization of the result obtained by Sliwa [27] and Collins and Gisin [20], namely, the appearance of only a finite number of families of Bell inequalities in measurement scenarios where one experimenter is allowed to choose between an arbitrary number of different measurements.

The variable-elimination method can also be applied to a family of correlation polytopes called complete probability polytopes. In this case, the method may not provide a significant
computational advantage with respect to other known algorithms. Nevertheless, we believe that such results are interesting and we collected them in the appendix.

The rest of the paper is organized as follows. In section 2, the notion of correlation polytope is reviewed, together with the result on extensions of probabilities associated with tree graphs. In section 3, we discuss their application to the problem of the computation of $\mathcal{H}$-representation for general correlation polytopes and provide some examples. Finally, in section 4, we discuss some possible further developments and computational results.

2. Correlation polytopes and extensions of measures

In this section, we recall the definition of correlation polytope and its probabilistic interpretation; then, we introduce the notion of a compatibility graph for a set of observables and present the main result about the classical probability space representation for the corresponding QM predictions.

We use the definition of a correlation polytope given in [17], which is the natural generalization of Pitowsky’s notion [12] to higher order correlations, but with a slightly different notation.

Given a set of propositions $\mathcal{G} = (A_1, \ldots, A_n)$ and a family $\mathcal{I}$ of subsets of $\mathcal{G}$, i.e. $\mathcal{I} \subset 2^{\mathcal{G}}$, we define the sets $\mathcal{S}_k$, $k = 2, \ldots, m$ with $m \leq n$, as the sets of logical conjunctions $\mathcal{S}_k = \{ A_{i_1} \wedge \cdots \wedge A_{i_k} | i_j \neq i_j, \{ A_{i_1}, \ldots, A_{i_k} \} \in \mathcal{I} \}$. The lines of the truth table associated with the above set of propositions and logical conjunctions between them, namely the $2^n$ vectors of $\mathbb{R}^{[|\mathcal{G}|+|\mathcal{S}_2|+\cdots+|\mathcal{S}_m|}$,

$$u_v = (\epsilon_1, \ldots, \epsilon_n, \epsilon_1\epsilon_j, \ldots, \epsilon_i \epsilon_1 \cdots \epsilon_i u_n, \ldots),$$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{0, 1\}^n$ are called the vertices of the correlation polytope. Their convex hull, i.e. the set of points generated by their convex combinations, is called the correlation polytope associated with $\mathcal{I}$ and denoted as COR$^\square(\mathcal{I})$.

It is convenient to introduce the following notation which makes apparent the correspondence between coordinates and joint probabilities. The coordinates of a point $p \in \mathbb{R}^{[|\mathcal{G}|+|\mathcal{S}_2|+\cdots+|\mathcal{S}_m|}$ will be denoted as

$$p = (p_1, \ldots, p_n, \ldots p_{ij}, \ldots, p_{i_1 \cdots i_n}, \ldots).$$

In order to clarify our notation, it is convenient to present a simple example, the CHSH scenario. Consider the bipartite scenario in which one experimenter, say Alice, can choose between two dichotomic measurements, associated with propositions $A_1$ and $A_2$, and another experimenter, say Bob, can choose between two dichotomic measurements, associated with propositions $A_3$ and $A_4$. The set of propositions is, therefore, $\mathcal{G} = (A_1, A_2, A_3, A_4)$. Moreover, the measurements associated with $A_i$ and $A_j$, for $i = 1, 2$ and $j = 3, 4$, can be performed jointly and, consequently, it makes sense to consider the following set of logical conjunctions: $\mathcal{S}_2 = \{ A_1 \wedge A_3, A_1 \wedge A_4, A_2 \wedge A_3, A_2 \wedge A_4 \}$. The associated polytope is described by $2^4 = 16$ vertices in $\mathbb{R}^3$, namely

$$u_v = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_1 \epsilon_3, \epsilon_1 \epsilon_4, \epsilon_2 \epsilon_3, \epsilon_2 \epsilon_4), \quad \epsilon_i \in \{0, 1\},$$

where $\epsilon_i$ represents a classical $[0, 1]$-valued assignment to proposition $A_i$ and $\epsilon_i \epsilon_j$ is the classical assignment for the logical conjunction $A_i \wedge A_j$.

As a consequence of the Weyl–Minkowski theorem (see, e.g., [12]), each convex polytope has a double description: one as the convex hull of its vertices $u_v$, i.e. the $\mathcal{Y}$-representation, and one as a (finite) intersection of half-spaces which generates it, each one given by a linear inequality, i.e. the $\mathcal{H}$-representation.
The convex hull of vertices (3) gives a set of linear inequalities constraining the coordinates of a generic point in $\mathbb{R}^8$, which we denote by

$$p = (p_1, p_2, p_3, p_4, p_{13}, p_{14}, p_{23}, p_{24}).$$

Such inequalities, known as CHSH inequalities, are presented in equations (7a)–(7d) and will be discussed in more detail. The interpretation of such a geometrical object is the following: given a vector belonging to the polytope, each component represents the joint probability for the corresponding subset of propositions, e.g., $p_{13}$ represents the joint probability $\text{Prob}(A_1 \wedge A_3)$; the whole polytope gives possible ranges for such joint probabilities if the underlying probabilistic structure is assumed to be classical, i.e. given by a probability space or, equivalently (since the number of propositions is finite) by a normalized measure on a Boolean algebra.

