Mathematical Programming 50 (1991) 395-414
North-Holland

Correlation polytopes: Their geometry and complexity

Itamar Pitowsky

Department of Philosophy, The Hebrew University, 91904 Jerusalem, Israel

Received 16 May 1988
Revised manuscript received 31 July 1989

A family of polytopes, correlation polytopes, which arise naturally in the theory of probability and propositional logic, is defined. These polytopes are tightly connected to combinatorial problems in the foundations of quantum mechanics, and to the Ising spin model. Correlation polytopes exhibit a great deal of symmetry. Exponential size symmetry groups, which leave the polytope invariant and act transitively on its vertices, are defined. Using the symmetries, a large family of facets is determined. A conjecture concerning the full facet structure of correlation polytopes is formulated (the conjecture, however, implies that NP = co-NP).

Various complexity results are proved. It is shown that deciding membership in a correlation polytope is an NP-complete problem, and deciding facets is probably not even in NP. The relations between the polytope symmetries and its complexity are indicated.

Key words: Convex polytopes, symmetries, NP, co-NP.

1. Correlation polytopes

1.1. Definitions and notations

Let $n \geqslant 2$ be a natural number and let $S \subseteq K_n = \{\{i, j\};\ 1 \leqslant i < j \leqslant n\}$. We shall denote by (n, S) the (simple) graph whose vertices are $1, 2, \ldots, n$ and edges $\{i, j\} \in S$. Let $\mathscr{R}(n, S)$ denote the real linear space of all functions $f: \{1, 2, \ldots, n\} \cup S \to \mathscr{R}$, clearly $\dim \mathscr{R}(n, S) = n + |S| \leqslant \frac{1}{2} n(n+1)$. We shall denote vectors in $\mathscr{R}(n, S)$ by $f = (f_1, f_2, \ldots, f_n, \ldots, f_{ij}, \ldots)$ where the numbers $f_{ij}, \{i, j\} \in S$, appear in lexicographic order on the i, j's.

For $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \in \{0, 1\}^n$, let u^ε denote the following vector of $\mathscr{R}(n, S)$:

$$u_i^\varepsilon = \varepsilon_i, \quad 1 \leqslant i \leqslant n, \qquad u_{ij}^\varepsilon = \varepsilon_i \varepsilon_j, \quad \{i, j\} \in S. \tag{1.1}$$

Definition 1.1. The correlation polytope $c(n, S)$ is the (closed) convex hull, in $\mathscr{R}(n, S)$, of the vectors u^ε, $\varepsilon \in \{0, 1\}^n$. In case $S = K_n$ is the set of all pairs, we put $c(n, K_n) = c(n)$. $c(n)$ is the "full correlation polytope".

The coordinates of each vector $p \in c(n, S)$ are bounded between zero and one. Hence, it is obvious that each vector u^ε is a vertex of $c(n, S)$. The name "correlation polytope" is justified by the following theorem.

Theorem 1.1. *Let $p \in \mathcal{R}(n, S)$ then $p \in c(n, S)$ if and only if there exists a probability space (X, Σ, μ) and (not necessarily distinct) events $A_1, A_2, \ldots, A_n \in \Sigma$ such that*

$$p_i = \mu(A_i), \quad 1 \leq i \leq n, \qquad p_{ij} = \mu(A_i \cap A_j), \quad \{i, j\} \in S.$$

Proof. Suppose there exists a probability space (X, Σ, μ), events $A_1, \ldots, A_n \in \Sigma$ s.t. $\mu(A_i) = p_i$, $i = 1, 2 \ldots n$, $\mu(A_i \cap A_j) = p_{ij}$, $\{i, j\} \in S$. For an arbitrary $B \in \Sigma$ put $B^1 = B$, $B^0 = \tilde{B} = X \setminus B$, and for $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0, 1\}^n$ put $A(\varepsilon) = A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap \cdots \cap A_n^{\varepsilon_n}$. Then clearly, $A(\varepsilon) \cap A(\varepsilon') = \emptyset$ for $\varepsilon \neq \varepsilon'$, and $\bigcup_{\varepsilon \in \{0,1\}^n} A(\varepsilon) = X$. Also, $A_i = \bigcup_{\varepsilon \in \{0,1\}^n} \{A(\varepsilon); \varepsilon_i = 1\}$. Put $\lambda(\varepsilon) = \mu(A(\varepsilon))$, then $\lambda(\varepsilon) \geq 0$ and $\sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon) = 1$, $p_i = \mu(A_i) = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i$, $p_{ij} = \mu(A_i \cap A_j) = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i\varepsilon_j$, for $i = 1, 2 \ldots n$ and $\{i, j\} \in S$, hence $p = \sum \lambda(\varepsilon) u^\varepsilon \in c(n, S)$. Conversely, suppose that $p \in c(n, S)$, then we can represent $p = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon) u^\varepsilon$ where $\lambda(\varepsilon) \geq 0$, $\sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon) = 1$. Let $X = \{0, 1\}^n$, let Σ be the power set of X, define $\mu(B) = \sum_{\varepsilon \in B} \lambda(\varepsilon)$ for all $B \subseteq X$. Put $A_i = \{\varepsilon \mid \varepsilon_i = 1\}$; then

$$\mu(A_i) = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i = p_i \quad \text{and} \quad \mu(A_i \cap A_j) = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i\varepsilon_j = p_{ij}. \qquad \square$$

From the "subjectivist" point of view of probability the polytope $c(c, S)$ can be interpreted as follows: let $x_1, x_2, \ldots, x_n$ be distinct Boolean variable, or "atomic propositions", and consider the conjunctions "x_i and x_j", $\{i, j\} \in S$. Each $\varepsilon \in \{0, 1)^n$ represents a truth value assignment to the propositions $x_1, \ldots, x_n$, and thus also for the pairs "x_i and x_j". Hence, the vertices of $c(n, S)$ are all the possible truth assignments, and every $p \in c(n, S)$ is nothing but a weighted average of these truth assignments. This intuition will enable us to associate membership in $c(n, S)$ with a satisfiability problem, since a (complex) proposition is satisfiable if and only if it can be assigned positive probability in some probability space.

The polytope $c(n, S)$ has non-empty interior. To see that, consider the vectors u^ε for

$$\varepsilon = (0, \ldots, 0, \overset{i}{1}, 0 \ldots 0), \quad 1 \leq i \leq n,$$

and

$$\varepsilon = (0 \ldots, \overset{i}{1}, 0 \ldots, \overset{j}{1}, 0 \ldots 0) \quad \text{for } \{i, j\} \in S.$$

There are $n + |S|$ such vectors, and it is clear that they are linearly independent in $\mathcal{R}(n, S)$.

1.2. Examples

(a) For $n = 2$, $S = \{\{1, 2\}\}$, $c(2, S)$ is the simplex in $\mathcal{R}(2, S)$ with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ and $(1, 1, 1)$. The inequalities for $c(2, S)$ are

$$0 \leq p_{12} \leq \min(p_1, p_2), \qquad p_1 + p_2 - p_{12} \leq 1.$$

(b) Let $n = 3$ and $S = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$. The inequalities for $c(3, S) = c(3)$ are

$$0 \leq p_{ij} \leq \min(p_i, p_j), \quad 1 \leq i < j \leq 3,$$

$$p_i + p_j - p_{ij} \leq 1, \quad 1 \leq i < j \leq 3,$$

$$
\begin{aligned}
&p_1+p_2+p_3-p_{12}-p_{13}-p_{23}\leqslant 1,\\
&p_1-p_{12}-p_{13}+p_{23}\geqslant 0,\\
&p_2-p_{12}-p_{23}+p_{13}\geqslant 0,\\
&p_3-p_{13}-p_{23}+p_{12}\geqslant 0.
\end{aligned}
\tag{1.2}
$$

Necessity follows directly from the fact that all u^{ε}, $\varepsilon=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\in\{0,1\}^3$ satisfy these inequalities. As for sufficiency, let p satisfy all these inequalities, and let η be a real number (which will subsequently play the role of $\mu(A_1\cap A_2\cap A_3)$ in the notations of Theorem 1.1),

$$\eta\leqslant\min\{p_{12},p_{13},p_{23},1-(p_1+p_2+p_3-p_{12}-p_{13}-p_{23})\},$$

$$\eta\geqslant\max\{0,-p_1+p_{12}+p_{13},-p_2+p_{12}+p_{23},-p_3+p_{13}+p_{23}\}.$$

Such a choice is possible because of the above inequalities. Now define $\lambda(\varepsilon)=\lambda(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ by

$$\lambda(0,0,0)=1-(p_1+p_2+p_3-p_{12}-p_{13}-p_{23})-\eta,\qquad \lambda(1,1,1)=\eta,$$

$$\lambda(100)=\eta+(p_1-p_{12}-p_{13}),\qquad \lambda(110)=p_{12}-\eta,$$

$$\lambda(010)=\eta+(p_2-p_{12}-p_{23}),\qquad \lambda(101)=p_{13}-\eta,$$

$$\lambda(001)=\eta+(p_3-p_{13}-p_{23}),\qquad \lambda(011)=p_{23}-\eta.$$

It is easy to see that

$$\lambda(\varepsilon)\geqslant 0,\qquad \sum_{\varepsilon\in\{0,1\}^3}\lambda(\varepsilon)=1,\qquad \sum_{\varepsilon\in\{0,1\}^3}\lambda(\varepsilon)u^{\varepsilon}=p.$$

I shall call this polytope the Bell–Wigner polytope. Inequalities (1.2), and some of their generalizations, play an important role in the controversy concerning the interpretation of quantum theory [1–8].

