ELSEVIER

Journal of Computational and Applied Mathematics 55 (1994) 217–247

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Applications of cut polyhedra — II

Michel Deza*, Monique Laurent

LIENS — Ecole Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05, France

Received 2 September 1992; revised 15 January 1993

Abstract

This is the continuation of Part I (this issue). In this second part, we present the following applications of cut polyhedra: the max-cut problem, the Boole problem and the multicommodity flow problems in combinatorial optimization, lattice holes in geometry of numbers, density matrices of many-fermions systems in quantum mechanics, as well as some other applications, in probability theory, statistical data analysis and design theory.

As we shall frequently use results, definitions and notation from Part I, the sections in this second part are numbered consecutively.

Keywords: Cut; Polyhedron; ℓ_1-metric; Hypermetric; Delaunay polytope; Probability; Boole problem; Combinatorial optimization; Max-cut problem; Multicommodity flow; Quantum mechanics; Design

We refer to the first part [31] for a general introduction, including also the topics treated here.

5. Applications in combinatorial optimization

5.1. The maximum-cut problem

Given a graph $G = (V_n, E)$ and nonnegative weights w_e, $e \in E$, assigned to its edges, the *max-cut problem* consists of finding a cut $\delta(S)$ whose weight $\sum_{e \in \delta(S)} w_e$ is as large as possible. The max-cut problem is a notorious NP-hard problem [48]. If we replace "as large" by "as small", then we obtain the min-cut problem which can be solved using network flow techniques [46]. Several classes of graphs are known for which the max-cut problem can be solved in polynomial time. This is the case, for instance, for planar graphs [53], for graphs not contractible to K_5 [6], for weakly bipartite graphs, i.e., the graphs G for which the polytope $\{x \in \mathbb{R}_+^E : x(C) \leqslant |C| - 1$ for all odd cycles

* Corresponding author.

SSDI 0377-0427(93)E0180-T

C of $G\}$ has all its vertices integral [52]. In fact, the class of weakly bipartite graphs includes the graphs not contractible to K_5 ([45] or [80]).

For definitions of the terms used in this section, see, e.g., [51, 86].

The max-cut problem can be reformulated as a linear programming problem over the cut polytope, namely, as

$$\begin{aligned} &\max && w^{\mathrm{T}}x \\ &\text{subject to} && x \in \mathrm{CUTP}(G). \end{aligned}$$

This is the polyhedral approach, classical in combinatorial optimization, which leads to the study of the facets of $\mathrm{CUTP}(G)$. This approach has been used in practice for solving large instances of the max-cut problem (see, e.g., [7, 8]). Its success depends, of course, on the degree of knowledge about the facets needed for the problem at hand and of their tractability, i.e., whether they can be separated in polynomial time or, at least, whether a good separation heuristic is available.

For instance, $\mathrm{CUTP}(G) = \mathrm{MET}(G)$, i.e., the inequalities

$$x(F) - x(C - F) \leqslant |F| - 1 \quad \text{for } F \subseteq C \text{ cycle with } |F| \text{ odd}$$

are sufficient for describing $\mathrm{CUTP}(G)$ if and only if G is not contractible to K_5 [9]. Moreover, the above inequalities can be separated in polynomial time, implying that the max-cut problem in graphs not contractible to K_5 is polynomially solvable [9].

The max-cut problem in an arbitrary graph G on n nodes can always be formulated as

$$\begin{aligned} &\max && w^{\mathrm{T}}x \\ &\text{subject to} && x \in \mathrm{CUTP}_n \end{aligned}$$

after setting $w_e = 0$ if e is not an edge of G. This permits to exploit fully the symmetry of the complete graph.

The max-cut problem has many applications in various fields. For instance, the problem of determining ground states of spin glasses with an exterior magnetic field, or the problem of minimizing the number of vias (holes on a printed circuit board) subject to pin assignment and layer preferences, can both be formulated as instances of the max-cut problem; they arise, respectively, in statistical physics and VLSI circuit design. We refer to [7] for a detailed description of these two applications, together with a computational treatment. In fact, the spin glass problem was already mentioned in [72] as an optimization problem over the boolean quadric polytope. Other applications can be found in [12].

Another application is to unconstrained quadratic 0–1 programming, which consists of solving

$$\begin{aligned} &\max && \sum_{1 \leqslant i \leqslant j \leqslant n} c_{ij}x_ix_j \\ &\text{subject to} && x \in \{0,1\}^n, \end{aligned}$$

where $c_{ij} \in \mathbb{R}$. If we set $p_{ij} = x_ix_j$ for $1 \leqslant i \leqslant j \leqslant n$, this problem can be equivalently formulated as a linear programming problem over the boolean quadric polytope

$$\begin{aligned} &\max && c^{\mathrm{T}}p \\ &\text{subject to} && p \in \mathrm{BQP}_n. \end{aligned}$$

Just as the points of the boolean quadric polytope and of the cut polytope are in one-to-one correspondence (via the covariance map; see [31, Section 2.4]), the max-cut problem and the unconstrained quadratic programming problem are equivalent.

Other approaches, besides the polyhedral approach, have been proposed for attacking the max-cut problem. In particular, an approach based on eigenvalue methods is investigated in [23, 81, 82]. We mention briefly some facts, permitting to connect it with polyhedral aspects.

The Laplacian matrix L of the graph G is the $n \times n$ matrix defined by $L_{ii} = \deg_G(i)$ for $i \in V_n$ and $L_{ij} = -a_{ij}$ for $i \neq j \in V_n$, where $A(a_{ij})_{1 \leqslant i,j \leqslant n}$ is the adjacency matrix of G. Set

$$\varphi(G) = \tfrac{1}{4}n \min\left(\lambda_{\max}(L + \operatorname{diag}(u))\colon u \in \mathbb{R}^n, \sum_{1 \leqslant i \leqslant n} u_i = 0\right),$$

where $\operatorname{diag}(u)$ is the diagonal matrix with diagonal entries $u_1, \ldots, u_n$ and $\lambda_{\max}(L + \operatorname{diag}(u))$ is the largest eigenvalue of the matrix $L + \operatorname{diag}(u)$. Set

$$\psi(G) = \max(\tfrac{1}{2}\operatorname{Trace}(AY)\colon \tfrac{1}{2}J - Y \textit{ is} \text{ positive semidefinite and}$$

$$Y \text{ a symmetric } n \times n \text{ matrix with } Y_{ii} = 0 \text{ for } 1 \leqslant i \leqslant n),$$

where J is the $n \times n$ matrix with all entries equal to 1. Let $\operatorname{mc}(G)$ denote the maximum cardinality of a cut in G. Then,

(i) $\operatorname{mc}(G) \leqslant \varphi(G)$ [23],

(ii) $\operatorname{mc}(G) \leqslant \psi(G)$ [87].

The quantity $\psi(G)$ can be easily reformulated as

$$\psi(G) = \max(\textstyle\sum_{1 \leqslant i \leqslant j \leqslant n} a_{ij}x_{ij}\colon x \text{ satisfies the inequalities (19) for all integers } b_1, \ldots, b_n),$$

$$\sum_{1 \leqslant i \leqslant j \leqslant n} b_i b_j x_{ij} \leqslant \frac{1}{4}\left(\sum_{1 \leqslant i \leqslant n} b_i\right)^2. \tag{19}$$

Inequalities (19) are clearly valid for the cut polytope $\operatorname{CUT}P_n$, but they are never facet defining since they are dominated by the gap inequalities (5) (defined in [31, Section 2.2]); however, inequalities (19) can be separated in polynomial time while the separation problem for the gap inequalities is probably hard.