It is interesting to define the complete probability polytope associated with $n$ propositions $G = \{A_1, \ldots, A_n\}$, it is the correlation polytope COR$_{(2^G)}$, i.e. the correlation polytope for a set of compatible proposition, in particular, $|G| + \sum_k |S_k| = \sum_{k=1}^n \binom{n}{k} = 2^n - 1$.

The importance of such polytopes has already been recognized in [17], namely, the fact that each COR$_{(I)}$ can be obtained as a projection of COR$_{(2^G)}$ (on the subspace $\mathbb{R}^{|I|}$). A more detailed discussion, containing also an algebraic derivation of the $\mathcal{H}$-representation for complete probability polytopes, can be found in the appendix.

We now recall a result on the extension of probability measures which allows for another projection-based approach to the computation of $\mathcal{H}$-representation of a correlation polytope.

More details can be found in [15].

Proposition 1 (Representation of tree graphs). Consider a set of probabilities $p_i$ on a set of $\{0, 1\}$-valued observables $A_i$ and correlations $p_{ij}$ on a subset of pairs $A_i, A_j$, defining a probability on each pair $A_i, A_j$, with $p_i = \langle A_i \rangle$, $p_j = \langle A_j \rangle$, $p_{ij} = \langle A_i A_j \rangle$; now depict observables $A_i$ as vertices and the above pairs as edges in a graph. Then, any set of predictions associated with a tree graph, i.e. a graph without closed loops, admits a classical representation.

Proposition 2. The same holds with yes/no observables $A_i$ substituted by free Boolean algebras $\mathcal{A}_i$, $p_i$ by probabilities on $\mathcal{A}_i$, $p_{ij}$ by probabilities on the Boolean algebra freely generated by the union of the sets of generators of $\mathcal{A}_i$ and $\mathcal{A}_j$.

We recall that a Boolean algebra is freely generated by $n$ generators $B_1, \ldots, B_n$ if such generators are as unconstrained as possible, i.e. they satisfy no conditions except those necessary conditions defining a Boolean algebra (e.g., distributive law). Since all Boolean algebras freely generated by $n < \infty$ generators are isomorphic, for the sake of simplicity, we can think of the algebra of subsets $2^X$ of the set $X = \{0, 1\}^n$, with set theoretic operations ($\cap, \cup, ^c$); then, the subsets $B_i = \{(x_1, \ldots, x_n) \in X \mid x_i = 1\}$, for $i = 1, \ldots, n$, can be taken as free generators. In terms of propositions and truth assignments, the free Boolean algebra assumption amounts to the assumption that each possible $\{0, 1\}$-valued assignment to propositions is admissible. For more details, see [28]. Note that this is precisely how we define the $2^n$ vertices of the correlation polytope.

In the next section, we shall show how to exploit the above results to obtain $\mathcal{H}$-representation for general correlation polytopes starting from lower dimensional polytopes and using the Fourier–Motzkin method.
3. Derivation of Bell inequalities

The above results provide a general method for the computation of $H$-representation for correlation polytopes which consists in solving the hull problem for a smaller polytope, then constructing a higher dimensional polytope from that solution and applying variable-elimination methods, such as, e.g., the Fourier–Motzkin method.

For the convenience of the reader, we briefly recall the Fourier–Motzkin method; for more details, see [26]. Consider a system of inequalities of the form $Ax \leq b$, where $A$ is a $m \times d$ real matrix, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $b \in \mathbb{R}^m$ and suppose we want to eliminate the variable $x_d$. After a proper normalization, each inequality will be of the form $a_i x_1 + \cdots + a_{id} x_d \leq b_i$, with $a_{id}$ being $+1$, $-1$ or $0$. By summing an inequality where $a_{id} = +1$ with an inequality where $a_{jd} = -1$, we obtain a new inequality not containing $x_d$. By repeating the above operation for every possible pair $i, j$ such that $a_{id} = +1$ and $a_{jd} = -1$, and considering also the inequalities where $a_{id} = 0$, we obtain a new system of inequalities not containing $x_d$. From a geometric point of view, since the system of linear inequalities can represent a polytope (more generally, a cone), the above operation amounts to a projection on the coordinates associated with the variables $x_1, \ldots, x_{d-1}$.

Our method consists in exploiting the automatically existing classical representations for subsets of observables with compatibility relations described by tree graphs, see propositions 1 and 2. Conditions for classical representability arise as consistency (i.e. coincidence on intersections) conditions for putting together partial extensions associated with subgraphs, giving rise to a description of the initial compatibility graph as a tree graph on such extended nodes. Such consistency conditions are expressed in terms of the existence of a solution for a set of linear inequalities. One of the main applications of the Fourier–Motzkin algorithm is precisely deciding whether a system of inequalities has a solution.

Moreover, we shall discuss how this result gives an account for the appearance of a finite number of families of Bell inequalities in cases in which one experimenter can choose between an arbitrary number of different measurements while such a number is fixed for the others, as recognized by Sliwa [27] and Collins and Gisin [20] for the case in which Alice performs two measurements and Bob an arbitrary number.