(c) Clauser et al. [3], and following them Clauser and Horne [4], extended (1.2) to the following case: let $n=4$ and $S=\{\{1,3\},\{1,4\},\{2,3\},\{2,4\}\}$. The following inequalities are necessary and sufficient for $p\in\mathscr{R}(4,S)$ to be an element of $c(4,S)$:

$$
\begin{aligned}
&0\leqslant p_{ij}\leqslant\min(p_i,p_j),\quad i=1,2,\ j=3,4,\\
&p_i+p_j-p_{ij}\leqslant 1,\quad i=1,2,\ j=3,4,\\
&-1\leqslant p_{13}+p_{14}+p_{24}-p_{23}-p_1-p_4\leqslant 0,\\
&-1\leqslant p_{23}+p_{24}+p_{14}-p_{13}-p_2-p_4\leqslant 0,\\
&-1\leqslant p_{14}+p_{13}+p_{23}-p_{24}-p_1-p_3\leqslant 0,\\
&-1\leqslant p_{24}+p_{23}+p_{13}-p_{14}-p_2-p_3\leqslant 0.
\end{aligned}
\tag{1.3}
$$

The inequalities (1.3) are called the Clauser–Horne inequalities.

Sufficiency was proved by Fine [5]. Mermin and his students attempted to generalize these inequalities to higher dimensional cases. In particular, one should mention [6] where the first connection between Bell-type inequalities and linear programming is made. Correlation polytopes were introduced in [7]. There, the relation between these polytopes, classical logic, and quantum logic is indicated. (See also [8].)

1.3. Ising spin

Another, less apparent source of correlation polytopes, has to do with the ising spin model. This model is used in a variety of applications: in statistical mechanics to model spin glasses [9], in the theory of computation to model "connectionist" computers, or even the collective action of neural networks [10]. The amount of published work on these issues is vast.

An ising spin system is a set of n cites, at each cite i there is a component capable of taking two ising spin values $\varepsilon_i = 0$ or $\varepsilon_i = 1$ (often the values ± 1 are taken, but the transformation is clear). Between the ith and jth cites there is an "action potential" J_{ij} (a real number), and sometimes the ith cite is subjected to an "external field" J_i. If, at a given moment the values of the ising spins are given by $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0, 1\}^n$, the energy of the system at that moment is

$$W(\varepsilon) = \sum_{i=1}^{n} J_i \varepsilon_i + \sum_{1 \leqslant i < j \leqslant n} J_{ij} \varepsilon_i \varepsilon_j.$$

A typical combinatorial problem associated with the model is to determine the minimum energy of a given ising spin system (minimum, that is, over all possible configurations $\varepsilon \in \{0, 1\}^n$). To see the connection with correlation polytopes, consider the vector $(J_1, \ldots, J_n, \ldots, J_{ij}, \ldots) \in \mathscr{R}(n, K_n)$, where K_n is the set of all pairs. Rather than taking the discrete optimization problem of the minimum energy, consider the linear program

$$\text{minimize} \quad \sum_{i=1}^{n} J_i p_i + \sum_{1 \leqslant i < j \leqslant n} J_{ij} p_{ij}$$

$$\text{subject to} \quad \text{the constraint that } p = (p_1, \ldots, p_n, \ldots, p_{ij}, \ldots) \in c(n).$$

The linear program and the discrete problem are clearly equivalent. If we attempt to use linear programming in order to establish the minimum energy of the ising spin system, we have to compute first the facet inequalities for $c(n)$. The minimum energy problem was proved to be NP-hard by Barahona [11]. I shall provide a very short proof of this fact in the third section. This seems to indicate that deriving all inequalities for $c(n)$ is a very difficult task.

1.4. Boole's problem

The search for the facet inequalities of $c(n)$ has a long history, though the problem had been phrased in a probablistic rather than geometric terminology. Let (X, Σ, μ)

be a probability space, and let $A_1, \ldots, A_n \in \Sigma$. A typical question asked by George Boole [12] is the following: suppose that we are given the values of $p_i = \mu(A_i)$ $1 \leq i \leq n$, but have no further information. What then is the best possible estimation of $\mu(A_1 \cup \cdots \cup A_n)$? The answer is

$$\max[p_1, \ldots, p_n] \leq \mu(A_1 \cup \cdots \cup A_n) \leq \min[1, p_1 + \cdots + p_n].$$

Boole considered other similar problems, attempting to derive the best possible bounds for the probability of certain Boolean functions of events, given the values of others. He often applied methods which can be identified today as primitive forms of linear programming. (Further details on this can be found in [13, 14].) Consider the following generalization of Boole's problem: given the values of $p_i = \mu(A_i)$, $1 \leq i \leq n$, and $p_{ij} = \mu(A_i \cap A_j)$, $1 \leq i < j \leq n$, but no further information, what then is the best lower bound for $\mu(A_1 \cup A_2 \cup \cdots \cup A_n)$? Following the notations of Theorem 1.1 we know that $\mu(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{\varepsilon \neq 0} \lambda(\varepsilon)$, where for each $0 \neq \varepsilon \in \{0, 1\}^n : \lambda(\varepsilon) \geq 0$ and $\sum_\varepsilon \lambda(\varepsilon)\varepsilon_i = p_i, \sum_\varepsilon \lambda(\varepsilon)\varepsilon_i\varepsilon_j = p_{ij}$. Hence, the best bound is given by the linear program

$$\min \sum_{\varepsilon \neq 0} \lambda(\varepsilon).$$

Constraints:

$$\lambda(\varepsilon) \geq 0, \qquad \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i = p_i, \qquad \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)\varepsilon_i\varepsilon_j = p_{ij}.$$

The dual program is

$$\max\left[\sum_{i=1}^{n} x_i p_i + \sum_{1 \leq i < j \leq n} x_{ij} p_{ij}\right].$$

Constraints:

$$-\infty < x_i, x_{ij}, < \infty, \qquad \sum_{i=1}^{n} x_i\varepsilon_i + \sum_{1 \leq i < j \leq n} x_{ij}\varepsilon_i\varepsilon_j \leq 1 \quad \text{for all } \varepsilon \in \{0, 1\}^n.$$

By assumption, $p = (p_1, \ldots, p_n, \ldots, p_{ij}, \ldots) \in c(n)$, hence the primal and the dual are both solvable, and have identical optimal value ≤ 1. Note that the constraints of the dual program define the polar $c^*(n)$ of $c(n)$:

$$c^*(n) = \left\{ x \in \mathcal{R}^*(n, K_n) \;\middle|\; \sum_{i=1}^{n} x_i\varepsilon_i + \sum_{1 \leq i < j \leq n} x_{ij}\varepsilon_i\varepsilon_j \leq 1, \varepsilon \in \{0, 1\}^n \right\}.$$

$c^*(n)$ is obviously unbounded. Since 0 is an element of $c(n)$ — albeit not an internal point — we have $(c^*(n))^* = c(n)$. Therefore, the extreme points of $c^*(n)$ define facets of $c(n)$, indeed, all the facets of $c(n)$ of which 0 is not an element. We shall see that unless NP = co-NP the determination of these facets is an intractable problem, and therefore the above form of Boole's problem is probably intractable as well.

Some lower bounds for $\mu(A_1 \cup \cdots \cup A_n)$, given the values of p_i, p_{ij} have been known for years. Bonferroni [15] proved

$$\sum_{i=1}^{n} p_i - \sum_{1 \leq i < j \leq n} p_{ij} \leq \mu(A_1 \cup \cdots \cup A_n). \tag{1.4}$$

Chung [16] generalized this formula:

$$\frac{2}{k+1} \sum_{i=1}^{n} p_i - \frac{2}{k(k+1)} \sum_{1 \leq i < j \leq n} p_{ij} \leq \mu(A_1 \cup \cdots \cup A_n) \tag{1.5}$$

whenever $1 \leq k \leq n-1$. We shall see that these inequalities give rise to exponentially many independent inequalities when we apply the symmetries of $c(n)$. Some of these facts were summarized in an influential monograph by Fréchet [17]. Research on this problem continues to the present day [18–21].

1.5. Statement of results

It is easy to see that every face of $c(n, S)$, $S \subseteq K_n$ induces a face of $c(n)$, so that $c(n)$ reflects in its face structure all the polytopes $c(n, S)$. $c(n)$ has a large symmetry group, of cardinality $n!2^n$, which operates transitively on its vertices. I shall identify these symmetries in Section 2.1. Using the symmetries we can prove that the edge graph of $c(n)$ is the complete graph on 2^n vertices.

If we manage to guess one inequality for $c(n)$, we automatically establish exponentially many facets by application of the symmetries. In Section 2.2 a large family of facets of $c(n)$ is determined, a conjecture regarding the total facet structure of $c(n)$ is formulated. As we shall see, this conjecture entails that NP = co-NP, so it is probably false. In Section 2.4 some generalizations are proved. Chapter 3 is devoted to a complexity study, where the following decision problems are considered

CORRELATION
Instance: a (rational) vector $p \in R(n, K_n)$ (K_n-is the set of all pairs).
Question: is $p \in c(n)$?

I shall prove that CORRELATION is NP-complete. This means that unless NP = co-NP, deriving all the inequalities for $c(n)$ is an impossible task. The situation is even worse in a sense; consider the following decision problem:

CORRELATION FACET
Instance: an inequality $\langle a, x \rangle \leq b$, where $a \in R(n, K_n)$ and b are integral.
Question: is it a facet of $c(n)$?