In fact, using general duality theory, it is shown that $\varphi(G) = \psi(G)$ holds [82]. Recently, it has been shown in [49] that the quantity $\psi(G)$ provides a good approximation for the max-cut problem, namely,

$$\frac{\psi(G)}{\operatorname{mc}(G)} \leqslant 1.138.$$

5.2. *Multicommodity flows*

An instance of the multicommodity flow problem consists of two graphs: the *supply graph* $G = (V_n, E)$ together with a *capacity* function $c\colon E \to \mathbb{R}_+$, and the *demand graph* $H = (T, U)$ together with a *demand* function $r\colon U \to \mathbb{R}_+$, where $T \subseteq V_n$ is the set of nodes spanned by U. Given

a pair of nodes (s,t), $\mathscr{P}_{st}$ denotes the set of st-paths in G and we set $\mathscr{P} = \bigcup_{(s,t)\in U} \mathscr{P}_{st}$. A *multiflow* is a function $f: \mathscr{P} \to \mathbb{R}_+$. The instance (G,H,c,r) is said to be *feasible* if there exists a feasible multiflow, i.e., a multiflow $f: \mathscr{P} \to \mathbb{R}_+$ satisfying the following capacity and demand requirements:

$$\sum_{P\in\mathscr{P}, e\in P} f_P \leqslant c_e \quad \text{for } e\in E, \qquad \sum_{P\in\mathscr{P}_{st}} f_P \geqslant r_{st} \quad \text{for } (s,t)\in U.$$

Using the Farkas lemma, it can be checked that the following holds.

Proposition 5.1. *The problem (G,H,c,r) is feasible if and only if $c^{\mathrm{T}}y - r^{\mathrm{T}}z \geqslant 0$ for all $(y,z)\in C(G,H)$, where $C(G,H)$ is the cone defined by*

$$C(G,H) = \left\{(y,z)\in \mathbb{R}_+^E \times \mathbb{R}_+^U : \sum_{e\in P} y_e - z_{st} \geqslant 0 \quad \text{for } P\in\mathscr{P}_{st} \text{ and } (s,t)\in U\right\}.$$

The cone $C(G,H)$ is studied in detail in [61] and, in particular, the fractionality of its extreme rays.

Without loss of generality, we can suppose that G is the complete graph K_n; then, r is extended to K_n by setting $r_e = 0$ for the edges $e\notin U$ and $U = \{e: r_e > 0\}$ is called the *support* of r and we simply say that the pair (c,r) is feasible. An alternative characterization for feasible multiflows is given by the following so-called Japanese theorem (from [56, 75], restated in [68, 69]).

Theorem 5.2. *The pair (c,r) is feasible if and only if*

$$(c-r)^{\mathrm{T}}d \geqslant 0 \quad \text{for all } d\in \mathrm{MET}_n. \tag{20}$$

Therefore, the metric cone MET_n is the dual cone to the cone of feasible multiflows.

An obvious way for testing feasibility of the pair (c,r) is to solve the linear program $\min((c-r)^{\mathrm{T}}d: d\in \mathrm{MET}_n)$ which has $\binom{n}{2}$ variables and $3\binom{n}{3}$ constraints (the triangle inequalities (1) in [31]). An alternative way is to check the condition (20) for all extreme rays d of MET_n. This approach leads to the study of extreme rays of the metric cone MET_n (see references on it in [31, Section 2.4]).

There are other variants of the Japanese theorem, in particular, in the more general setting of binary matroids (see [89]). In particular, the metric cone $\mathrm{MET}(G)$ (defined in [31, relation (11)]) arises naturally when studying multicommodity flows. It is shown in [89] that all extreme rays of $\mathrm{MET}(G)$ are 0-, 1-valued (i.e., $\mathrm{MET}(G) = \mathrm{CUT}(G)$) if and only if G is not contractible to K_5. The graphs for which all extreme rays of $\mathrm{MET}(G)$ are 0-, 1-, 2-valued are characterized in [88]. The graphs for which all the vertices of the metric polytope $\mathrm{METP}(G)$ (defined in [31, relation (12)]) are $\frac{1}{3}$-integral are studied in [66] (x is said to be $\frac{1}{3}$-integral if $3x$ is integral).

Since the cut cone CUT_n is contained in the metric cone MET_n, a necessary condition for the existence of a feasible multiflow is the following *cut condition*:

$$\sum_{e\in\delta(S)} (c_e - r_e) \geqslant 0 \quad \text{for all } S\subseteq V_n. \tag{21}$$

The well-known Ford–Fulkerson theorem [46] states that the cut condition is, in fact, also sufficient for feasibility in the case of single commodity flows, i.e., when $|U| = 1$. We give below some other results of this type. An *integral multiflow* is a multiflow f with integral values.

Theorem 5.3. *Assume that the support of the demand function r is K_4, C_5, or the union of two stars (i.e., all edges are covered by two nodes). Then, the pair (c,r) is feasible if and only if the cut condition (21) holds [76]. Moreover, if c, r are integral, $(c-r)^T\delta(S)$ is even for all cuts and (21) holds, then there exists an integral multiflow (see [69] and references therein).*

Theorem 5.4 (Karzanov [60, 62]). *If the support graph of the demand function r is a subgraph of K_5 (including K_5), c, r are integral and $(c-r)^T\delta(S)$ is even for all cuts, then there exists an integral multiflow if and only if (20) holds or, equivalently, if and only if (21) holds and $(c-r)^T d \geqslant 0$ holds for all 0-extensions of the path metrics of $K_{2,3}$.*

There is a close connection between these results and L_1-embeddability, as noted in [4]. Given a semimetric d on V_n, an *extremal graph* [68, 69] for d is a minimal graph $K = (V_0, W)$ such that, for each $x, y \in V_n$, there exists $(s,t) \in W$ satisfying $d_{sx} + d_{xy} + d_{yt} = d_{st}$, and V_0 is the set of nodes covered by W. The extremal graph is unique if $d_{ij} > 0$ for all $i, j \in V_n$. The notion of extremal graph is a key notion for testing feasibility of multiflows.

Proposition 5.5 (Lomonossov [68, 69]). *The pair (c,r) is feasible if and only if $(c-r)^T d \geqslant 0$ holds for all $d \in \mathrm{MET}_n$ having an extremal graph $K = (V_0, W)$ such that W is a subset of the support of the demand function r.*

Theorem 5.6 (Karzanov [59]). *If $d \in \mathrm{MET}_n$ has an extremal graph which is K_4, C_5, or a union of two stars, then $d \in \mathrm{CUT}_n$. Moreover, if d satisfies the parity condition [31, (16)], then d is a non-negative integer sum of cuts, i.e., d is h-embeddable.*

Note that the latter two results imply the first part of Theorem 5.3.

We conclude with some additional related results.

Given a supply graph G, a capacity function c and a demand graph H, the *maximum multiflow problem* consists of finding a multiflow f not exceeding the capacity constraints whose value $\sum_{P \in \mathscr{P}} f_P$ is as large as possible. By linear programming duality, this problem is equivalent to the linear programming problem

$$\min(c^T y:\ y \in \mathbb{R}_+^E, y(P) \geqslant 1 \text{ for all } P \in \mathscr{P}).$$

This leads to the study of the polytope $P(G,H) = \{y \in \mathbb{R}_+^E : y(P) \geqslant 1 \text{ for all } P \in \mathscr{P}\}$. The fractionality of the vertices of $P(G,H)$ is studied in detail in [61]; in particular, the demand graphs H for which all vertices of $P(G,H)$ are $\frac{1}{4}$-integral for an arbitrary demand graph G with $V(H) \subseteq V(G)$ are characterized.

5.3. The Boole problem

Let $(\Omega, \mathscr{A}, \mu)$ be a probability space and let $A_1, \ldots, A_n$ be n events of $\mathscr{A}$. A classical question, which goes back to Boole [11], is the following:

Suppose we are given the values $p_i = \mu(A_i)$ for $1 \leqslant i \leqslant n$, what is the best estimation of $\mu(A_1 \cup \cdots \cup A_n)$?

It is easy to see that the answer is

$$\max(p_1, \dots, p_n) \leqslant \mu(A_1 \cup \cdots \cup A_n) \leqslant \min\left(1, \sum_{1 \leqslant i \leqslant n} p_i\right).$$

More generally, let $\mathscr{I}$ be a collection of subsets of $\{1, \dots, n\}$.