We recall that, although our method is general, different strategies are possible corresponding to different partitions of the initial graph into subgraphs. We shall discuss our method by means of some simple examples. The first one is the derivation of the CHSH polytope. It is interesting to note that it is analogous to that presented by Fine [10]; our method can be seen as a generalization of his idea to an arbitrary number of observables.

3.1. CHSH polytope from the Bell–Wigner polytope

The CHSH polytope is generated by the following set of vertices:

$$u_v = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_1 \varepsilon_3, \varepsilon_1 \varepsilon_4, \varepsilon_2 \varepsilon_3, \varepsilon_2 \varepsilon_4), \quad \varepsilon_i \in \{0, 1\}.$$  (5)

It is associated with a bipartite measurement scenario in which Alice can choose between two measurements, associated with propositions $A_1$ and $A_2$, and Bob can choose between two measurements, associated with propositions $A_3$ and $A_4$.

As already recognized by Fine [10] and also discussed in [15, 29], the existence of a classical description for the four observables is equivalent to the existence of classical descriptions for the two subsystems $\{A_1, A_2, A_3\}$ and $\{A_1, A_2, A_4\}$, coinciding on $\{A_1, A_2\}$. In fact, the two classical descriptions would give rise to an extension of the probability assignment to the four observables satisfying the hypothesis of proposition 2 (see figures (a)–(c)).
The constraints on the subsystem \{A_1, A_2\} imposed by the third observable, \(A_3\) or \(A_4\), are described by the Bell–Wigner polytope, i.e. the correlation polytope associated with three propositions and their pairwise logical conjunctions, which is given by the following inequalities, obtained in \([12]\):

\[
0 \leq p_{ij} \leq p_i, \quad 0 \leq p_{ij} \leq p_j, \quad ij = 12, 1s, 2s,
\]

\[
p_i + p_j - p_{ij} \leq 1, \quad ij = 12, 1s, 2s,
\]

\[
1 - p_1 - p_2 - p_s + p_{12} + p_{1s} + p_{2s} \geq 0,
\]

\[
p_1 - p_{12} - p_{1s} + p_{2s} \geq 0,
\]

\[
p_2 - p_{12} - p_{2s} + p_{1s} \geq 0,
\]

\[
p_s - p_{1s} - p_{2s} + p_{12} \geq 0,
\]

for \(s = 3\) or 4. For fixed \(s\) (3 or 4), the above inequalities give ranges for the classical probabilities for a system of three proposition and their pairwise joint probabilities. It is interesting to note that, by proposition 1, each subset of three observables always admits a classical representation (see also figure 1(b)); therefore, for fixed \(s\), the above system of inequalities admits a solution, i.e. a value can be assigned to the joint probability \(p_{12}\) (which is assumed to be experimentally unmeasurable, since it is associated with two noncommuting observables) consistently with the constraints imposed by the classical model and the measurable joint probabilities \(p_{1s}\) and \(p_{2s}\). This implies that, for fixed \(s\), the elimination of variable \(p_{12}\) gives rise only to trivial inequalities, i.e. inequalities always satisfied by quantum predictions. An analysis of the constraints imposed on such ‘experimentally unmeasurable’ joint probabilities has been presented in \([29]\). Analogous situations always arise in the bipartite scenario and this additional property just gives more information about the triviality of certain inequalities. However, our method can also be applied to the general multipartite case.

Classical representability for QM prediction in the CHSH scenario amounts, therefore, to the existence of a common solution for the two systems of linear inequalities, namely, it amounts to the existence of a value for \(p_{12}\) consistent with the constraints imposed by classical descriptions for the two subsystems of three observables.
As a consequence of the general properties of the Fourtié–Mottkin method (see [26]), a
system of inequalities admits a solution if and only if the projected system, i.e. the system
obtained by eliminating one or more variables, admits a solution. It follows that the above
set of measurements admits a classical description if and only if measured correlations, i.e.
correlations between compatible observables, satisfy the system of inequalities obtained by
eliminating \( p_{12} \).

In order to eliminate \( p_{12} \), we just combine inequalities where \( p_{12} \) appears with opposite
sign and keep inequalities where it does not appear. As discussed before, by proposition 1,
combining inequalities with the same index \( s \) gives rise only to trivial inequalities, in particular,
they are also redundant. A posteriori, we know that the same happens also for inequalities (6a)
and (6b).

The interesting inequalities are those obtained from the combination of (6c)–(6f) for
different \( s \), namely
\[
-1 \leq p_{13} + p_{14} + p_{24} - p_{23} - p_1 - p_4 \leq 0, \tag{7a}
\]
\[
-1 \leq p_{23} + p_{24} + p_{14} - p_{13} - p_2 - p_4 \leq 0, \tag{7b}
\]
\[
-1 \leq p_{14} + p_{13} + p_{23} - p_{24} - p_1 - p_3 \leq 0, \tag{7c}
\]
\[
-1 \leq p_{24} + p_{23} - p_{14} - p_2 - p_3 \leq 0. \tag{7d}
\]

For instance, \(-1 \leq p_{13} + p_{14} + p_{24} - p_{23} - p_1 - p_4 \) is obtained as a sum of inequality (6c)
for \( s = 4 \), which contains the term \(+p_{12}\), and inequality (6e) for \( s = 3 \), which contains the
term \(-p_{12}\).