It follows from Barahona's result [11], in conjunction with the theorem of Karp and Papadimitriou [22], that if CORRELATION FACET $\in$ NP, then NP = co-NP.

A short proof of Barahona's theorem and this consequence is given in Section 3.4.

2. Geometry

2.1. The groups $I(n, S)$ and $A(n, S)$

Let $p \in c(n, S)$, by Theorem 1.1 there is a probability space (X, Σ, μ), events $A_1, \ldots, A_n \in \Sigma$, such that $\mu(A_i) = p_i$, $\mu(A_i \cap A_j) = p_{ij}$, for $1 \leq i \leq n$ and $\{i, j\} \in S$. Consider the transformation $A_i \to \tilde{A}_i = X \setminus A_i$ for some fixed i. We have $\mu(\tilde{A}_i) = 1 - \mu(A_i)$ and $\mu(\tilde{A}_i \cap A_j) = \mu(A_j) - \mu(A_i \cap A_j)$, hence the affine transformation σ_i defined on $\mathcal{R}(n, S)$ by

$$
\begin{aligned}
&(\sigma_i p)_i = 1 - p_i, \\
&(\sigma_i p)_j = p_j, \quad j \neq i, \\
&(\sigma_i p)_{ij} = p_j - p_{ij} \text{ for all } j \text{ such that } \{i, j\} \in S, \\
&(\sigma_i p)_{jk} = p_{jk}, \quad \{j, k\} \in S,\ j, k \neq i,
\end{aligned}
\tag{2.1}
$$

leaves the polytope $c(n, S)$ invariant. Obviously, σ_i^2 is the identity, and σ_i, $i = 1, 2, \ldots, n$, generate a commutative group of involutions isomorphic to $Z_2^{(n)}$. We denote this group by $I(n, S)$. For $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \in \{0, 1\}^n$ put $\sigma(\varepsilon) = \prod_{i=1}^n \sigma_i^{\varepsilon_i}$, and the $\sigma(\varepsilon)$, $\varepsilon \in \{0, 1\}^n$, are all the distinct elements of $I(n, S)$. Obviously, the elements of $I(n, S)$ transform a vertex of $c(n, S)$ to another vertex. Moreover, the group $I(n, S)$ acts transitively on the vertices of $c(n, S)$: if $u^0 = 0$ then $\sigma(\varepsilon)u^0 = u^\varepsilon$.

Consider next the automorphism group of the graph (n, S), that is, the group of permutations $\pi : \{1, 2. \ldots, n\} \to \{1, 2, \ldots, n\}$ such that $\{i, j\} \in S$ if and only if $\{\pi(i), \pi(j)\} \in S$. Each such automorphism π induces a linear map on $\mathcal{R}(n, S)$, which we also denote by π, by $f_i \to f_{\pi(i)}$, $1 \leq i \leq n$, and $f_{ij} \to f_{\pi(i), \pi(j)}$ for $\{i, j\} \in S$. Clearly, the transformation π leaves the polytope $c(n, S)$ invariant, since its effect is to permute the events $A_1, \ldots, A_n$. Obviously, π maps a vertex of $c(n, S)$ to another vertex. Denote by $A(n, S)$ the group of all transformations induced by the automorphisms of the graph (n, S), and let $G(n, S)$ be the group generated by both $A(n, S)$ and $I(n, S)$. The mixed action of the groups $I(n, S)$ and $A(n, S)$ is easy to identify, if $\pi \in A(n, S)$ and $\sigma(\varepsilon) = \prod_{i=1}^n \sigma_i^{\varepsilon i} \in I(n, S)$, then $\pi\sigma\pi^{-1} = \prod_{i=1}^n \sigma_{\pi(i)}^{\varepsilon i}$ so that $I(n, S)$ is normal in $G(n, S)$.

In case $S = K_n$ is the set of all pairs, we denote $I(n, K_n) = I(n)$, $A(n, K_n) = A(n)$ is the full permutation group $A(n) \simeq S_n$, and $G(n, K_n) = G(n)$ has cardinality $n!2^n$.

Using the symmetries we can easily prove:

Lemma 2.1. *Let $\varepsilon^1, \varepsilon^2 \in \{0, 1\}^n$, $\varepsilon^1 \neq \varepsilon^2$; then the interval joining u^{ε^1} and u^{ε^2} is a face of $c(n)$, so that $c(n)$ is a 1-neighbourly polytope.*

Proof. The lemma is valid if and only if the interval joining ηu^{ε^1} and ηu^{ε^2} is a face of $c(n)$, whenever $\eta \in G(n)$. Take $\eta = \sigma(\varepsilon^1)$, then $\sigma(\varepsilon^1)u^{\varepsilon^1} = 0$, apply a suitable permutation π so that $\pi\sigma(\varepsilon^1)u^{\varepsilon^2} = u^\delta$, for δ of the form $\delta = (1, \ldots, 1, 0, \ldots, 0)$, that is $\delta_i = 1$ for $1 \leq i \leq k$ and $\delta_i = 0$ for $k < i \leq n$, $k \geq 1$. Since $\pi 0 = 0$, it is sufficient to

prove that the interval joining the origin with u^δ is a face of $c(n)$, where δ has the above form. Put

$$p = \tfrac{1}{2}u^\delta + \tfrac{1}{2}0. \tag{2.2}$$

We shall show that (2.2) is the unique representation of p as a convex combination of the vertices of $c(n)$, and conclude.

We have $p_i = \frac{1}{2}$, $1 \leqslant i \leqslant k$, $p_i = 0$, $k < i \leqslant n$, $p_{ij} = \frac{1}{2}$, $1 \leqslant i < j \leqslant k$, and $p_{ij} = 0$ for $i > k$ or $j > k$. Suppose that $p = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon) u^\varepsilon$, where $\lambda(\varepsilon) \geqslant 0$ and $\sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon) = 1$. Then $\lambda(\varepsilon) = 0$ whenever there is $k < i \leqslant n$ for which $\varepsilon_i = 1$. If $k = 1$, then $\lambda(\varepsilon) = 0$ for all $\varepsilon \neq 0, \delta$, and $\lambda(0) = \lambda(\delta) = \frac{1}{2}$, so the representation is unique. If $k \geqslant 2$, then $p_i = p_{ij} = p_j = \frac{1}{2}$ for all $1 \leqslant i < j \leqslant k$. Let $1 \leqslant i < j \leqslant k$; then $\lambda(\varepsilon)\varepsilon_i \neq 0$ entails $\lambda(\varepsilon)\varepsilon_i\varepsilon_j \neq 0$, therefore $\lambda(\varepsilon) > 0$ and $\varepsilon_i = 1$ entails $\varepsilon_j = 1$ for $1 \leqslant i < j \leqslant k$, hence $\lambda(\varepsilon) \neq 0$ if and only if $\varepsilon = 0$ or $\varepsilon = \delta$, in which case $\lambda(\varepsilon) = \frac{1}{2}$. □

2.2. *Some facets of* $c(n)$

The polytope $c(n, S)$ is obtained from $c(n)$ by a projection; simply by dropping the coordinates $p_{ij}\{i, j\} \notin S$. It follows that every face of $c(n, S)$ induces a face of $c(n)$. Hence, the face structure of $c(n)$ reflects the face structure of $c(n, S)$ for $S \subseteq K_n$. Let $1 \leqslant k \leqslant n$ and let $A_1, A_2, \ldots, A_k$ be events in a probability space (X, Σ, μ), then we have by Bonferroni inequalities (1.4),

$$\sum_{i=1}^{k} \mu(A_i) - \sum_{1 \leqslant i < j \leqslant k} \mu(A_i \cap A_j) \leqslant \mu(A_i \cup A_2 \cup \cdots \cup A_k) \leqslant 1.$$

Hence if $p \in c(n)$, we must have $\sum_{i=1}^{k} p_i - \sum_{1 \leqslant i < j \leqslant k} p_{ij} \leqslant 1$ for all $0 \leqslant k \leqslant n$. If this is valid for p this must also be valid for σp and πp for $\sigma \in I(n)$ and $\pi \in A(n)$. Summing up these facts we conclude:

$$\sum_{i \in \alpha} (\sigma p)_i - \sum_{\substack{i<j \\ ij \in \alpha}} (\sigma p)_{ij} \leqslant 1 \tag{2.3}$$

for all-non empty subsets $\alpha \subseteq \{1, 2, \ldots, n\}$ and all involutions $\sigma \in I(n)$. To see how these inequalities work take $\alpha = \{i\}$ then (2.3) reads $p_i \leqslant 1$, take $\alpha = \{i, j\}$ then $p_i + p_j - p_{ij} \leqslant 1$; applying σ_i to this inequality we get $(1 - p_i) + p_j - (p_j - p_{ij}) \leqslant 1$, or $p_{ij} \leqslant p_i$, similarity $p_{ij} \leqslant p_j$; applying $\sigma_i\sigma_j$ we get $(1 - p_i) + (1 - p_j) - (1 - p_i - p_j + p_{ij}) \leqslant 1$, or $p_{ij} \geqslant 0$. Take $\alpha = \{i, j, k\}$, then $p_i + p_j + p_k - p_{ij} - p_{ik} - p_{jk} \leqslant 1$, which is the first of the Bell inequalities (1.2); apply σ_i to obtain $p_i - p_{ij} - p_{ik} - p_{jk} \geqslant 0$, which is the type of the other three Bell inequalities (1.2).

Let $c^*(n)$ denote the polar of $c(n)$. We have already noted that since $0 \in c(n)$, we have $(c^*(n))^* = c(n)$. Using this fact we shall prove:

Theorem 2.2. *If* $|\alpha| \geqslant 2$, *inequality* (2.3) *represents a facet of* $c(n)$.