Suppose we are given the values of the joint probabilities $p_I = \mu(\bigcap_{i \in I} A_i)$ for all $I \in \mathscr{I}$. What is the best estimation of $\mu(A_1 \cup \cdots \cup A_n)$ in terms of the p_I's ?

In fact, the answer to this problem is given by the facet defining inequalities for the polytope $\mathrm{BQP}_n^{\mathscr{I}}$ (defined in [31, Section 2.4]). Namely,

$$\mu(A_1 \cup \cdots \cup A_n) \geqslant \max(w^{\mathrm{T}} p: w^{\mathrm{T}} z \leqslant 1 \text{ is facet defining for } \mathrm{BQP}_n^{\mathscr{I}})$$

(see Proposition 5.8 and relation (26)). In particular, when $\mathscr{I}$ consists of all pairs and singletons, then the lower bound for $\mu(A_1 \cup \cdots \cup A_n)$ is in terms of the facets of the boolean quadric polytope BQP_n.

Estimations for $\mu(A_1 \cup \cdots \cup A_n)$ via linear programming

First, we observe that [31, Theorem 3.2] remains valid for the polytope $\mathrm{BQP}_n^{\mathscr{I}}$ for an arbitrary nonempty set family $\mathscr{I}$.

Theorem 5.7. *Let $\mathscr{I}$ be a nonempty collection of subsets of $\{1, \dots, n\}$ and let $p = (p_I)_{I \in \mathscr{I}} \in \mathbb{R}^{\mathscr{I}}$. The following assertions are equivalent.*

(i) $p \in \mathrm{BQ}_n^{\mathscr{I}}$ *(resp. $p \in \mathrm{BQP}_n^{\mathscr{I}}$).*

(ii) *There exist a nonnegative measure space (resp. a probability space) $(\Omega, \mathscr{A}, \mu)$ and $A_1, \dots, A_n \in \mathscr{A}$ such that $p_I = \mu(\bigcap_{i \in I} A_i)$ for all $I \in \mathscr{I}$.*

Proof. It is identical to that of [31, Theorem 3.2]. □

Given $p \in \mathrm{BQ}_n^{\mathscr{I}}$, consider the following two linear programming problems:

$$\begin{aligned} &\text{minimize} && \sum_{\emptyset \neq S \subseteq \{1,\dots,n\}} \lambda_S && \\ &\text{subject to} && \sum_{\emptyset \neq S \subseteq \{1,\dots,n\}} \lambda_S \pi^{\mathscr{I}}(S) = p, && (22)\\ &&& \lambda_S \geqslant 0 \qquad \text{for } \emptyset \neq S \subseteq \{1, \dots, n\}, && \end{aligned}$$

$$\begin{aligned} &\text{maximize} && \sum_{\emptyset \neq S \subseteq \{1,\dots,n\}} \lambda_S && \\ &\text{subject to} && \sum_{\emptyset \neq S \subseteq \{1,\dots,n\}} \lambda_S \pi^{\mathscr{I}}(S) = p, && (23)\\ &&& \lambda_S \geqslant 0 \qquad \text{for } \emptyset \neq S \subseteq \{1, \dots, n\}. && \end{aligned}$$

Let $z_{\min}$ (resp. $z_{\max}$) denote the optimum value of the program (22) (resp. (23)).

So, the program (22) (resp. (23)) is evaluating the minimum value (resp. the maximum value) of $\sum_S \lambda_S$ for a decomposition $p = \sum_S \lambda_S \pi^{\mathscr{I}}(S)$, $\lambda_S \geqslant 0$, of $p \in \mathrm{BQ}_n^{\mathscr{I}}$. In particular, in the case $\mathscr{I} = \mathscr{I}_{\leqslant 2}$, if

we set $d = \varphi_{c_0}^{-1}(p)$, then $d \in \mathrm{CUT}_{n+1}$ and $z_{\min}$ coincides with the minimum size $s(d)$ (defined in [31, Section 2.5]). This approach, in the case of $\mathscr{I}_{\leqslant 2}$, is considered in [63, 79].

Proposition 5.8. $z_{\min} \leqslant \mu(A_1 \cup \cdots \cup A_n) \leqslant z_{\max}$.

Proof. For $S \subseteq \{1, \ldots, n\}$, set $A^S = \bigcap_{i \in S} A_i \cap \bigcap_{i \notin S} (\Omega - A_i)$. Then, $\bigcap_{i \in I} A_i = \bigcup_{I \subseteq S \subseteq \{1,\ldots,n\}} A^S$, $\Omega = \bigcup_S A^S$ and $A_1 \cup \cdots \cup A_n = \bigcup_{S \neq \emptyset} A^S$. We have $p_I = \mu(\bigcap_{i \in I} A_i)$ for each $I \in \mathscr{I}$. Therefore, $p = \sum_{S \neq \emptyset} \mu(A^S) \pi^{\mathscr{I}}(S)$ holds, with $\mu(A^S) \geqslant 0$. Hence $(\mu(A^S)\colon \emptyset \neq S \subseteq \{1, \ldots, n\})$ is a feasible solution to the program (22), or (23), with objective value $\mu(A_1 \cup \cdots \cup A_n)$. This proves the result. □

The dual programs to (22) and (23) are the following programs (24) and (25), respectively:

$$\begin{aligned} &\text{maximize} && w^{\mathrm{T}} p \\ &\text{subject to} && w^{\mathrm{T}} \pi^{\mathscr{I}}(S) \leqslant 1 \text{ for } \emptyset \neq S \subseteq \{1, \ldots, n\}, \end{aligned} \tag{24}$$

$$\begin{aligned} &\text{minimize} && w^{\mathrm{T}} p \\ &\text{subject to} && w^{\mathrm{T}} \pi^{\mathscr{I}}(S) \geqslant 1 \text{ for } \emptyset \neq S \subseteq \{1, \ldots, n\}. \end{aligned} \tag{25}$$

By linear programming duality, we have

$$z_{\min} = \max(w^{\mathrm{T}} p\colon w^{\mathrm{T}} z \leqslant 1 \text{ is a valid inequality for } \mathrm{BQP}_n^{\mathscr{I}}) \tag{26}$$

and it is easily verified that, in relation (26), it is sufficient to consider facet defining inequalities. Similarly,

$$z_{\max} = \min(w^{\mathrm{T}} p\colon w^{\mathrm{T}} z \geqslant 1 \text{ is facet defining for the polytope } \mathrm{Conv}(\{\pi^{\mathscr{I}}(S)\colon \emptyset \neq S \subseteq V_n\})).$$

(The latter polytope is distinct from $\mathrm{BQP}_n^{\mathscr{I}}$ since it does not contain the origin.)

Therefore, by (26), every valid inequality for $\mathrm{BQP}_n^{\mathscr{I}}$ yields a lower bound for $\mu(A_1 \cup \cdots \cup A_n)$ in terms of the joint probabilities $p_I = \mu(\bigcap_{i \in I} A_i)$ for $I \in \mathscr{I}$. Examples of such lower bounds are exposed below (after Proposition 5.9).

The case when the collection $\mathscr{I}$ of index sets is $\mathscr{I}_{\leqslant m}$ is considered in [13]. The following estimations for $\mu(A_1 \cup \cdots \cup A_n)$ are given there:

$$y_{\min} \leqslant \mu(A_1 \cup \cdots \cup A_n) \leqslant y_{\max}, \tag{27}$$

where $y_{\min}$ is the optimum value of the linear program (28) below and $y_{\max}$ is the optimum value of (29) below, setting

$$S_k = \sum_{1 \leqslant i_1 \leqslant i_2 < \cdots < i_k \leqslant n} \mu(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \quad \text{for } 1 \leqslant k \leqslant n.$$

$$\begin{aligned} &\text{minimize} && \sum_{1 \leqslant i \leqslant n} v_i \\ &\text{subject to} && \sum_{1 \leqslant i \leqslant n} \binom{i}{k} v_i = S_k \quad \text{for } 1 \leqslant k \leqslant m, \\ & && v_i \geqslant 0 \qquad\qquad \text{for } 1 \leqslant i \leqslant n, \end{aligned} \tag{28}$$

$$\begin{aligned} &\text{maximize} && \sum_{1 \leqslant i \leqslant n} v_i \\ &\text{subject to} && \sum_{1 \leqslant i \leqslant n} \binom{i}{k} v_i = S_k \quad \text{for } 1 \leqslant k \leqslant m, \\ & && v_i \geqslant 0 \qquad\quad \text{for } 1 \leqslant i \leqslant n. \end{aligned} \tag{29}$$

In fact, the programs (22), (23) give sharper bounds than the programs (28), (29), respectively. Namely, we have the following proposition.