Inequalities (7a)–(7d), together with the inequalities (6a), (6b) in which \( p_{12} \) does not
appear, give the \( \mathcal{H} \)-representation of the CHSH polytope (compare with, e.g., [12]).

3.2. Bipartite (2, \( n \)) scenario

An analogous argument applies to the scenario in which Alice can choose between two
measurements and Bob can choose among \( n \geq 2 \) measurements, associated with propositions
\( A_1, \ldots, A_{n+2} \): the initial system of inequalities is still given by equations (6a)–(6f), but with \( s \)
taking values in \([3, 4, \ldots, n+2]\) (see figure 1(d)).

Since only one variable (i.e. \( p_{12} \)) has to be eliminated, at most two inequalities with
different index \( s \) can be combined to give a valid inequality. Therefore, the final set of
inequalities is given by (7a)–(7d) with the pair 3,4 substituted by any pair \( i, j \) with
\( i, j \in [3, \ldots, n+2] \) and \( i < j \). This is precisely the result obtained in [20, 27].

3.3. Two parties, three settings

Now consider the bipartite scenario in which Alice can choose among three measurements,
associated with propositions \( A_1, A_2 \) and \( A_3 \), and Bob can choose among three measurements,
associated with propositions \( A_4, A_5 \) and \( A_6 \).

Analogously to the previous case, see figure 2, the existence of a classical description
for the six observables is equivalent to the existence of classical descriptions for the three
subsystems \( \{A_1, A_2, A_3, A_4\}, \{A_1, A_2, A_3, A_5\} \) and \( \{A_1, A_2, A_3, A_6\} \), coinciding on \( \{A_1, A_2, A_3\} \).
A probability on \( \{A_1, A_2, A_3\} \) is completely defined, see lemma 5, once the probabilities
\( p_1, p_2, p_3, p_{12}, p_{13}, p_{23}, p_{123} \) are given.

It is therefore sufficient to calculate the correlation polytope associated with probabilities
\( p_1, p_2, p_3, p_{12}, p_{13}, p_{23}, p_{123}, p_{123} \); then, consider the system given by all the above
inequalities for \( s = 4, 5, 6 \) and eliminate the variables \( p_{12}, p_{13}, p_{23}, p_{123} \).
3.4. Bipartite \((3, n)\) scenario

Again, by adding observables only on Bob’s side, one just obtains more copies of the initial system of inequalities, but with different indices \(s\). The situation is analogous to that depicted in figure 1(d), but with \(A_{12}\) substituted by \(A_{123}\).

In particular, since four variables have to be eliminated, at most \(2^4 = 16\) inequalities with different indices \(s\) can be combined. As a result, all families of valid inequalities for the general case \((3, n)\) already arise in the case in which Alice and Bob perform 3 and 16 measurements, respectively.

3.5. Multipartite \((m, \ldots, m, n)\) scenario

The above argument can be extended to the case of \(p\) parties in which the first \(p - 1\) can choose among \(m\), with \(m\) fixed, measurements, while the last one can choose among \(n > m\), with \(n\) arbitrary, measurements: all families of inequalities can be obtained by studying the case in which the last experimenter performs \(2^k\) measurements, where \(k\) is the number of variables to be eliminated.

4. Conclusions and computational results

We have presented an alternative method for the computation of half-space representation for correlation polytopes based on algebraic conditions and variable elimination. A reasonable question is: does it provide any advantage with respect to existing methods? In order to show the advantages of the tree graph method, we have computed the \(\mathcal{H}\)-representation for some simple polytopes, with (i) our tree graph method using existing software implementing the Fourier–Motzkin algorithm, specifically, we used \texttt{porta} [23] and \texttt{FM} library [30], and (ii) using standard software for solving the hull problem, specifically, we used \texttt{cdd}.

For simple cases like the \((2, 2)\) (i.e. the CHSH) and \((3, 3)\) scenarios, the computation is equally fast with both methods. However, remarkably, our tree graph method is noticeably faster to compute asymmetric scenarios.

For the \((3, 4)\) scenario, the tree graph method implemented with \texttt{porta} completed the calculation in \(\approx 11\) min, while \texttt{cdd} needed \(\approx 20\) min. The 11 min include the time (s) required to calculate the initial polytope (see section 3.3).

For the \((3, 5)\) scenario, the tree graph method implemented with \texttt{porta} completed the calculation in \(\approx 72\) min, while \texttt{cdd} was still running after a week and we had to stop it. All computations were performed on the same machine with an Intel Xeon CPU running at 3.20 GHz.
However, beyond the practical advantage that the tree graph method has to compute some scenarios, the main purpose of this paper is to point out that, for some polytopes, the \( \mathcal{H} \)-representation can be easily expressed by means of simple algebraic arguments, and then these polytopes can be related to those associated with polytopes arising in physically interesting scenarios by means of variable-elimination methods. For example, in 1948 the Boolean algebraic approach in [16] provided complete sets of Bell inequalities. To our knowledge, this approach has never been investigated in relation to Bell inequalities, and we hope that our work will stimulate further developments in this direction.

Acknowledgments

The authors thank Louis-N\OE nel Pouchet for helpful discussions about his software and the Fourier–Motzkin algorithm in general, Emilio Pinna for his help in writing some of the programs used for the calculations and an anonymous referee for useful comments and suggestions. This work was supported by the Spanish projects FIS2008-05596 and FIS2011-29400, and the Wenner-Gren Foundation.