Proof. We assume $n \geqslant 3$ (for $n = 2$ the characterization of $c(2)$ is trivial). Let $f_i \in \mathcal{R}(n, Sn)$ be the vector which is 1 in the coordinate i and zero elsewhere, and

$f_{ij} \in \mathcal{R}(n, S_n)$ the vector which is 1 in the coordinate $\{i, j\}$ and zero elsewhere, let $2 \leq k \leq n$, put $F^{(k)} = \sum_{i=1}^{k} f_i - \sum_{1 \leq i < j \leq k} f_{ij}$. We shall prove $F^{(k)}$ is an extreme point of the polyhedron $c^*(n)$. Suppose $F^{(k)} = \frac{1}{2}a + \frac{1}{2}b$ for $a, b \in c^*(n)$, that is $\langle a, u^\varepsilon \rangle \leq 1$ and $\langle b, u^\varepsilon \rangle \leq 1$ for $\varepsilon \in \{0, 1\}^n$, we shall show $a = b = F^{(k)}$. Let $1 \leq i \leq k$, and take $\varepsilon \in \{0, 1\}^n$, which is 1 in the ith coordinate and zero elsewhere, we get $\langle a, u^\varepsilon \rangle = a_i \leq 1$, and similarly, $b_i \leq 1$ but $\frac{1}{2}(a_i + b_i) = 1$, so that $a_i = b_i = 1$, $1 \leq i \leq k$. Also, suppose $k < n$, and let $1 \leq i \leq k$, $k < l \leq n$. Let ε be 1 on i and l and zero elsewhere, then $a_i + a_l + a_{il} \leq 1$, or $a_l + a_{il} \leq 0$; similarly, $b_l + b_{il} \leq 0$, but $a_l + b_l = 0$, $a_{il} + b_{il} = 0$, hence $a_l + a_{il} = 0$ for $1 \leq i \leq k$ and $k < l \leq n$. We know $k > 1$. Let $1 \leq i < j \leq k$, then $a_i + a_j + a_{ij} \leq 1$, $a_i = a_j = 1$, hence $a_{ij} \leq -1$; also, $b_{ij} \leq -1$, but $\frac{1}{2}(a_{ij} + b_{ij}) = -1$, hence $a_{ij} = b_{ij} = -1$; therefore, the case is proved for $k = n$. Take $1 \leq i < j \leq k$, $k < l \leq n$, then $a_i + a_j + a_l + a_{ij} + a_{il} + a_{jl} \leq 1$, but $a_i = a_j = 1$, $a_{ij} = -1$, therefore $a_l + a_{il} + a_{jl} \leq 0$; also, $b_l + b_{il} + b_{jl} \leq 0$ and $a_l + b_l = 0$, $a_{il} + b_{il} = a_{jl} + b_{jl} = 0$, hence $a_l + a_{il} + a_{jl} = 0$; but we proved $a_l + a_{il} = 0$, hence $a_{jl} = 0$, and thus $a_l = 0$; and the case is proved for $k = n - 1$. For $k < n - 1$ take $k < l < r \leq n$ and $1 \leq i \leq k$, then $a_i + a_l + a_r + a_{il} + a_{ir} + a_{lr} \leq 1$; substituting the values already obtained we get $a_{rl} \leq 0$, since $a_{rl} + b_{rl} = 0$ we obtain by standard reasoning $a_{rl} = 0$, and thus $F^{(k)}$ is an extreme point of $c^*(n)$. Since $(c^*(n)))^* = c(n)$ we have

$$\sum_{i=1}^{k} p_i - \sum_{1 \leq i < j \leq k} p_{ij} \leq 1, \quad 2 \leq k \leq n,$$

represents a facet of $c(n)$. Since the operations of the group $G(n)$ (involutions and permutations) take a facet to a facet the claim follows. □

Note that the inequality $p_1 \leq 1$ is not a facet, since it is a consequence of other inequalities: from $p_1 + p_2 - p_{12} \leq 1$ we conclude (by applying σ_1) that $p_{12} \leq p_1$ and (by applying $\sigma_1\sigma_2$) that $p_{12} \geq 0$, hence $p_1 \geq 0$; applying σ_1 we get $1 - p_1 \geq 0$ or $p_1 \leq 1$.

We have proved (examples a, b) that inequalities (2.3) are sufficient for $n = 2, 3$. I believe that they are sufficient for $n = 4$ as well. In any case, inequalities (2.3) do not represent all the facets of $c(n)$ in the general case. Consider Chung inequality (1.5) for $k = 2$:

$$2 \sum_{i=1}^{n} p_i - \sum_{1 \leq i < j \leq n} p_{ij} \leq 3. \tag{2.4}$$

It is satisfied by all vertices of $c(n)$, and therefore by all vectors $p \in c(n)$. If u^ε is a vertex of $c(n)$, then equality holds in (2.4) if and only if $\sum_{i=1}^{n} \varepsilon_i = 2$, or $\sum_{i=1}^{n} \varepsilon_i = 3$. Hence, the convex hull of $\{u^\varepsilon \mid 2 \leq \sum \varepsilon_i \leq 3\}$ is a face of $c(n)$. It is easy to see that, for $n \geq 5$, this face is not a subset of any of the facets of the form (2.3).

2.3. The correlation conjecture

We can generalize inequalities (2.3), and (2.4) in the following way. Call a quadruple (k, a, b, c) an n-adequate quadruple if the following conditions hold:

(a) $2 \leq k \leq n$, $a, b, c \in \mathbf{Z}$.

(b) For all $1 \leq j \leq k$: $a + b\binom{j}{2} \leq c$.

(c) For at least one index $1 \leq j \leq k$ there is equality in (b).

If (k, a, b, c) is an n-adequate quadruple and $p \in c(n)$ we have

$$a \sum_{i \in \alpha} (\sigma p)_i + b \sum_{\substack{ij \in \alpha \\ i<j}} (\sigma p)_{ij} \leq c \tag{2.5}$$

for all $\alpha \subseteq \{1, 2, \ldots, n\}$ such that $|\alpha| \leq k$, and all $\sigma \in I(n)$. Thus I propose:

Conjecture. *Let $p \in \mathcal{R}(n, K_n)$. Then $p \in c(n)$ if and only if inequality* (2.5) *holds for all n-adequate quadruples (k, a, b, c), all $\alpha \subseteq \{1, 2, \ldots, n\}$, $|\alpha| \leq k$, and all $\sigma \in I(n)$.*

We shall see in the next chapter that the conjecture entails that NP = co-NP, so it is quite probably false. In any case, let $l(n)$ denote the polytope generated by inequalities (2.5). (It is a polytope since it is a bounded polyhedron.)

Clearly, $c(n) \subseteq l(n)$, and the group $G(n)$ (involutions and permutations) acts as a symmetry group of $l(n)$. Moreover, for all $\varepsilon \in \{0, 1\}^n$, u^ε *is a vertex of* $l(n)$. This is easy to see: first the origin is a vertex of $l(n)$, since all elements of $l(n)$ are non-negative. But $u^\varepsilon = \sigma(\varepsilon)0$ and the operations of $I(n)$ transform a vertex to a vertex. From (2.5) it follows, in particular, that each $p \in l(n)$ satisfies $0 \leq p_{ij} \leq \min(p_i, p_j)$ and $p_i + p_j - p_{ij} \leq 1$ for all $1 \leq i < j \leq n$.

Lemma 2.3. *The following statements are equivalent*:

(1) $l(n) = c(n)$ *for* $n = 2, 3, \ldots$,

(2) $l(n)$ *is an integral polytope,*

(3) *if* $p \in l(n)$ *is a vertex, then the restriction of p to* $\mathcal{R}(n-1, K_{n-1})$ *is a vertex of* $l(n-1)$ *(for* $n \geq 3$*),*

where by "restriction to $\mathcal{R}(n-1, K_{n-1})$*", I mean the vector* $\bar{p}$ *obtained from p by dropping the coordinates* p_n, p_{in}, $1 \leq i \leq n-1$.

Proof. (1) $\Leftrightarrow$ (2), $c(n)$ is an integral polytope. If $l(n)$ is an integral polytope then the coordinates of its vertices should be in $\{0, 1\}$, since $0 \leq p_{ij} \leq \min\{p_i, p_j\} \leq 1$ for all $p \in l(n)$. But, if $p \in l(n)$ is a vertex, then $p_{ij} = p_i p_j$, for otherwise $p_i = p_j = 1$, $p_{ij} = 0$ entails $p_i + p_j - p_{ij} = 2 > 1$.

(1) $\Leftrightarrow$ (3) If $l(n) = c(n)$ then (3) is valid, since the vertices of $c(n)$ have that property $\bar{u}^\varepsilon = u^\delta$ for $\delta = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-1}) \in \{0, 1\}^{n-1}$. Conversely, suppose that (3) is valid, then since $l(2) = c(2)$ we can proceed by induction. Take a vertex $p \in l(n)$, $n \geq 3$, then $\bar{p} \in l(n-1)$ is a vertex, by the induction hypothesis, that is, $\bar{p} = u^\delta$ for $\delta = (\delta_1, \delta_2, \ldots, \delta_{n-1}) \in \{0, 1\}^{n-1}$. Thus we should have $p_{in} = \delta_i p_n$. To see that assume $\delta_i = 0$, then $0 \leq p_{in} \leq \delta_i = 0 = p_n \delta_i$. If $\delta_i = 1$ we have $p_n + 1 - p_{in} \leq 1$, or $p_n \leq p_{in}$. Since the reversed inequality holds in $l(n)$ we have $p_{in} = p_n = p_n \delta_i$. Now, if $0 < p_n < 1$ we can represent p as a convex combination $p = p_n u^\varepsilon + (1 - p_n) u^{\varepsilon'}$, where $\varepsilon = (\delta_1, \ldots, \delta_{n-1}, 1)$ and $\varepsilon' = (\delta_1, \ldots, \delta_{n-1}, 0)$. But since p is a vertex of $l(n)$, we obtain a contradiction. Hence, $p_n = 0$ or 1, and the claim follows by induction. □

Statement (2) of Lemma 2.3 may assist in proving (or disproving) the conjecture by using well known characterizations of integral polytopes.