Proposition 5.9. *In the case* $\mathscr{I} = \mathscr{I}_{\leqslant m}$ *for some integer* m, $1 \leqslant m \leqslant n$, *we have* $y_{\min} \leqslant z_{\min} \leqslant \mu(A_1 \cup \cdots \cup A_n) \leqslant z_{\max} \leqslant y_{\max}$.

Proof. Indeed, every feasible solution for (22) yields a feasible solution for (28) with the same objective value. Namely, let $(\lambda_S, \emptyset \neq S \subseteq \{1, \ldots, n\})$ be a feasible solution for (22), i.e., $\lambda_S \geqslant 0$ and $p = \sum_S \lambda_S \pi^{\mathscr{I}_{\leqslant m}}(S)$. Set $v_i = \sum_{S:|S|=i} \lambda_S$ for $1 \leqslant i \leqslant n$. Then,

$$\begin{aligned} \sum_{1 \leqslant i \leqslant n} \binom{i}{k} v_i &= \sum_{1 \leqslant i \leqslant n} \binom{i}{k} \sum_{S:|S|=i} \lambda_S \\ &= \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \; \sum_{S: i_1, \ldots, i_k \in S} \lambda_S \\ &= \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} p_{\{i_1, \ldots, i_k\}} \\ &= \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mu(A_{i_1} \cap \cdots \cap A_{i_k}) \\ &= S_k. \end{aligned}$$

Therefore, $(v_1, \ldots, v_n)$ is a feasible solution for (28) with $\sum_{1 \leqslant i \leqslant n} v_i = \sum_S \lambda_S$. This shows that $y_{\min} \leqslant z_{\min}$. The inequality $z_{\max} \leqslant y_{\max}$ follows from the same argument. □

Examples of bounds for $\mu(A_1 \cup \cdots \cup A_n)$

The best lower bound for $\mu(A_1 \cup \cdots \cup A_n)$ is given by $z_{\min}$, defined by relation (26), whose evaluation relies on the knowledge of the facets of the polytope $\mathrm{BQP}_n^{\mathscr{I}}$. In the case $\mathscr{I} = \mathscr{I}_{\leqslant 2}$, the facet structure of the boolean quadric polytope BQP_n has been extensively studied (directly or indirectly, via the covariance map, through the cut polytope). We describe below several examples of valid inequalities for BQP_n, together with the lower bounds they yield for $\mu(A_1 \cup \cdots \cup A_n)$.

First, note that, if $p = \sum_S \lambda_S \pi(S)$ with $\lambda_S \geqslant 0$, then $n \sum_{1 \leqslant i \leqslant n} p_i - 2 \sum_{1 \leqslant i < j \leqslant n} p_{ij} = \sum_S \lambda_S |S|(n + 1 - |S|)$, where $n \leqslant |S|(n + 1 - |S|) \leqslant \lfloor \frac{1}{2}(n+1) \rfloor \lceil \frac{1}{2}(n+1) \rceil$ if $S \neq \emptyset$. Hence, we have:

$$\begin{aligned} \frac{n \sum_{1 < i < n} p_i - 2 \sum_{1 < i < j < n} p_{ij}}{\lfloor \frac{1}{2}(n+1) \rfloor \lceil \frac{1}{2}(n+1) \rceil} &\leqslant \sum_{\emptyset \neq S \subseteq \{1, \ldots, n\}} \lambda_S, \\ \frac{n \sum_{1 < i < n} p_i - 2 \sum_{1 < i < j < n} p_{ij}}{n} &\geqslant \sum_{\emptyset \neq S \subseteq \{1, \ldots, n\}} \lambda_S \end{aligned} \tag{30}$$

and, therefore, from the definition of $z_{\min}$, $z_{\max}$ and from Proposition 5.8,

$$\begin{aligned} &\frac{n\sum_{1<i<n} p_i - 2\sum_{1<i<j<n} p_{ij}}{\lfloor \frac{1}{2}(n+1)\rfloor \lceil \frac{1}{2}(n+1)\rceil} \leqslant \mu(A_1 \cup \cdots \cup A_n), \\ &\frac{n\sum_{1<i<n} p_i - 2\sum_{1<i<j<n} p_{ij}}{n} \geqslant \mu(A_1 \cup \cdots \cup A_n). \end{aligned} \tag{31}$$

Note that the inequalities equivalent to (30) in the context of the cut cone are the bounds on the minimum size of $d \in \mathrm{CUT}_{n+1}$ given in [31, relation (13)].

The inequality

$$2k \sum_{1 \leqslant i \leqslant n} p_i - 2 \sum_{1 \leqslant i < j \leqslant n} p_{ij} \leqslant k(k+1) \tag{32}$$

is valid for the boolean quadric polytope BQP_n for $1 \leqslant k \leqslant n-1$; it is facet defining if $1 \leqslant k \leqslant n-2$ and $n \geqslant 4$. Setting $b_0 = 2k+1-n$ and $b_1 = \cdots = b_n = 1$, inequality (32) corresponds (via the covariance map) to the inequality

$$\sum_{0 \leqslant i < j \leqslant n} b_i b_j x_{ij} \leqslant k(k+1), \tag{33}$$

which is valid for the cut polytope CUTP_{n+1}; (33) is a switching of the hypermetric inequality $\mathrm{Hyp}_{n+1}(2k+1-n, 1, \ldots, 1, -1, \ldots, -1)$ (with $n-k$ coefficients $+1$ and k coefficients -1). (See, e.g., [29].) Therefore, we have the following lower bound for $\mu(A_1 \cup \cdots \cup A_n)$:

$$\frac{2}{k+1} \sum_{1 \leqslant i \leqslant n} p_i - \frac{2}{k(k+1)} \sum_{1 \leqslant i < j \leqslant n} p_{ij} \leqslant \mu(A_1 \cup \cdots \cup A_n) \tag{34}$$

for each k, $1 \leqslant k \leqslant n-1$; it was found independently by several authors, including [16, 22, 47]. Note that (34) coincides with the lower bound of (31) in the case $n = 2k$.

More generally, given integers $b_1, \ldots, b_n$ and $k \geqslant 0$, the inequality

$$\sum_{1 \leqslant i \leqslant n} b_i(2k+1-b_i)p_i - 2 \sum_{1 \leqslant i < j \leqslant n} b_i b_j p_{ij} \leqslant k(k+1) \tag{35}$$

is valid for BQP_n. This yields the bound

$$\frac{1}{k(k+1)}\left(\sum_{1 \leqslant i \leqslant n} p_i b_i(2k+1-b_i) - 2 \sum_{1 \leqslant i < j \leqslant n} b_i b_j p_{ij} \right) \leqslant \mu(A_1 \cup \cdots \cup A_n).$$

The programs (28), (29) provide weaker bounds than the programs (22), (23), but they present the advantage of being easier to handle, especially for small values of m. Exploiting their special structure, the bounds $y_{\min}$ and $y_{\max}$ were explicitly described in [13] in terms of the S_k's (defined in relation (27)), as we recall briefly.