Appendix

In this appendix, we recall some results, proposition 4 and lemmas 3, 5, arising from the translation of Pitowsky’s notion of correlation polytope into the Boolean framework and use them to derive the \( \mathcal{H} \)-representation for the complete correlation polytope. For more details on such results and the basic notions of Boolean algebra involved, see [15].

After completing our work, we realized that (i) obtaining a correlation polytope as a projection of the corresponding complete probability polytope need approximatively the same amount of time and resources as the usual convex hull algorithms based on the Fourier–Motzkin elimination and (ii) similar results were already known in the field of integer 0/1 programming. In fact, in [31], the \( \mathcal{H} \)-representation of the complete probability polytope is obtained by finding the explicit form of the matrix \( N \) that we defined in lemma 9. Nevertheless, we believe our approach is interesting, since it is based on logical and probabilistic, rather than geometrical, notions and provides an alternative point of view which has not been explored, to our knowledge, in the context of Bell inequalities.

First, we present the \( \mathcal{H} \)-representation for the complete probability polytope associated with \( n \) propositions. It is given by the following set of \( 2^n \) tight inequalities:

\[
\begin{align*}
  h(\varepsilon_1, \ldots, \varepsilon_n) & \geq 0, \quad \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0, 1\}^n, \\
  \text{where } h(\varepsilon_1, \ldots, \varepsilon_n) \text{ is defined as } h(1, \ldots, 1, 0, \ldots, 0) &= p_{1-s} - \sum_{s+1 \leq l \leq s+k} p_{1-sl} + \sum_{s+1 \leq l_1 < l_2 \leq s+k} p_{1-sl_1l_2} + \cdots \\
  & \cdots + (-1)^{s-1} \sum_{s+1 \leq l_1 < \cdots < l_k \leq s+k} p_{1-sl_1l_2\cdots l_k} + (-1)^k p_{1-n},
\end{align*}
\]  

(A.1)

for all \( 1 \leq s \leq n \) and \( s+k = n \); all other values, except \( h(0, \ldots, 0) \), are defined by a permutation of indices \( \{1, \ldots, n\} \) providing the above form, i.e. \( (1, \ldots, 1, 0, \ldots, 0) \), for \( (\varepsilon_1, \ldots, \varepsilon_n) \); in the case \( s = 0 \), the first term on the rhs is 1.

The notion of correlation polytope has been introduced by Pitowsky [12] in terms of propositional logic; however, the same problem can be expressed in terms of free Boolean algebras with the identification (see [28]) of atomic proposition with free generators,
logic operations with Boolean operations, truth assignments with two-valued measures and probability assignments with normalized measures.

To prove that the above representation is correct, we first need the following basic fact.

Lemma 3. Let $\mathcal{B}$ be a finite Boolean algebra with $k$ atoms $\{a_1, \ldots, a_k\}$; then, a function $f : X \subset \mathcal{B} \rightarrow [0, 1]$ can be extended to normalized measure $\mu$ on $\mathcal{B}$ if and only if there are $k$ numbers $\lambda_1, \ldots, \lambda_k$, with $\lambda_i \geq 0$ and $\sum_{i=0}^{k} \lambda_i = 1$, such that

$$f = \sum_{i=1}^{k} \lambda_i \delta_{a_i|X},$$

(A.3)

where $\delta_{a_i}$ is the two-valued measure that is 1 on $a_i$, $\mu$ is given by

$$\mu = \sum_{i=1}^{k} \lambda_i \delta_{a_i},$$

(A.4)

and the coefficients correspond to measure of atoms, i.e. $\lambda_i = \mu(a_i)$.

As a consequence, Pitowsky’s result can be expressed as follows.

Proposition 4. Let $\mathcal{B}$ be a Boolean algebra freely generated by $\mathcal{G} = \{A_1, \ldots, A_n\}$ and consider $X \subset \mathcal{B}$ with $X = \{A_1, \ldots, A_n, A_1 \cap A_j, \ldots, A_i \cap A_j \cap A_k, \ldots\}$, i.e. $X = \mathcal{G} \cup \mathcal{S}_2 \cup \cdots \mathcal{S}_m$, $m \leq n$, where elements of $\mathcal{S}_l$ are the intersections of $l$ distinct generators, but not necessarily all of those possible, i.e. $|\mathcal{S}_l| \leq \binom{n}{l}$. Now consider $f : X \rightarrow [0, 1]$ and define the vector

$$p = (p_1, \ldots, p_m, \ldots p_{ij}, \ldots, p_{i_1 \cdots i_m}, \ldots) \in \mathbb{R}^{|X|},$$

(A.5)

which has as components the values assumed by $f$ on $X$, namely

$$p_i = f(A_i), \quad i = 1, \ldots, n,$$

(A.6)

$$p_{ij} = f(A_i \cap A_j), \quad A_i \cap A_j \in \mathcal{S}_2, \ldots$$

(A.7)

$$p_{i_1 \cdots i_m} = f(A_{i_1} \cap \cdots \cap A_{i_m}), \quad A_{i_1} \cap \cdots \cap A_{i_m} \in \mathcal{S}_m.$$ (A.8)

For every $\varepsilon \in \{0, 1\}^n$ define the vector $u_\varepsilon \in \{0, 1\}^{|X|}$ given by

$$u_\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m, \ldots, \varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_m}, \ldots),$$

(A.9)

i.e. for every component $p_{i_1 \cdots i_m}$ of $p$ there is a corresponding component of $u_\varepsilon$ given by $\varepsilon_{i_1}, \ldots, \varepsilon_{i_m}$.