2.4. Generalizations, the simplex $C(n)$

Correlation polytopes can easily be generalized. Let S denote in this section (and only in this section) an arbitrary family of non-empty subsets of $\{1, 2, \ldots n\}$, let $\mathcal{R}(S)$ denote the real linear space of all functions $f: S \to \mathcal{R}$ so that $\dim \mathcal{R}(S) = |S|$. For $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0, 1\}^n$ let U^ε denote the following vector in $\mathcal{R}(S)$:

$$U^\varepsilon(\alpha) = \prod_{i \in \alpha} \varepsilon_i, \quad \alpha \in S, \tag{2.6}$$

and let $C(S)$ stand for the closed convex hull of $\{U^\varepsilon;\ \varepsilon \in \{0, 1\}^n\}$. By repeating the argument of Theorem 1.1 we can prove that $p \in C(S)$ if and only if there exists a probability space (X, Σ, μ), events $A_1, \ldots, A_n \in \Sigma$, such that $p(\alpha) = \mu(\bigcap_{i \in \alpha} A_i)$ for $\alpha \in S$.

In the following I will be concerned with the simple case where S is the family of all non-empty subsets of $\{1, 2, \ldots, n\}$. Denote for this case $C(S) = C(n)$; it is a $(2^n - 1)$-dimensional polytope with 2^n vertices, in other words a simplex.

In order to derive the 2^n facet inequalities for $C(n)$, first consider its symmetries. For a fixed $1 \leqslant i \leqslant n$ let ζ_i denote the affine transformation, defined for $x \in \mathcal{R}^{(2^n-1)}$, in the following way: $(\zeta_i x)(\{i\}) = 1 - x(\{i\})$, $(\zeta_i x)(\{j\}) = x(\{j\})$ for $j \neq i$, and for $\alpha \subseteq \{1, 2, \ldots, n\} | \alpha | \geqslant 2$, put $(\zeta_i x)(\alpha) = x(\alpha \backslash \{i\}) - x(\alpha)$ in case $i \in \alpha$, and $(\zeta_i x)(\alpha) = x(\alpha)$ in case $i \notin \alpha$.

The transformation ζ_i leaves $C(n)$ invariant, since its effect is to transform the event A_i to its complement: $\tilde{A}_i = X \backslash A_i$. Let $J(n)$ be the group generated by ζ_i, $i = 1, 2, \ldots, n$; then $J(n) \simeq Z_2^{(n)}$.

Theorem 2.4. *Let $p \in \mathcal{R}^{(2^n-1)}$. Then the following conditions are equivalent:*

(i) $p \in C(n)$.

(ii) $(\zeta p)(\{1, 2, \ldots, n\}) \geqslant 0$ *for all* $\zeta \in J(n)$.

(iii) $\sum_{k=1}^{n} (-1)^{k+1} \sum_{|\alpha|=k} (\zeta p)(\alpha) \leqslant 1$ *for all* $\zeta \in J(n)$.

Proof. (i)$\Rightarrow$(ii): Suppose $p \in C(n)$, then there exists a probability space (X, Σ, μ), events $A_1, \ldots, A_n \in \Sigma$ such that $p(\alpha) = \mu(\bigcap_{i \in \alpha} A_i)$ for all $\emptyset \neq \alpha \subseteq \{1, 2, \ldots, n\}$. For $B \in \Sigma$ let $B^1 = B$, $B^0 = \tilde{B} = X \backslash B$. If $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0, 1\}^n$, denote $\zeta(\varepsilon) = \prod_{i=1}^{n} \zeta_i^{\varepsilon_i} \in J(n)$, and let $\not{\varepsilon} \in \{0, 1\}^n$ stand for the vector $\not{\varepsilon}_i = 1 - \varepsilon_i$. Then

$$(\zeta(\not{\varepsilon})p)(\{1, 2, \ldots, n\}) = \mu(A_2^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap \cdots \cap A_n^{\varepsilon_n}) \geqslant 0.$$

(ii)$\Rightarrow$(i): For $\varepsilon \in \{0, 1\}^n$ put $\lambda(\varepsilon) = (\zeta(\not{\varepsilon})p)(\{1, 2, \ldots, n\}) \geqslant 0$. We shall prove by induction on k, $0 \leqslant k < n$, that if $\alpha \subseteq \{1, 2, \ldots, n\}$, $|\alpha| = n - k$, then $(\zeta p)(\alpha) = \sum_{\varepsilon \in \{0,1\}^n} \lambda(\varepsilon)(\zeta U^\varepsilon)(\alpha)$ for all $\zeta \in J(n)$.

Indeed, for $k=0$ we have

$$(\zeta(\varkappa)p)(\{1,2,\ldots,n\}) = \lambda(\varepsilon) = \sum_{\delta\in\{0,1\}^n} \lambda(\delta)(\zeta(\varkappa)U^\delta)(\{1,2,\ldots,n\}),$$

since $(\zeta(\varkappa)U^\delta)(\{1,2,\ldots,n\}) = 0$ if $\delta \neq \varepsilon$ and 1 for $\delta = \varepsilon$.

Suppose the claim is valid for $1 \leq k < n-1$, let $\alpha \subseteq \{1,2,\ldots,n\}$ be such that $|\alpha| = n-k-1$. Let $1 \leq j \leq n$ be an integer such that $j \notin \alpha$, then by induction hypothesis

$$(\zeta p)(\alpha \cup \{j\}) = \sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon)(\zeta U^\varepsilon)(\alpha \cup \{j\}) \quad \text{for all } \zeta \in J(n).$$

Take $\zeta = \zeta_j$, then $p(\alpha \cup \{j\}) + (\zeta_j p)(\alpha \cup \{j\}) = p(\alpha)$ by definition, also $U^\varepsilon(\alpha \cup \{j\}) + (\zeta_j U^\varepsilon)(\alpha \cup \{j\}) = U^\varepsilon(\alpha)$, hence $p(\alpha) = \sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon) U^\varepsilon(\alpha)$. If $\zeta \in J(n)$ is arbitrary we can repeat the argument for ζp instead of p, and hence the claim follows.

Finally, $\sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon) = 1$. To see that note that by the above proof we have for $\alpha = \{1\}$:

$$p(\{1\}) = \sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon)\varepsilon_1 \quad \text{and} \quad (\zeta_1 p)(\{1\}) = 1 - p(\{1\}) = \sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon)(1-\varepsilon_1).$$

Hence,

$$\sum_{\varepsilon\in\{0,1\}^n} \lambda(\varepsilon) = p(\{1\}) + (\zeta_1 p)(\{1\}) = 1.$$

(ii) $\Leftrightarrow$ (iii): We have

$$\begin{aligned} 1 - \sum_{k=1}^{n} (-1)^{k+1} \sum_{|\alpha|=k} (\zeta(\varkappa)p)(\alpha) &= (\zeta(\varkappa)p)(\{1,2,\ldots,n\}) \\ &= \mu(A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap \cdots \cap A_n^{\varepsilon_n}) \\ &= 1 - \mu(A_1^{1-\varepsilon_1} \cup A_2^{1-\varepsilon_2} \cup \cdots \cup A_n^{1-\varepsilon_n}). \end{aligned}$$ □

3. Complexity

3.1. Introduction

Correlation polytopes exhibit a great deal of symmetry, a fact which may facilitate facet determination. Thus, if we manage to derive one inequality for $c(n)$, we automatically obtain exponentially many by the application of the group operations. Correlation polytopes, on the other hand, are very complex. This chapter is devoted to the study of their complexity.

In Section 3.2 I shall prove that deciding membership in $c(n, S)$ is NP-complete for some particular $S \subseteq K_n$. This is established by a transformation from a particular SATISFIABILITY problem. By a further transformation from GRAPH THREE COLORABILITY, I shall demonstrate in Section 3.3 that deciding membership in $c(n)$ is NP-complete.

In Section 3.4 I shall prove, using the same SATISFIABILITY problem, that if CORRELATION FACET $\in$ NP then NP = co-NP. All this indicates that, unless NP = co-NP, we shall not be able to derive all inequalities of $c(n)$. In the concluding Section 3.5 I shall comment on this situation and on some further open problems.

It is easy to see that deciding membership in $c(n, S)$ is an NP-problem. Let p be a rationally valued vector in $\mathcal{R}(n, S)$ for some set of pairs S. By Caratheodory's theorem, $p \in c(n, S)$ if and only if it is a convex combination of $\dim \mathcal{R}(n, S)+1 = n+|S|+1$ vertices of $c(n, S)$.