Let M denote the matrix corresponding to the program (22) or (23). Its columns are the n vectors a_i, where $a_i = (\binom{i}{1}, \binom{i}{2}, \ldots, \binom{i}{m})$ for $1 \leqslant i \leqslant n$. Set $b = (S_1, \ldots, S_m)$. The matrix M is full rank, hence a basis B consists of a set of m linearly independent vectors among $a_1, \ldots, a_n$. The basis B is called *dual feasible* if the vector $y = 1_m^{\mathrm{T}} M_B^{-1}$ is feasible for the dual program of (28), i.e. $y^{\mathrm{T}} a_i \leqslant 1$ for

$i \in \{1, \dots, n\} - B$, since equality holds for the indices $i \in B$ (M_B is the submatrix of M whose columns are those vectors a_i belonging to the basis B; 1_m has m coordinates equal to 1). If M is dual feasible, then the inequality $1_B^T M_B^{-1} b \leqslant \mu(A_1 \cup \cdots \cup A_n)$ holds. The dual feasible bases are explicitly described in [13] together with the corresponding bounds for $\mu(A_1 \cup \cdots \cup A_n)$.

For example, for m even, $\{a_1, a_2, \dots, a_m\}$ is a dual feasible basis, yielding the bound

$$\mu(A_1 \cup \cdots \cup A_n) \geqslant S_1 - S_2 + S_3 - S_4 \cdots + (-1)^{m-1} S_m,$$

which was first given in [10]. For $m = 2$, this is the special case $k = 1$ of the bound (34); another choice of basis yields the general bound (34).

In fact, the method from [13] also works for finding estimates of the probabilities $\mu(\{v \geqslant r\})$ and $\mu(\{v = r\})$, where v denotes the random variable counting the number of events that occur among $A_1, \dots, A_n$.

Inequality (35) can alternatively be written as

$$\left(\sum_{1 \leqslant i \leqslant n} b_i p_i - k \right) \left(\sum_{1 \leqslant i \leqslant n} b_i p_i - k - 1 \right) \geqslant 0 \tag{36}$$

with the convention that, when developing the product, the expression $p_i p_j$ is replaced by the variable p_{ij} (setting $p_{ii} = p_i$). This inequality (or special cases of it) was considered in this form by many authors (e.g., [40, 63, 72, 79, 94]). This suggests naturally the following generalization of inequality (36) in the case $\mathcal{I}_{\leqslant m}$, when m is an even integer. Given integers $b_1, \dots, b_n$ and $k_1, \dots, k_m \geqslant 0$, the inequality

$$\prod_{1 \leqslant l \leqslant m} \left(\sum_{1 \leqslant i \leqslant n} b_i p_i - k_l \right) \left(\sum_{1 \leqslant i \leqslant n} b_i p_i - k_k - 1 \right) \geqslant 0 \tag{37}$$

is clearly valid for the polytope $\mathrm{BQP}_n^{\mathcal{I}_{\leqslant 2m}}$. Thus arises the question of determining the parameters $b_1, \dots, b_n, k_1, \dots, k_m$ for which (37) defines a facet of $\mathrm{BQP}_n^{\mathcal{I}_{\leqslant 2m}}$. This problem is, however, already difficult for the case $m = 1$ of the boolean quadric polytope.

6. Hypermetrics and geometry of numbers

6.1. L-polytopes

We recall here some definitions about lattices and L-polytopes. A detailed treatment can be found in [20, 27].

Given $x, y \in \mathbb{R}^k$, we set $d_0(x, y) = (\|x - y\|_2)^2$ (the square of the euclidean distance). Recall that the hypermetric cone HYP_n is defined by the hypermetric inequalities:

$$\sum_{1 \leqslant i < j \leqslant n} b_i b_j x_{ij} \leqslant 0 \quad \text{for } b_1, \dots, b_n \text{ integers with } \sum_{1 \leqslant i \leqslant n} b_i = 1. \tag{38}$$

For $d \in \mathrm{HYP}_n$, $(V_n = \{1, \dots, n\}, d)$ is called a *hypermetric space*. It is convenient to work with the hypermetric cone HYP_{n+1} defined on the $n + 1$ points $0, 1, 2, \dots, n$.

A subset $L \subseteq \mathbb{R}^k$ is a *lattice* if, up to translation, L is a discrete subgroup of $\mathbb{R}^k$. So, the notion of lattice considered in this section is distinct from the notion of lattice (as partially ordered set) used in [31, Section 4.4]. A subset $B = \{v_0, v_1, \ldots, v_m\} \subseteq L$ is *generating* for L if, for each $v \in L$, there exist integers $z_0, z_1, \ldots, z_m$ such that $\sum_{0 \leqslant i \leqslant m} z_i = 1$ and $v = \sum_{0 \leqslant i \leqslant m} z_i v_i$. If, moreover, there is unicity of the integers z_i, then B is an (affine) *basis* of L; in this case, $m = |B| - 1$ is called the *dimension* of L.

Let L be a k-dimensional lattice in $\mathbb{R}^k$. Let $S = S(c, r)$ denote the sphere with center c and radius r. The sphere S is called an *empty sphere* (in Russian literature), or *hole* (in English literature), in L if the following two conditions hold:

- $\|v - c\|_2 \geqslant r$ holds for all $v \in L$,
- $S \cap L$ has affine rank $k + 1$.

Then, the polytope P defined as the convex hull of $S \cap L$ is called an *L-polytope* (or *Delaunay polytope*, or *constellation*); S is its circumscribed sphere and c is its center. The L-polytope P is *generating* if its set of vertices $V(P)$ generates L, and *basic* if $V(P)$ contains an affine basis of L. Actually all known generating L-polytopes are basic.

For $v \in S$, let $v^* = 2c - v$ denote its antipode on S. Every L-polytope P is either *asymmetric*, i.e., $v^* \notin V(P)$ for each vertex $v \in V(P)$, or *centrally symmetric*, i.e., $v^* \in V(P)$ for each $v \in V(P)$.

Two L-polytopes P, P' have the same *type* if they are affinely equivalent, i.e., $P' = T(P)$ for some affine bijective map T.

Examples of L-polytopes include the n-dimensional simplex α_n, hypercube γ_n, cross polytope $\beta_n := \mathrm{Conv}(\pm e_i : 1 \leqslant i \leqslant n)$ (where $e_1, \ldots, e_n$ are the unit vectors in $\mathbb{R}^n$). Both β_n and γ_n are centrally symmetric, α_n is asymmetric. All types of L-polytopes in dimension $k \leqslant 4$ have been classified in [42]:

- for $k = 1$, there is only $\alpha_1 = \beta_1 = \gamma_1$;
- for $k = 2$, they are α_2 and $\beta_2 = \gamma_2$;
- for $k = 3$, they are α_3, β_3, γ_3, the prism (with triangular base) and the pyramid (with square base);
- for $k = 4$, there are 19 polytopes.

Remark that the polytopes α_n, β_n, γ_n are L-polytopes for any n.

The following polytope $P^m_{p,q}$ was studied and named *repartitioning polytope* by Voronoi (see also [5]). Let P be a polytope and let v be a point which does not lie in the affine space spanned by P; the convex hull of P and v is called the *pyramid* with base P and apex v and is denoted by $\mathrm{Pyr}(P)$. We define iteratively $\mathrm{Pyr}_m(P)$ as $\mathrm{Pyr}(\mathrm{Pyr}_{m-1}(P))$, setting $\mathrm{Pyr}_0(P) = P$. Let S_p, S_q be two simplices of respective dimensions p, q and lying in affine spaces which intersect in one point. Then, $P^m_{p,q} := \mathrm{Pyr}_m(\mathrm{Conv}(S_p \cup S_q))$ is called a *repartitioning polytope*; it has dimension $m + p + q$ and $m + p + q + 2$ vertices. In fact, $P^m_{p,q}$ does not denote a concrete polytope, but corresponds to a class of affinely equivalent polytopes of the same type.

A construction of symmetric L-polytopes is given in [25]. Let L be an integral lattice (i.e. $u^{\mathrm{T}} v$ integer for all $u, v \in L$) and set $m = \min(u^{\mathrm{T}} u : u \in L, u \neq 0)$. For $c \in L$, $c \neq 0$, set $P_c = \mathrm{Conv}(\{u \in L : u^{\mathrm{T}} u = m \text{ and } 2u^{\mathrm{T}} c = (\|c\|_2)^2\})$. Then, P_c is a symmetric L-polytope. Moreover, under some condition, the set of diagonals of P_c is a set of equiangular lines. (See Section 6.4.)