Then, $f$ can be extended to a normalized measure on $\mathcal{B}$ if and only if there are $2^n$ numbers $\lambda(\varepsilon)$, $\varepsilon \in \{0, 1\}^n$, such that

$$p = \sum_{\varepsilon \in \{0, 1\}^n} \lambda(\varepsilon)u_\varepsilon,$$

(A.10)

with

$$\lambda(\varepsilon) \geq 0 \quad \forall \varepsilon \in \{0, 1\}^n \quad \text{and} \quad \sum_{\varepsilon \in \{0, 1\}^n} \lambda(\varepsilon) = 1.$$ (A.11)

The convex hull of the vectors $u_\varepsilon$ defined above is known as the correlation polytope; clearly, each $A_i$ is associated with a QM observable, each set $\mathcal{S}_l$ to a joint measurements of $k$ observables, and numbers $p_{i_1 \cdots i_k}$ to probabilities for the corresponding measurements.

If $X$ contains all possible intersections between generators, i.e. $|X| = \sum_{l=1}^{n} \binom{n}{l} = 2^n - 1$, it will be called a complete set of measurements and the corresponding polytope will be called the complete probability polytope, and it will be denoted as $\mathcal{CP}_n$; such names come from the fact that there is a bijection between the points of the polytope and the normalized (i.e. probability) measures on the Boolean algebra freely generated by $A_1, \ldots, A_n$, as a consequence of the following general fact [15].
Lemma 5. Let $\mathcal{B}$ be a Boolean algebra freely generated by $\{A_1, \ldots, A_n\}$, let $\mu$ be a normalized measure on $\mathcal{B}$ and let $X \subset \mathcal{B}$ be a complete set of measurements. Then, the measure $\mu$ is uniquely defined by the values it assumes on the set $X$.

Now consider a Boolean algebra $\mathcal{B}$ freely generated by $\{A_1, \ldots, A_n\}$, a complete set of measurement $X \subset \mathcal{B}$ and a function $f : X \rightarrow [0, 1]$. If $f$ admits extension to a normalized measure $\mu$, then the measure of the atoms of $\mathcal{B}$, i.e. the elements $A_i^j \cap \cdots \cap A_n^j$, where $\epsilon_i \in \{0, 1\}$ and $A_i^j$ is defined as $A_i$ if $\epsilon_i = 1$ and $A_i^j$ (the Boolean complement) otherwise, can be written as

$$
\begin{align*}
\mu(A_1 \cap \cdots \cap A_n) &= f(A_1 \cap \cdots \cap A_n), \\
\mu(A_1 \cap \cdots \cap A_n^c) &= f(A_1 \cap \cdots \cap A_n) - f(A_1 \cap \cdots \cap A_n^c), \\
& \vdots \\
\mu(A_i^j \cap \cdots \cap A_n^j) &= 1 - \sum_i f(A_i) + \sum_{i<j} f(A_i \cap A_j) \\
& - \sum_{i<j<k} f(A_i \cap A_j \cap A_k) + \cdots + (-1)^{k-1} f(A_1 \cap \cdots \cap A_n), 
\end{align*}
$$

(A.12)

where the above expressions are obtained by using the identities $\mu(A) = \mu(A \cap B) + \mu(A \cap B^c)$ for all $A, B \in \mathcal{B}$ (additivity of measure on disjoint elements) and $\mu(1) = 1$, with 1 being the identity in $\mathcal{B}$ (normalization).

The exact expression for the values of the measure on atoms in terms of values on intersections is given by the following general result.

Lemma 6. Given a Boolean algebra $\mathcal{B}$ and $n \geq 2$ elements $B_1, \ldots, B_n \in \mathcal{B}$ and a measure $\mu$ on $\mathcal{B}$, it holds

$$
\begin{align*}
\mu(B_1 \cap \cdots \cap B_s \cap B_{s+1}^c \cap \cdots \cap B_{s+k}^c) &= \mu(B_1 \cap \cdots \cap B_s) \\
& - \sum_{s+1 \leq 1 \leq s+k} \mu(B_1 \cap \cdots \cap B_s \cap B_l) + \sum_{s+1 \leq 1 \leq s+k} \mu(B_1 \cap \cdots \cap B_s \cap B_l) \\
& + \cdots + (-1)^{k-1} \sum_{s+1 \leq 1 \leq s+k} \mu(B_1 \cap \cdots \cap B_s \cap B_l \cap B_{l+1} \cap \cdots \cap B_{n-1}) \\
& + (-1)^k \mu(B_1 \cap \cdots \cap B_n^c), 
\end{align*}
$$

(A.13)

with $s$ and $k$ such that $1 \leq s \leq n$, and $k = n - s$. In the case $s = 0$, i.e. $\mu(B_1^c \cap \cdots \cap B_n^c)$, an analogous expression holds, but with the first term $\mu(B_1 \cap \cdots \cap B_s)$ substituted by $\mu(1)$, which is equal to 1 if $\mu$ is a normalized measure.