At the guessing stage our non-deterministic Turing machine (NDTM) produces $r = n+|S|+1$ vectors $u^{\varepsilon^1}, u^{\varepsilon^2}, \ldots, u^{\varepsilon^r}$, where $\varepsilon^1, \ldots, \varepsilon^r \in \{0, 1\}^n$. Then the machine turns to solve the following instance of LINEAR PROGRAMMING:

Are there $\lambda_l \geqslant 0$, $l = 1, 2, \ldots, r$ such that

$$\sum_{l=1}^{r} \lambda_l = 1,$$

$$\sum_{l=1}^{r} \lambda_l \varepsilon_i^l = p_i, \quad 1 \leqslant i \leqslant n, \ n \text{ equations},$$

$$\sum_{l=1}^{r} \lambda_l \varepsilon_i^l \varepsilon_j^l = p_{ij}, \quad \{i, j\} \in S, \ |S| \text{ equations}.$$

Since LINEAR PROGRAMMING $\in P$ (see [23, 24]), the computation stage terminates after a number of steps which is less than fixed polynomial in the number of code bits for r and p. By Caratheodory's theorem, $p \in c(n, S)$ if and only if the machine stops on YES after some such guess, hence deciding membership in $c(n, S)$ is an NP-problem. (In fact we do not have to use a polynomial time machine for linear programming, for our Turing machine can guess the λ_ls as well.)

3.2. *Deciding membership in $c(n, S)$ is NP-complete*

Let $1 \leqslant k < n$ and let $S_{k,n}$ be the set of all pairs $\{i, j\} 1 \leqslant i < j \leqslant n$ *except* for $\{1, n\}$, $\{2, n\}, \ldots, \{k, n\}$. I shall show in this section that deciding membership in $c(n, S_{k,n})$ is NP-complete. For this I shall use a transformation from the problem ONE IN THREE 3-SATISFIABILITY (see [25, 26]). Let $k \geqslant 4$ and m be integers, $3 \leqslant m \leqslant \binom{k}{3}$. In the present discussion *a proposition over k of length m*, will mean a set of triples

$$\Psi = \{\{a_1^1, a_2^1, a_3^1\}, \{a_1^2, a_2^2, a_3^2\}, \ldots, \{a_1^m, a_2^m, a_3^m\}\}$$

such that for $i = 1, 2, \ldots, m$ and $j = 1, 2, 3$, a_j^i is a natural number $1 \leqslant a_1^i < a_2^i < a_3^i \leqslant k$, for $i = 1, 2, \ldots, m$, and such that for each $1 \leqslant b \leqslant k$ there exists $1 \leqslant i \leqslant m$, for which $b \in \{a_1^i, a_2^i, a_3^i\}$. In such a case we shall say that b *occurs* in $\{a_1^i, a_2^i, a_3^i\}$. Let $l(b)$ denote the number of distinct triples in which b occurs. A truth assignment for Ψ is any function $t: \{1, 2, \ldots, k\} \to \{0, 1\}$; a truth assignment is called a *solution* for Ψ if $t(a_1^i)+t(a_2^i)+t(a_3^i) = 1$ for all $1 \leqslant i \leqslant m$. The decision problem from which a transformation will be defined is:

Instance: A proposition Ψ over k.
Question: Is there a solution for Ψ?

Let Ψ be a fixed proposition over k of length m,put $n=k+m+1$; let $S=S_{k,n}$, that is S is the set of all pairs $\{\{i,j\}; i\neq j,\ 1\leqslant i<j\leqslant n\}$ except for $\{1,n\}$, $\{2,n\},\ldots,\{k,n\}$. Let $J\geqslant 0$ be a natural number, and consider the following vector $p^J\in\mathscr{R}(n,S_{k,n})$:

Domain	Value
$1\leqslant a\leqslant k$	$p^J_a=3^{-l(a)}$
$1\leqslant i\leqslant m$	$p^J_{k+i}=3^{-l(a^i_1)}+3^{-l(a^i_2)}+3^{-l(a^i_3)}$
n	$p^J_n=J\cdot 3^{-m}$
$1\leqslant a<b\leqslant k$	$p^J_{ab}=\begin{cases}0 & \text{there is } 1\leqslant i\leqslant m \text{ such that } \{a,b\}\subseteq\{a^i_1,a^i_2,a^i_3\}\\ 3^{-l(a)-l(b)} & \text{otherwise}\end{cases}$

In the following if $a=b$, $p^J_{ab}=p^J_a$:

$1\leqslant a\leqslant k, 1\leqslant i\leqslant m$	$p^J_{a,k+i}=p^J_{a,a^i_1}+p^J_{a,a^i_2}+p^J_{a,a^i_3}$
$1\leqslant i<j\leqslant m$	$p^J_{k+i,k+j}=p^J_{a^i_1,a^j_1}+p^J_{a^i_2,a^j_1}+p^J_{a^i_3,a^j_1}+p^J_{a^i_1,a^j_2}+p^J_{a^i_2,a^j_2}+p^J_{a^i_3,a^j_2}+p^J_{a^i_1,a^j_3}+p^J_{a^i_2,a^j_3}+p^J_{a^i_3,a^j_3}$
$1\leqslant i\leqslant m, n$	$p^J_{k+i,n}=J\cdot 3^{-m}$

Note that for $J\neq J'$, p^J and $p^{J'}$ are identical except for the coordinates n and $\{k+i,n\}$ for $1\leqslant i\leqslant m$.

Lemma 3.1. (i) *If there exists $J>0$ such that $p^J\in c(n,S_{k,n})$, then Ψ has a solution*
(ii) *Let J be the number of distinct solutions of Ψ ($J=0$ in case Ψ has no solution); then $p^J\in c(n,S_{k,n})$.*

Proof. (i) Suppose that $p^J\in c(n,S_{k,n})$ for $J>0$; then by Theorem 1.1 there exists a probability space (X,Σ,μ), and events $A_1,\ldots,A_k,B_1,\ldots,B_m,C\in\Sigma$ such that

$$p^J_a=\mu(A_a),\quad 1\leqslant a\leqslant k,\qquad p^J_{k+1}=\mu(B_i),\quad 1\leqslant i\leqslant m,$$

$$p^J_n=\mu(C)=J\cdot 3^{-m}>0,\qquad p^J_{a,b}=\mu(A_a\cap A_b),\quad 1\leqslant a<b\leqslant k,$$

$$p^J_{a,k+i}=\mu(A_a\cap B_i),\qquad p^J_{k+i,k+j}=\mu(B_i\cap B_j)\quad\text{and}\quad p^J_{k+i,n}=\mu(B_i\cap C).$$

In the following equality (inclusion) relation between sets will refer to equality (inclusion) up to sets of μ-measure zero.

Claim: $B_i=A_{a^i_1}\cup A_{a^i_2}\cup A_{a^i_3}$. Indeed, by definition, $p^J_{a,k+i}=p^J_{a,a^i_1}+p^J_{a,a^i_2}+p^J_{a,a^i_3}$. If $a=a^i_1$, then, since $p^J_{a^i_1a^i_2}=p^J_{a^i_1a^i_3}=0$, we have: $p^J_{a^i_1,k+i}=p^J_{a^i_1a^i_1}=p^J_{a^i_1}$ (by the convention introduced in the above table), hence $A_{a^i_1}\subseteq B_i$. Similarly, $A_{a^i_2}\subseteq B_i$, $A_{a^i_3}\subseteq B_i$, hence

$A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i} \subseteq B_i$. Also, $\mu(B_i) = p^J_{k+i} = p^J_{a_1^i} + p^J_{a_2^i} + p^J_{a_3^i}$, and $A_{a_1^i}, A_{a_2^i}, A_{a_3^i}$ are pairwise disjoint, hence the claim follows. Now $p^J_{k+i,n} = p^J_n$, that is $\mu(B_i \cap C) = \mu(C)$, therefore $C \subseteq B_i$ for all i and hence

$$C \subseteq \bigcap_{i=1}^{m} B_i = \bigcap_{i=1}^{m} (A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i}).$$

But $\mu(C) > 0$, therefore $C \neq \emptyset$; *let* $x \in C$ and define a truth assignment $t: \{1, 2, \ldots, k\} \to \{0, 1\}$ by

$$t(a) = 1 \quad \text{iff} \quad x \in A_a, \quad 1 \leqslant a \leqslant k.$$

For every $1 \leqslant i \leqslant m$ we have $x \in A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i}$, hence $t(a_1^i) + t(a_2^i) + t(a_3^i) \geqslant 1$, but $A_{a_1^i}, A_{a_2^i}, A_{a_3^i}$ are pairwise disjoint (up to a set of μ-measure zero, which makes no difference here since we can eliminate it anyway), therefore we must have $t(a_1^i) + t(a_2^i) + t(i_3^i) = 1$.