Finally, we mention the connection between L-polytopes and Voronoi polytopes. Given $v_0 \in L$, the *Voronoi polytope* $P_{\mathrm{V}}(v_0)$ is the set $\{x \in \mathbb{R}^k : \|x - v_0\|_2 \leqslant \|x - v\|_2 \text{ for all } v \in L\}$. The vertices of $P_{\mathrm{V}}(v_0)$ are exactly the centers of the L-polytopes in L which contain v_0.

6.2. Hypermetrics and L-polytopes

We state here the beautiful connection existing between hypermetrics and L-polytopes.

Theorem 6.1 (Assouad [3]). (i) *Let P be an L-polytope with set of vertices $V(P)$. Then, $(V(P), d_0)$ is a hypermetric space.*

(ii) *Let $d \in \mathrm{HYP}_{n+1}$. Then, there exist a lattice $L_d \subseteq \mathbb{R}^k$ of dimension $k \leqslant n$, an L-polytope P_d in L_d and a map $f_d: \{0, 1, \ldots, n\} \to V(P_d)$, $f_d(i) = v_i$ for $0 \leqslant i \leqslant n$, such that*

- *$\{v_0, v_1, \ldots, v_n\}$ generates L_d,*
- *$d_{ij} = d_0(v_i, v_j) = (\|v_i - v_j\|_2)^2$ for $0 \leqslant i \leqslant j \leqslant n$.*

Moreover, the triple (L_d, P_d, f_d) is unique, up to translation and orthogonal transformation.

Therefore, hypermetrics on $n+1$ points correspond to generating L-polytopes of dimension $k \leqslant n$.

Proof. (i) Let $S(c, r)$ denote the empty sphere circumscribed to P. Let b_v, $v \in V(P)$, be integers with $\sum_{v \in V(P)} b_v = 1$. Then,

$$\begin{aligned}\sum_{u,v \in V(P)} b_u b_v d_0(u,v) &= \sum_{u,v \in V(P)} b_u b_v (\|(u-c)+(c-v)\|_2)^2 \\ &= \sum_{u,v \in V(P)} b_u b_v (2r^2 + 2(u-c)^{\mathrm{T}}(c-v)) \\ &= 2r^2 - 2\left(\left\|\sum_{u \in V(P)} b_u u - c\right\|_2\right)^2 \leqslant 0,\end{aligned}$$

because $\sum_{u \in V(P)} b_u u \in L$.

We now give a sketch of the proof of (ii). One of the basic tools used in the proof is the covariance map φ_{c_0}. Define $p = \varphi_{c_0}(d)$, $p = (p_{ij})_{1 \leqslant i \leqslant j \leqslant n}$. By [31, relation (8)], $d \in \mathrm{HYP}_{n+1}$ if and only if $\sum_{1 \leqslant i,j \leqslant n} b_i b_j p_{ij} - \sum_{1 \leqslant i \leqslant n} b_i p_{ii} \geqslant 0$ for all integers $b_1, \ldots, b_n$. Therefore, if $d \in \mathrm{HYP}_{n+1}$, then the symmetric matrix $(p_{ij})_{1 \leqslant i,j \leqslant n}$ is positive semidefinite and, thus, $p_{ij} = v_i^{\mathrm{T}} v_j$, $1 \leqslant i \leqslant j \leqslant n$, for some vectors $v_1, \ldots, v_n \in \mathbb{R}^k$, where k is the rank of the matrix $(p_{ij})_{1 \leqslant i,j \leqslant n}$, $k \leqslant n$.

Moreover, one can show the existence of $c \in \mathbb{R}^k$ such that $2c^{\mathrm{T}} v_i = (\|v_i\|_2)^2$ for $1 \leqslant i \leqslant n$. Therefore, $v_0 = 0, v_1, \ldots, v_n$ lie on the sphere $S(c, r := \|c\|_2)$. Remains only to show that $\{v_1, \ldots, v_n\}$ generates a lattice L and that the sphere S is empty in L. □

Proposition 6.2 (Deza et al. [27]). *Let P be an L-polytope and let V be a subset of its set of vertices $V(P)$. Let P' be the L-polytope associated with the hypermetric space (V, d_0). Then, $V(P') \subseteq V(P)$ with equality if and only if V is a generating subset of $V(P)$.*

In particular, every face of an L-polytope is an L-polytope.

We summarize in Table 2 the correspondences between some special hypermetrics and their associated L-polytopes. Given $d \in \mathrm{HYP}_{n+1}$, $F(d)$ denotes the smallest face of HYP_{n+1} containing d.

Table 2

Hypermetric d		Associated L-polytope P
	[3]	
d is an ℓ_1-metric		Vertices of a parallepiped
d is a cut	$\Leftrightarrow$	$P = \alpha_1$
	[3]	
$F(d) = \mathrm{HYP}_{n+1}$	$\Leftrightarrow$	$P = \alpha_n$
	[5]	
$F(d)$ is a facet	$\Leftrightarrow$	P is a repartitioning polytope
	[27]	
$F(d)$ is an extreme ray	$\Leftrightarrow$	P is extreme
	[27]	
$F(d) = F(d')$	$\Leftrightarrow$	P, P' are affinely equivalent

The hypermetric cone is defined by an infinite list of inequalities. Thus arises naturally the question of deciding whether it is a polyhedral cone, i.e., whether among the infinite list of inequalities (38) only a finite number is nonredundant. The answer is yes, as stated in the following result.

Theorem 6.3 (Deza et al. [28]). *The hypermetric cone* HYP_n *is polyhedral.*

The proof given in [28] is based on the following two facts:
- the correspondence between the hypermetrics of HYP_{n+1} and the L-polytopes of dimension $k \leqslant n$,
- the fact that, in given dimension, the number of types of L-polytopes is finite [93] (a direct proof is given in [28]).

Let $b^n_{\max}$ denote the largest value of $\max_i |b_i|$ for which inequality (38) defines a facet of HYP_n. Then, $b^n_{\max} < (2^{n-2}(n-1)!)/(n+1)$ is shown in [5].

6.3. Rank of an L-polytope

Let $d \in \mathrm{HYP}_{n+1}$ and let $F(d)$ denote the smallest face of HYP_{n+1} containing d. The dimension of $F(d)$ is called the *rank* of d and is denoted as $r(d)$, or $r(V_{n+1}, d)$. Hence, $r(d) = 1$ if d lies on an extreme ray of HYP_{n+1}, $r(d) = \binom{n+1}{2}$ if d lies in the interior of HYP_{n+1} and $r(d) = \binom{n+1}{2} - 1$ if $F(d)$ is a facet of HYP_{n+1}.

Let P be an L-polytope. The *rank* $r(P)$ of P is defined as the rank of the hypermetric space $(V(P), d_0)$. In fact, the rank of a hypermetric d is an invariant of the associated L-polytope P_d, namely, $r(d) = r(P_d)$.

Proposition 6.4 [27]. *Let P be an L-polytope and let $V \subseteq V(P)$ be a generating subset. Then, $r(V, d_0) = r(V(P), d_0) = r(P)$ holds.*

Proposition 6.5 [27]. *Let P be an L-polytope. Then, $r(P) = 1$ if and only if the only affine bijective transformations T (up to translation and orthogonal transformation) for which $T(P)$ is an L-polytope are the homotheties.*

The *extreme* L-polytopes, i.e., those having rank 1, are of special importance since they correspond to the extreme rays of the hypermetric cone. For $n \leqslant 5$, $\mathrm{HYP}_{n+1} = \mathrm{CUT}_{n+1}$, i.e., the only extreme rays are the cut vectors. Therefore, the only extreme L-polytope of dimension $k \leqslant 5$ is α_1.

Proposition 6.6 [27]. *Let P_i, $i = 1,2$, be an L-polytope in $\mathbb{R}^{k_i}$. Then, $P_1 \times P_2$ is an L-polytope in $\mathbb{R}^{k_1+k_2}$ with rank $r(P_1 \times P_2) = r(P_1) + r(P_2)$.*

For instance, $r(\gamma_k) = kr(\gamma_1) = k$. An important consequence of Proposition 6.6 is that, if P is an extreme L-polytope in a lattice L, then L must be irreducible.