Proof. Fixed $n \geq 2$, we proceed by induction on $k$. For $k = 0$, we have nothing to show. For $k = 1$, it is sufficient to use the identity $\mu(A \cap B^c) = \mu(A) - \mu(A \cap B)$.

For the inductive step, it is sufficient to define $B'_1 = B_1 \cap B_{s+k}$, apply the inductive hypothesis to $\mu(B'_1 \cap B_2 \cap \cdots \cap B_{s+k}^c)$ and then apply again the identity $\mu(A \cap B^c) = \mu(A) - \mu(A \cap B)$ to eliminate $B_{s+k}^c$.

It is obvious that the above argument does not depend on $n$. The proof for $s = 0$ is analogous.

Now consider a generic function $f : X \rightarrow [0, 1]$ and define $\tilde{X}$ as the set of atoms $\tilde{X} = \{A_1^i \cap A_2^j \cap \cdots \cap A_n^k \mid \epsilon_i \in \{0, 1\}\}$ and a function $\tilde{f} : \tilde{X} \rightarrow R$ in terms of $f$ in analogy with the result of lemma 6, i.e. for each $A_1^i \cap A_2^j \cap \cdots \cap A_n^k$, $\tilde{f}$ is defined as the rhs of the
Lemma 7. \( f \) admits an extension to a measure \( \mathcal{A} \). The first implication follows from the definition of \( f \). Moreover, if the two extensions exist, they coincide.

Proof. The first implication follows from the definition of \( \tilde{f} \); the other follows from equation (A.14) together with the identity

\[
1 = \sum_{\varepsilon_{i_1}, \ldots, \varepsilon_{i_n}} \tilde{f}(A_{i_1}^{\varepsilon_1} \cap \cdots \cap A_{i_n}^{\varepsilon_n}),
\]

where \((i_1, \ldots, i_n)\) is a permutation of \((1, \ldots, n)\). In particular, each of the two functions \( f \) and \( \tilde{f} \) is completely determined by the other one.

Extensions of \( f \) and \( \tilde{f} \) are related by the following.

\[
\boxed{\text{Lemma 7. } f \text{ admits an extension to a measure } \iff \tilde{f} \text{ admits an extension to a measure. Moreover, if the two extensions exist, they coincide.}}
\]

Proof. The first implication follows from the definition of \( \tilde{f} \); the other follows from equation (A.14) together with the identity

\[
A_{i_1} \cap \cdots \cap A_{i_n} = \bigcup_{\varepsilon_{i_1}, \ldots, \varepsilon_{i_n}} A_{i_1} \cap \cdots \cap A_{i_n} \cap A_{\varepsilon_{i_1}^{i_1+1} \cap \cdots \cap A_{\varepsilon_{i_n}^{i_n}},
\]

for all \((i_1, \ldots, i_n)\) permutation of \((1, \ldots, n)\) and all \(1 \leq k \leq n\), since the above is a disjoint union.

To conclude, it is sufficient to note that a measure is completely defined by the values it assumes on the atoms of the Boolean algebra (since each element can be written in a unique way as a disjoint union of atoms) and that each of the two functions is uniquely determined by the other one.

We are now able to prove the main result.

Theorem 8 (H-representation of \( \mathcal{CP}_n \)).

A vector \( p = (p_1, \ldots, p_n, \ldots, p_{n-1}, \ldots, p_1) \in \mathbb{R}^{2^n-1} \) belongs to \( \mathcal{CP}_n \) \( \iff \) its components satisfy the following \( 2^n \) inequalities:

\[
\begin{align*}
0 & \leq p_1 - n \\
0 & \leq p_1-(n-1) - p_1-(n-1)n \\
& \quad \vdots \\
0 & \leq 1 - \sum_{i=n}^{i=n} p_i + \sum_{i=n}^{i=n} p_{ij} - \sum_{i=k} p_{ijk} + \cdots,
\end{align*}
\]

where the rhs of inequalities (A.17) are obtained from the rhs of equations (A.12) by substituting \( f(A_{i_1} \cap \cdots \cap A_{i_k}) \) with \( p_{i_1\ldots i_k} \) for all \( \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\} \).

Proof. Given \( p \in \mathbb{R}^{2^n-1} \), consider the function \( f : X \subset \mathcal{B} \rightarrow \mathbb{R} \), where \( \mathcal{B} \) is the Boolean algebra freely generated by \( \{A_1, \ldots, A_n\} \) and \( X \) is a complete set of measurements, defined as

\[
f(A_{i_1} \cap \cdots \cap A_{i_k}) = p_{i_1\ldots i_k}, \text{ for all } \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}.
\]

Let us assume that \( f \) takes values in the interval \([0, \ 1]\); we shall prove it later. Then, by proposition 4, the above problem amounts to the problem of the existence of a measure extending \( f \) on \( \mathcal{B} \) and, therefore, by lemma 7, to the problem of the existence of a measure extending \( f \). Since \( f \) is defined on the atoms of \( \mathcal{B} \) and each element can be written uniquely as a disjoint union of atoms, \( f \) can be defined on each \( B \in \mathcal{B} \) as the sum of the valued assumed on the atoms \( A_{i_1}^{\varepsilon_1} \cap \cdots \cap A_{i_n}^{\varepsilon_n} \subset B \). In this way, we obtain a unique extension of \( f \) on \( \mathcal{B} \) which is
additive on disjoint elements and normalized by construction (see equation (A.15)). It follows that such an extension is a measure if it is non-negative if \( f \) is non-negative on the elements of \( X \). By construction of \( f \), see equation (A.12), such a condition is equivalent to the system of inequalities (A.17).