(ii) Let J be the number of distinct solutions for Ψ. Let $X = \{1, 2, 3\}^m$, and let Σ be the power set of X. Let μ be the uniform probability measure on X so that $\mu(\{x\}) = 3^{-m}$ for all $x \in X$. For $1 \leqslant a \leqslant k$, let $A_a = D(a, 1) \times D(a, 2) \times \cdots \times D(a, m)$ where $D(a, i) \subseteq \{1, 2, 3\}$ is defined as follows:

$$D(a, i) = \begin{cases} \{1\}, & a = a_1^i, \\ \{2\}, & a = a_2^i, \\ \{3\}, & a = a_3^i, \\ \{1, 2, 3\}, & \text{otherwise.} \end{cases}$$

Then $\mu(A_a) = 3^{-l(a)} = p^J_a$. Let $1 \leqslant a < b \leqslant k$, then $\mu(A_a \cap A_b) = 0$ in case there is i, such that $a, b \in \{a_1^i, a_2^i, a_3^i\}$, otherwise $\mu(A_a \cap A_b) = \mu(A_a)\mu(A_b)$; put $B_i = A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i}$ and $C = \bigcap_{i=1}^m (A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i})$, then clearly $\mu(B_i) = p^J_{k+1}$, $\mu(B_i \cap A_a) = p^J_{a,k+i}$, $p^J_{k+i,k+j} = \mu(B_i \cap B_j)$, $\mu(B_i \cap C) = \mu(C)$. To complete the proof we have to demonstrate that $\mu(C) = J3^{-m}$ (where J is the number of solutions of Ψ). Let t be a solution for Ψ; define an element $x(t) = (x_1(t), \ldots, x_m(t)) \in X$ as follows: $x_i(t) = 1$ if $t(a_1^i) = 1$, $x_i(t) = 2$ if $t(a_2^i) = 1$ and $x_i(t) = 3$ in case $t(a_3^i) = 1$. Since $t(a_1^i) + t(a_2^i) + t(a_3^i) = 1$ for all $i = 1, 2, \ldots, m$ the point $x(t)$ is well defined. Also, if $t \neq t'$ then $x(t) \neq x(t')$ (remember we assumed that for all $1 \leqslant b \leqslant k$ there is $1 \leqslant i \leqslant m$ such that $b \in \{a_1^i, a_2^i, a_3^i\}$). Let $1 \leqslant a \leqslant k$, then $x(t) \in A_a$ if and only if $t(a) = 1$. To see that put $A_a = D(a, 1) \times \cdots \times D(a, m)$ as above, and let $a = a^{i_1}_{j_1} = a^{i_2}_{j_2} = \cdots = a^{i_r}_{j_r}$, $r = l(a)$ be all the distinct occurrences of a in Ψ, then $t(a) = 1$ if and only if $x_{i_l}(t) = j_l$ for $l = 1, 2, \ldots, r$ if and only if $D(a, i_l) = \{j_l\}$ for $l = 1, 2, \ldots, r$ if and only if $x(t) \in A_a$. But if t is a solution $t(a_1^i) + t(a_2^i) + t(a_3^i) = 1$, hence $x(t) \in A_{a_1^i} \cup A_{a_2^i} \cup A_{a_3^i}$ for all $i = 1, 2, \ldots, m$, hence $x(t) \in C$ and therefore $|C| \geqslant J$. Also, if $x \in C$ we can define a truth assignment $t: \{1, 2, \ldots, k\} \to \{0, 1\}$ by $t(a) = 1$ if and only if $x \in A_a$, and as before, it is easy to see that t is a solution. Hence, $|C| = J$ and $\mu(C) = J3^{-m}$. □

Let p^0, p^1 denote p^J for $J = 0$, $J = 1$; then we can now easily prove:

Theorem 3.2. (i) *We always have* $p^0 \in c(n, S_{k,n})$.

(ii) Ψ *has a solution if and only if* $p^1 \in c(n, S_{k,n})$.

Proof. (i) Choose (X, Σ, μ) as in Lemma 3.1 part (ii), except take $C = \emptyset$ (instead of $C = \bigcap_{i=1}^{m} B_i$), and the claim follows.

(ii) Suppose that $p^1 \in c(n, S_{k,n})$; then Ψ has a solution by Lemma 3.1 part (i). Conversely, suppose Ψ has a solution, let J be the number of solutions of Ψ then $p^J \in c(n, S_{k,n})$ by Lemma 3.1 part (ii). But $p^0 \in c(n, S_{k,n})$, and $p^1 = (1/J)p^J + (1-1/J)p^0$ hence, since $c(n, S_{k,n})$ is convex, the claim follows. □

From Theorem 3.2 it is clear that deciding membership in $c(n, S)$, $S \subseteq K_n$, is NP-complete. For it we had a polynomial time, deterministic Turing machine to decide membership in $c(n, S)$ we would be able to decide ONE IN THREE 3-SATISFIABILITY in polynomial time as well. (It is easy to see that constructing p^1, given a proposition Ψ, can actually be performed in $O(m(m+k)^2) = O(k^9)$ steps.) It should be noted that by a similar, though somewhat more cumbersome technique, a transformation from 3-SATISFIABILITY can be defined.

3.3. Correlation is NP-complete

In this section I shall show that deciding membership in $c(n)$ is NP-complete. To establish that, I shall use the same transformation as in the case of the previous section, only this time the proposition Ψ will be such that p^1_{an} can be uniquely defined also for $1 \leq a \leq k$.

Let $G = (V, E)$ be a simple graph, where V is the set of vertices, and E the set of edges. As is well known (see [23]), deciding whether G is 3-colourable is an NP-complete problem. Given a simple graph G we shall construct a proposition Ψ. The number of variables is $k = 3|V| + 3|E|$, and the number of triples is $m = |V| + 3|E|$. For each vertex $v \in V$ we shall have three variables v_1, v_2, v_3, where v_i is interpreted as: "the colour of the edge v is i" $(i = 1, 2, 3)$. For each edge $e \in E$ we shall have three variables e_1, e_2, e_3 where e_i stands for: "the colour i is missing from the edge e" $(i = 1, 2, 3)$. The proposition Ψ consists of the triples of the form $\{v_1, v_2, v_3\}$ for all $v \in V$, and triples of the form $\{v_i, v'_i, e_i\}$ for $i = 1, 2, 3$ and all $e = \{v, v'\} \in E$. It is easy to see that G is 3-colourable if and only if Ψ has a solution, that is, if and only if there is a truth function which satisfies one, and only one variable in each triple of Ψ.

Now let $n = k + m + 1 = 4|V| + 6|E| + 1$. Define p^J as in the previous section. In this way the value of p^J is given for all pairs in $S_{k,n}$. Note that if G is 3-colourable, then any permutation of the colours $i = 1, 2, 3$ is a 3-colouring of G as well. Hence, the proportion of the solutions of Ψ (i.e., 3-colourings of G), in which a variable (e_i or v_i) is true, is $\frac{1}{3}$. Thus we can put $p^J_{an} = \frac{1}{3}p_n = J \cdot 3^{-(m+1)}$ for $1 \leq a \leq k$. By this p^J is defined for all pairs, $p^J \in \mathcal{R}(n, K_n)$, and we can proceed as before and prove that G is three colourable (or equivalently, Ψ has a solution) if and only if $p^1 \in c(n)$. Consequently, deciding membership in $c(n)$ is NP-complete.

Having established that, it is easy to see that the conjecture of Section 2.3 entails that NP = co-NP. For let $p \in \mathcal{R}(n, K_n)$ then, if the conjecture is true, we can have a polynomial time, non-deterministic Turing machine, which decides whether $p \notin c(n)$. At the guessing stage the machine produces a quadruple (k, a, b, c), a set $\alpha \subseteq \{1, 2, \ldots, n\}$, $|\alpha| \leq k$, and an involution $\sigma \in I(n)$. Then the machine checks whether (k, a, b, c) is n-adequate, and if so, whether $a \sum_{i \in \alpha} (\sigma p)_i + b \sum_{ij \in \alpha, i<j} (\sigma p)_{ij} > c$. All this clearly takes only polynomial time. The only subtle point here is the size of the coefficients a, b, c, which should not be of exponential complexity. Since, however, $c(n)$ is a zero-one polytope, this is guaranteed by Karp and Papadimitriou [22, Lemma 1], which asserts that the complexity of the coefficients of the facet inequalities of polytopes such as $c(n)$ is bounded by a fixed polynomial in the dimension. Hence, if the conjecture is valid, then deciding membership in the *complement* of $c(n)$ is in NP. Since deciding membership in $c(n)$ is NP-complete, we conclude that the conjecture of Section 2.3 entails that NP = co-NP.

3.4. *Optimization and* CORRELATION FACET

Consider the following discrete optimization problem:

SPIN GLASS

Instance: An integral vector $J \in \mathcal{R}(k, K_k)$ and an integer M.

Question: Is

$$\max_{\varepsilon \in \{0,1\}^k} \left[\sum_{i=1}^{k} J_i \varepsilon_i + \sum_{1 \leq i < j \leq k} J_{ij} \varepsilon_i \varepsilon_j \right] \geq M ?$$

This decision problem is equivalent to the question: Is

$$\max_{p \in c(k)} \left[\sum_{i=1}^{k} J_i p_i + \sum_{1 \leq i < j \leq k} J_{ij} p_{ij} \right] \geq M ?$$

Thus if we were able to derive the facet inequalities for $c(k)$ we could have used techniques of linear programming to decide SPIN GLASS. As expected, however, this problem is NP-complete as well.

Lemma 3.3. SPIN GLASS *is NP-complete.*

Proof. We shall use ONE IN THREE 3-SATISFIABILITY, as in Section 3.2. Let Ψ be a proposition with k variables and m triples. For $1 \leq a \leq k$, $l(a)$ denotes the number of triples $\{a_1^i, a_2^i, a_3^i\}$ of which a is an element. For $1 \leq a < b \leq k$ let $l(a, b) = 1$ in case there exist $1 \leq i \leq m$ such that $a, b \in \{a_1^i, a_2^i, a_3^i\}$, and $l(a, b) = 0$ otherwise. We shall prove that Ψ has a solution if and only if there exists $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_k) \in \{0, 1\}^k$ such that

$$\sum_{a=1}^{k} l(a) \varepsilon_a - 3m \sum_{1 \leq a < b \leq k} l(a, b) \varepsilon_a \varepsilon_b \geq m. \tag{3.1}$$

Suppose that Ψ has a solution $t:\{1,2,\ldots,k\}\to\{0,1\}$. Put $\varepsilon_a = t(a)$ for $1\leq a\leq n$. Since $t(a_1^i)+t(a_2^i)+t(a_3^i)=1$ for $1\leq i\leq m$, we have $l(a,b)\varepsilon_a\varepsilon_b=0$ for all $1\leq a<b\leq k$. Also,

$$\sum_{a=1}^{k} l(a)\varepsilon_a = \sum_{i=1}^{m}[t(a_1^i)+t(a_2^i)+t(a_3^i)] = m.$$

Therefore, equality holds in (3.1).