Proposition 6.7 [27]. *Let P be a basic L-polytope of dimension k. Then,*
(i) $\binom{k+2}{2} \leqslant r(P) \leqslant \binom{k+2}{2} - |V(P)|$,
(ii) *for P centrally symmetric,*

$$r(P) \geqslant \binom{k+1}{2} - \tfrac{1}{2}|V(P)| + 1.$$

For instance, for α_k, $r(\alpha_k) = k + 1$, yielding equality in both inequalities of (i); for β_k, $r(\beta_k) = \binom{k+1}{2} - k + 1$ yielding equality in (ii).

6.4. Extreme L-polytopes

A direct application of Proposition 6.7 yields the following bounds for an extreme basic L-polytope of dimension k:

$$|V(P)| \geqslant \tfrac{1}{2}k(k+3), \tag{39}$$

$$|V(P)| \geqslant k(k+1) \quad \text{if } P \text{ is centrally symmetric.} \tag{40}$$

There is a striking analogy between the bounds (39) and (40) and some known upper bounds (see [67]) for the number N_p of points in a spherical two-distance set of dimension k and the number N_l of lines in a set of equiangular lines of dimension k, namely,

$$N_p \leqslant \tfrac{1}{2}k(k+3) \quad \text{and} \quad N_l \leqslant \tfrac{1}{2}k(k+1).$$

Moreover, if $N_l = \frac{1}{2}k(k+1)$, then $k+2 = 4, 5$, or $k+2 = q^2$ for some odd integer $q \geqslant 3$ (see [67]). The first case of equality is for $q = 3$, $k = 7$, $N_l = 28$; it corresponds to the set of 28 equiangular lines defined by the diagonals of the Gosset polytope 3_{21}. The next case of equality is

for $q = 5$, $k = 23$, $N_1 = 276$; it corresponds to the set of 276 equiangular lines defined by the diagonals of the extreme L-polytope P_{23} constructed from the Leech lattice (see below). For $q = 7$, $k = 47$, $N_l = 1128$, it is not known whether such a set of equiangular lines exists.

However, there are examples of extreme L-polytopes realizing equality in the bounds (39) or (40), but not arising from some spherical two-distance set or from some equiangular set of lines; this is the case for the polytopes P^8, P^{16} constructed from the Barnes–Wall lattice (see below). There are also examples of extreme L-polytopes not realizing equality in the bounds (39) or (40).

We have given in [27] several examples of extreme L-polytopes achieving or not equality in the bounds (39) or (40). We refer to [27] for a detailed account and to [20] for details on lattices.

Extreme L-polytopes in root lattices

All the extreme L-polytopes in root lattices are classified. Indeed, by Witt's theorem, the only irreducible root lattices are A_n ($n \geqslant 0$), D_n ($n \geqslant 4$) and E_n ($n = 6, 7, 8$). All types of L-polytopes in a root lattice are given in [90] or [41]. They are the half-cube $h\gamma_n$, the cross polytope β_n, the simplex α_n, the Gosset polytope 3_{21} and the Schläfli polytope 2_{21} (whose 1-skeletons are, respectively, the half-cube graph $\frac{1}{2}H(n, 2)$, the cocktail party graph $K_{n\times 2}$, the complete graph K_{n+1}, the Gosset graph G_{56} and the Schläfli graph G_{27}). Among them, the extreme polytopes are: the segment α_1, the Schläfli polytope 2_{21} and the Gosset polytope 3_{21}, of respective dimensions 1, 6, 7. The polytope 2_{21} is asymmetric with 27 vertices, realizing equality in the bound (39). The polytope 3_{21} is centrally symmetric with 56 vertices, realizing equality in the bound (40). Both are basic. We do not know any other extreme L-polytope of dimension $k \leqslant 7$ besides α_1, 2_{21}, 3_{21}.

Extreme L-polytopes in sections of the Leech lattice Λ_{24}

The Leech lattice Λ_{24} is a lattice of dimension 24. By taking suitable sections of the sphere of minimal vectors of Λ_{24}, two extreme L-polytopes are constructed in [27]:

- P_{23}, centrally symmetric, with 552 vertices, dimension 23, realizing equality in the bound (40),
- P_{22}, asymmetric, with 275 vertices, dimension 22, realizing equality in the bound (39).

Extreme L-polytopes in sections of the Barnes–Wall lattice Λ_{16}

The Barnes–Wall lattice Λ_{16} is a lattice of dimension 16. Several examples of extreme L-polytopes are constructed from Λ_{16} in [27]:

- P, centrally symmetric (constructed from a deep hole of Λ_{16}), with 512 vertices, dimension 16 (equality does not hold in (40)),
- Q, centrally symmetric, with 272 vertices, dimension 16, realizing equality in the bound (40),
- P^8, P^{16}, asymmetric, with 135 vertices, dimension 15, realizing equality in the bound (39),
- Q', asymmetric, with 1080 vertices, dimension 15 (equality does not hold in (39)).

Extreme hypermetric graphs

Let G be a hypermetric graph on n nodes, i.e., whose path metric d_G is hypermetric, and let P_G denote the L-polytope associated with d_G. It is shown in [26] that, if G is an extreme hypermetric graph, i.e., if d_G lies on an extreme ray of the hypermetric cone HYP_n and if $G \neq K_2$, then G is of one of the following two types:

Type I: $P_G = 3_{21}$, implying that $8 \leqslant n \leqslant 56$ and G has diameter 2 or 3.

Type II: $P_G = 2_{21}$, implying that $7 \leqslant n \leqslant 27$ and G has diameter 2.

Moreover, for G of diameter 2, G is extreme of type II if and only if its suspension ∇G is extreme of type I (recall that ∇G is obtained from G by adding a new node adjacent to all nodes of G).

In particular, the number of extreme hypermetric graphs is finite.

More details can be found in [26]; for instance, all regular extreme hypermetric graphs belong to the list from [14] of 187 regular graphs which have smallest eigenvalue -2 and are not line graphs; in particular, all nine maximal graphs of this list are extreme hypermetrics.

7. Applications in quantum mechanics

7.1. Preliminaries on quantum mechanics

The object of (nonrelativistic) quantum mechanics is to study microscopic objects, e.g., molecules, atoms, or any elementary particles. One of the fundamental differences with classical (Newtonian) mechanics is that many physical quantities can take only discrete values at the microscopic level and that the state of microscopic objects is disturbed by observations. Moreover, identical particles, i.e., with the same physical characteristics as mass, size, charge, etc., can be distinguished in classical mechanics (for instance, by following their trajectories) but they are undistinguishable within quantum mechanics. Von Neumann [92] laid the foundations for a rigorous mathematical account of quantum mechanics. We recall below some basic definitions and facts from quantum mechanics needed for our treatment. Useful references containing a detailed account of these facts include [43, 50, 70, 72, 94].

Consider a system of $N \geqslant 2$ identical particles. Each particle is represented by a vector $x = (r, s)$ composed by a space coordinate $r \in \mathbb{R}^3$ and a spin coordinate $s \in \mathbb{Z}_2$; $X = \mathbb{R}^3 \times \mathbb{Z}_2$ denotes the space of the coordinates. The physical state of the system is represented by a normalizable complex valued function ψ defined on X^N, called the *wavefunction*. Using the fact that no physical observation can be made that permits to distinguish the particles, it can be shown that either all wavefunctions are symmetric, or all of them are antisymmetric. In the symmetric case, the particles are called *bosons*, and in the antisymmetric case, they are called *fermions*. We consider here the case of a system of N fermions, i.e., the wavefunctions are antisymmetric.

Let $H(N)$ denote the set of the measurable complex-valued antisymmetric functions defined on X^N; $H(N)$ is a Hilbert space, called the *Fock space*, with inner product

$$\langle \psi_1, \psi_2 \rangle = \int_{x \in X^N} \psi_1^*(x) \psi_2(x) \, dx$$

for $\psi_1, \psi_2 \in H(N)$. Hence, the physical states of a system of fermions are represented by functions $\psi \in H(N)$ with $\langle \psi, \psi \rangle = 1$. In fact, the case of bosons can be treated in a similar way if the antisymmetry condition is replaced by the symmetry condition and the determinants by permanents in the Slater determinants (defined below).