To conclude, we just have to show that the function \( f \) defined above takes values in \([0, 1] \). Since inequalities (A.17) amount to the condition \( f \geq 0 \), by equation (A.14), it follows that \( f \geq 0 \). This implies, by equation (A.15), also that \( f \leq 1 \).

We now prove that the above inequalities are tight, i.e. they cannot be represented as a positive weighted sum of other inequalities and define facets of the polytope. First, we need the following.

**Lemma 9.** \( CP_n \) is full dimensional in \( \mathbb{R}^{2^n-1} \) for all \( n \).

**Proof.** Given \( p \in CP_n \), it can be written, with the notation of proposition 4, as \( p = \sum \lambda(\varepsilon) u_\varepsilon \). Now, consider the vectors \( p', u'_\varepsilon, \lambda \in \mathbb{R}^2 \) defined as \( p' = (p, 1), u'_\varepsilon = (u_\varepsilon, 1) \) and \( \lambda \) as the vector given by the \( 2^n \) values \( \lambda(\varepsilon) \) for \( \varepsilon \in \{0, 1\}^n \). Note that, by lemma 3, the vector \( \lambda \) is given by the values assumed by \( f \).

Given \( p \in CP_n \), by definition of \( CP_n \) and equation (A.15), there exists a \( 2^n \times 2^n \) matrix \( M \), which columns are given by the vectors \( u'_\varepsilon \), such that \( p' = M \lambda \). On the other hand, by definition of \( f \), each component of \( \lambda \) can be written as a linear combination of the components of \( p' \), i.e. there exists a matrix \( N \), such that \( \lambda = N p' \). Since the set of admissible values for \( \lambda \), see proof of theorem 8, contains a basis for \( \mathbb{R}^2 \), e.g., the canonical basis, it follows that \( N = M^{-1} \).

As a consequence, the columns of \( M \), i.e. the vectors \( u'_\varepsilon \), are linearly independent. This implies that the dimension of the subspace spanned by \( u_\varepsilon \) is \( 2^n - 1 \). □

In order to show that inequalities (A.17) are tight, it is sufficient to note that \( CP_n \) has exactly \( 2^n \) facets. It is sufficient to count the number of affine hyperplanes, i.e. affine subspaces of dimension \( 2^n - 2 \), generated by the vertices \( u_\varepsilon \). By the above lemma, the vectors \( u_\varepsilon \) are affinely independent; therefore, each subset of \( 2^n - 1 \) vectors defines an affine hyperplane. Therefore, the number of facets of \( CP_n \) is \( \binom{2^n-1}{2^n-2} = 2^n \).

**References**

[1] Bell J S 1964 *Physics* 1 195
[2] Brukner Č, Žukowski M, Pan J-W and Zeilinger A 2004 *Phys. Rev. Lett.* 92 127901
[3] Buhrman H, Cleve R, Massar D and de Wolf R 2010 *Rev. Mod. Phys.* 82 665
[4] Barrett J, Hardy L and Kent A 2005 *Phys. Rev. Lett.* 95 010503
[5] Acin A et al 2007 *Phys. Rev. Lett.* 98 230501
[6] Pironio S et al 2010 *Nature* 464 1021
[7] Klyachko A A, Can M A, Binicoglu S and Shumovsky A S 2008 *Phys. Rev. Lett.* 101 20403
[8] Budroni C and Morchio G 2010 *J. Math. Phys.* 51 122205
[9] Horn A and Tarski A 1948 *Trans. Am. Math. Soc.* 64 467
[10] Deza M M and Laurent M 1997 *Geometry of Cuts and Metrics* (Algorithms and Combinatorics vol 15) (Berlin: Springer)
[18] Gühne O et al 2010 Phys. Rev. A 81 022121
[19] Pitowsky I and Svozil K 2001 Phys. Rev. A 64 014102
[20] Collins D and Gisin N 2004 J. Phys. A: Math. Gen. 37 1775
[21] cdd http://www.ifor.math.ethz.ch/fukuda/cdd_home/cdd.html
[22] lrs http://cgm.cs.mcgill.ca/avis/C/lrs.html
[23] porta http://www.zib.de/Optimization/Software/porta/
[24] Avis D, Imai H, Ito T and Sasaki Y 2004 arXiv:quant-ph/0404014
[25] Avis D, Imai H, Ito T and Sasaki Y 2005 J. Phys. A: Math. Gen. 38 10971
[26] Ziegler G M 1995 Lectures on Polytopes (New York: Springer)
[27] Sliwa C 2003 Phys. Lett. A 317 165
[28] Givant S and Halmos P 2009 Introduction to Boolean Algebras (New York: Springer)
[29] Budroni C and Morchio G 2012 Found. Phys. 42 544
[30] FM http://www.cse.ohio-state.edu/pouchet/software/fm/
[31] Laurent M and Rendl F 2005 Semidefinite programming and integer programming Handbook on Discrete Optimization (Amsterdam: Elsevier)