Conversely, suppose that $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_k)\in\{0,1\}^k$ satisfies inequality (3.1). Put $t(a)\varepsilon_a$, $1\leq a\leq k$. Then

$$\sum_{a=1}^{k} l(a)\varepsilon_a = \sum_{i=1}^{m}[t(a_1^i)+t(a_2^i)+t(a_3^i)]\leq 3m.$$

Suppose there exists $1\leq a<b\leq k$ such that $l(a,b)\varepsilon_a\varepsilon_b=1$. Then

$$\sum_{a=1}^{k} l(a)\varepsilon_a - 3m \sum_{1\leq a<b\leq k} l(a,b)\varepsilon_a\varepsilon_b \leq 0 < m.$$

Since the reverse inequality holds, by assumption, we conclude that $\varepsilon_a\varepsilon_b=0$ whenever there exist $1\leq i\leq m$ such that $a, b, \in\{a_1^i, a_2^i, a_3^i\}$. Therefore, $t(a_1^i)+t(a_2^i)+t(a_3^i)\leq 1$ for all $1\leq i\leq m$. But since ε satisfies (3.1) we have $t(a_1^i)+t(a_2^i)+t(a_3^i)=1$ for all $1\leq i\leq m$, so that t is a solution. Since ONE IN THREE 3-SATISFIABILITY is an NP-complete problem it follows that SPIN GLASS is NP-complete as well. □

A variant of this lemma has been previously proved by Barahona [11]. His paper deals with a spin glass system with ±1 values rather than 0, 1. The components of the system are located on a cubic lattice. The minimum energy problem is shown to be NP-complete even if $J_i=0$, $1\leq i\leq n$, and for each i: $J_{ij}=0$ for all j, except the lattice neighbours. Other variants of the problem are also demonstrated to be NP-complete.

Consider now the problem CORRELATION FACET: given an integral vector $J\in\mathscr{R}(n, K_n)$, and an integer M, does the inequality $\sum_{i=1}^{n} J_i p_i + \sum_{i\leq i<j\leq n} J_{ij}p_{ij}\leq M$ represent a facet of $c(n)$?

Theorem 1 in Karp and Papadimitiriou [22] provides the connection between NP-complete discrete optimization problems and facet determination in the corresponding linear programs. Since SPIN GLASS is NP-complete by Lemma 3.3, a direct application of this theorem gives:

Corollary 3.4. *If* CORRELATION FACET ∈ NP, *then* NP = co-NP.

From this conclusion we can see once again that the Conjecture of Section 2.3 entails NP = co-NP. In fact, Corollary 3.4 entails that if NP ≠ co-NP, then CORRELATION FACET is not even in NP. Papadimitiriou and Yannakakis [27] introduced the class D_p of all languages $\mathscr{L}$ which have the form $\mathscr{L}=\mathscr{L}_1\cap\mathscr{L}_2$, where $\mathscr{L}_1\in$ NP and $\mathscr{L}_2\in$ co-NP. Clearly, CORRELATION FACET $\in D_p$. One can probably prove also that it is D_p-complete, though I was not able to establish that.

3.5. Conclusion

Many NP-complete problems have the structure of massive linear programming: optimization on the travelling salesperson polytope, the clique polytope, and a host of others (see [27] for details). CORRELATION has some advantages. Firstly, the interior points of $c(n)$, not just the vertices, have direct interpretation in terms of probability assignments. In fact, all facet inequalities for $c(n)$ should follow from "Venn diagrams", that is, the possible relations among n events in a probability space.

Secondly, the symmetries, which enable us to simplify facet determination. Let $a^{(1)}, \ldots, a^{(r)}$ be vectors in $\mathcal{R}(n, S_n)$, and let $b_1, \ldots, b_r$ be integers. We shall say that $a^{(1)}, \ldots, a^{(r)}, b_1, \ldots, b_r$ is *a generating set of facets* for $c(n)$, if all facet inequalities of $c(n)$ have the form $\sum_{i=1}^{n} a_i^{(l)}(\eta p)_i + \sum_{1 \leq i < j \leq n} a_{ij}^{(l)}(\eta p)_{ij} \leq b_l$, for some group element $\eta \in G(n)$ and some $a^{(l)}, b_l$ in the generating set. Since $|G(n)| = n!2^n$, it is possible that a polynomial size generating set can be found (that is, a set such that r is bounded by a fixed polynomial in n). Moreover, if a non-deterministic, polynomial time program can be found, which identifies the elements of the generating set, then NP = co-NP. This follows from Corollary 3.4 in conjunction with the Theorem of [22]. All in all, I believe that, apart from their intrinsic interest, correlation polytopes may assist us in clarifying the relations between NP and co-NP.

Acknowledgments

I would like to thank Amos Nevo for helpful discussions, B. Weiss for calling my attention to some earlier work on the subject, and the anonymous referee for suggesting the proof in Section 3.3. This research is supported in part by the Edelstein Center for the philosophy of science at the Hebrew University.

References

[1] J.S. Bell, "On the Einstein-Podolsky-Rosen Paradox", *Physics* 1 (1964) 195–200.
[2] E.P. Wigner, "On hidden variables and quantum mechanical probabilities," *American Journal of Physics* 38 (1970) 1005–1009.
[3] J.F. Clauser, M.A. Horne, A. Shimony and R.A. Holt, "Proposed experiment to test local hidden variable theories," *Physical Review Letters* 23 (1969) 880–884.
[4] J.F. Clauser and M.A. Horne, "Experimental consequences of objective local theories," *Physical Review D* 10 (1974) 526–535.
[5] A. Fine, "Hidden variables, joint probability and Bell inequalities," *Physical Review Letters* 48 (1982) 291–295.
[6] A. Garg and N.D. Mermin, "Farkas lemma and the nature of reality, statistical implications of quantum correlations," *Foundations of Physics* 14 (1984) 1–39.
[7] I. Pitowsky, "The range of quantum probability," *Journal of Mathematical Physics* 27 (1986) 1556–1565.
[8] I. Pitowsky, *Quantum Probability, Quantum Logic, Lecture Notes in Physics No. 321* (Springer, Berlin, 1989).
[9] S. Kirkpatrick and D. Sherrington," Infinite-ranged models of spin glasses," *Physical Review B* 17 (1978) 4384–4403.

[10] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in: *Proceedings at the National Academy of Sciences USA* 79 (1982) 2554–2558.
[11] F. Barahona, "On the computational complexity of ising spin glass models," *Journal of Physics A* 15 (1982) 3241–3253.
[12] G. Boole, *The Laws of Thought* (Dover, New York (original edition 1854)).
[13] T. Hailperin, "Best possible inequalities for the probability of a logical function of events," *American Mathematical Monthly* 72 (1962) 343–359.
[14] T. Hailperin, *Boole's Logic and Probability* (North-Holland, Amsterdam, 1986, 2nd ed.).
[15] C.E. Bonferroni, "Il calcolo delle assicurazioni su grouppi di teste," *Studi in Onore del Professor S.O. Carboni* (Roma, 1936).
[16] K.L. Chung, "On the probability of the occurrence of at least m events among n arbitrary events," *Annals of Mathematical Statistics* 12 (1941) 328–338.
[17] M. Fréchet, *Les Probabilités Associées a un Système D'événements Compatible et Dépendants* (Hermann, Paris, *Vol I* 1940, *Vol II* 1943).
[18] S. Kounias and J. Marin, "Best linear Bonferroni bounds," *SIAM Journal of Applied Mathematics* 30 (1976) 307–323.
[19] D. Hunter, "An upper bound for the probability of a union," *Journal of Applied Probability* 13 (1976) 597–603.
[20] M. Cerasoli, "On the probability of a Boolean polynomial of events," *Discrete Mathematics* 44 (1983) 221–227.
[21] W. Maurer, "Bivalent trees and forests or upper bounds for the probability of a union revisited," *Discrete Applied Mathematics* 6 (1983) 157–171.
[22] R.M. Karp and C.H. Papadimitriou, "On linear characterization of combinatorial optimization problem," in: *Proceedings of the 21 Symposium on the Foundations of Computer Science* (1980) pp. 1–9.
[23] L.G. Khacian, "A polynomial algorithm for linear programming," *Dolkady Academie Nauk SSSR* 244(5) (1979) 1093–1096.
[24] C.H. Papadimitriou and D. Steiglitz, *Combinatorial Optimization, Algorithms and Complexity* (Prentice-Hall, Englewood Cliffs, NJ, 1982).
[25] T.J. Schafer, "The complexity of satisfiability problems," in: *Proceedings of the 10th Annual Symposium on Theory of Computing* (ACM, New York, 1978) pp. 216–226.
[26] M.R. Garey and D.S. Johnson, *Computers and Intractability, A Guide to the Theory of NP completeness* (Freeman, New York, 1979).
[27] C.H. Papadimitriou and M. Yannakakis, "The complexity of facets (and some facets of complexity)," *Journal of Computing System Science* 28 (1984) 244–259.