A physical quantity of the system, or *observable*, is represented by a Hermitian operator A of the space $H(N)$ and the expected value of A in the state ψ is given by

$$\langle A \rangle_\psi := \langle \psi, A\psi \rangle = \int \psi^*(x) A\psi(x) \, dx.$$

Among the observables of the system, the simplest ones are those that the system may have (then the expected value of the observable is equal to one), or lack (then the expected value is zero). Such observables are represented by orthogonal projections on subspaces of $H(N)$.

Every observable A being a Hermitian operator admits a spectral decomposition. For simplicity, we assume that A can be decomposed as $A = \sum_{i \geq 1} \lambda_i E_i$, where the λ_i's are the eigenvalues of A and E_i denotes the projection on the eigenspace associated with the eigenvalue λ_i. So, the projection E_i corresponds to the property "The observable A has value λ_i". If the system is in the state ψ, then it has the property associated with E_i if $\langle E_i \rangle_\psi = 1$, i.e., if $A\psi = \lambda_i \psi$, that is, ψ is an eigenvector of A corresponding to the eigenvalue λ_i.

The *standard deviation* of the observable A in the state ψ is given by

$$\Delta_\psi(A) = |\langle A^2 \rangle_\psi - (\langle A \rangle_\psi)^2|^{1/2}.$$

Heisenberg's uncertainty principle states that, if A, B are two observables of the system in the state ψ, then $\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{2}|\langle \psi, (AB - BA)\psi \rangle|$, i.e., A, B cannot be simultaneously measured with precision if they do not commute.

An important observable of the system is its energy, represented by the Hamiltonian operator and denoted by Ω. The average energy of the system in the state ψ is given by $\langle \Omega \rangle_\psi$. A fundamental problem in quantum mechanics is to derive bounds on the average energy of the system without knowing explicitly the state ψ of the system. In fact, as we shall explain below, this problem has some tight connections with the problem of finding the linear description of the boolean quadric polytope.

The *density matrix of order* p of $\psi \in H(N)$ is the complex-valued function $\Gamma_\psi^{(p)}$ defined on $X^p \times X^p$ by:

$$\Gamma_\psi^{(p)}(x'_1 \dots x'_p \mid x_1 \dots x_p) = \binom{N}{p} \int_{y \in X^{N-p}} \psi^*(x'_1, \dots, x'_p, y)\psi(x_1, \dots, x_p, y)\,\mathrm{d}y. \tag{41}$$

Density matrices were introduced in [55] (see also [70]); Dirac [32] already considered density matrices of order $p = 1$. Density matrices have a simpler and more direct physical meaning than the wavefunction itself, in particular, the diagonal elements $\Gamma_\psi^{(p)}(x_1 \dots x_p \mid x_1 \dots x_p)$ which are of special importance. Indeed, $N^{-1}\Gamma_\psi^{(1)}(x_1 \mid x_1)\,\mathrm{d}v_1$ is the probability of finding a particle with spin s_1 within the volume $\mathrm{d}v_1$ around the point r_1, when all other particles have arbitrary positions and spins. Similarly, $\binom{N}{2}^{-1}\Gamma_\psi^{(2)}(x_1 x_2 \mid x_1 x_2)\,\mathrm{d}v_1\,\mathrm{d}v_2$ is the probability of finding a particle with spin s_1 within the volume $\mathrm{d}v_1$ around the point r_1, and another particle with spin s_2 within the volume $\mathrm{d}v_2$ around the point r_2, when all other particles have arbitrary spins and positions, etc.

From the antisymmetry of the wavefunction ψ, $\Gamma_\psi^{(p)}(x_1 \dots x_p \mid x_1 \dots x_p) = 0$ if $x_i = x_j$ for distinct i, j. In other words, particles with parallel spins are kept apart. This phenomenon is a consequence of the Pauli principle.

Density matrices have been widely studied. In particular, they were the central topic of several conferences held at Queen's University, Kingston, Canada, yielding three volumes of proceedings [19, 39, 43].

Every Hermitian operator A of $H(N)$ can be expanded as

$$A = A_0 + \sum_{1 \leq i \leq N} A_i + \frac{1}{2!} \sum_{1 \leq i \neq j \leq N} A_{ij} + \cdots, \tag{42}$$

where the nth term is an $(n-1)$-particle operator. Therefore, the expected value of A in the state ψ can be expressed, in terms of the density matrices, as follows:

$$\langle A\rangle_\psi = A_0 + \int \{A_1 \Gamma_\psi^{(1)}(x_1' \mid x_1)\}_{x_1' = x_1} \mathrm{d}x_1 + \int \{A_{12}\Gamma_\psi^{(2)}(x_1'x_2' \mid x_1x_2)\}_{x_1' = x_1, x_2' = x_2} \mathrm{d}x_1\,\mathrm{d}x_2 + \cdots \tag{43}$$

with the following convention for the notation $\{A_1\Gamma_\psi^{(1)}(x_1' \mid x_1)\}_{x_1' = x_1}$: A_1 operates only on the unprimed coordinate x_1, not on x_1', but after the action of A_1 has been carried out, one sets again $x_1' = x_1$. The same convention applies to the other terms.

By the Hartree–Fock approximation (see [50]), one can assume that the expansion of the Hamiltonian Ω in relation (42) has only terms involving two particles at most, i.e., $\Omega = \Omega_0 + \sum_{1 \leqslant i \leqslant N} \Omega_i + \frac{1}{2}\sum_{i \neq j} \Omega_{ij}$. In other words, one takes only into account pairwise interactions between the particles and the interaction of each particle with an exterior potential. Observe that Ω can then be expressed as $\Omega = \frac{1}{2}\sum_{i \neq j} G_{ij}$, where

$$G_{ij} = \Omega_{ij} + \frac{1}{N-1}(\Omega_i + \Omega_j) + \frac{2}{N(N-1)}\Omega_0.$$

Therefore, from relation (43), the average energy depends only on the second-order density matrices $\Gamma_\psi^{(2)}$. Hence, the question of finding bounds on the average energy reduces to the question of finding the boundary conditions on the second-order density matrices. In fact, the density matrices of first and second order contain already most of the useful information about the physical state of the system accessible to physicists.

Let Φ_k, $k \geqslant 1$, be an orthonormal set (assumed to be discrete for the sake of simplicity) of functions of $H(1)$ such that each function $f \in H(1)$ can be expanded as

$$f = \sum_{k \geqslant 1} \langle \Phi_k, f\rangle \Phi_k. \tag{44}$$

The functions Φ_k are called the *spin-orbitals*. Given a set $K = \{k_1, \ldots, k_N\}$, with $1 \leqslant k_1 < \cdots < k_N$, the *Slater determinant* Φ_K is defined by

$$\Phi_K(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \det(\Phi_{k_1}(x), \ldots, \Phi_{k_N}(x)), \tag{45}$$

where $\Phi_k(x)$ denotes the vector $(\Phi_k(x_1), \ldots, \Phi_k(x_N))$. Equivalently,

$$\Phi_K(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathrm{Sym}(N)} \mathrm{sign}(\sigma)\Phi_{k_{\sigma(1)}}(x_1) \ldots \Phi_{k_{\sigma(N)}}(x_N). \tag{46}$$

Then, each wavefunction $\psi \in H(N)$ can be expanded as

$$\psi(x_1, \ldots, x_N) = \sum_{K = \{k_1, \ldots, k_N\}, 1 \leqslant k_1 < \cdots < k_N} C_K \Phi_K, \tag{47}$$

where

$$C_K = \langle \Phi_K, \psi \rangle = \sqrt{N!} \int \psi(x_1, \ldots, x_N) \Phi_{k_1}^*(x_1) \ldots \Phi_{k_N}^*(x_N) \, dx_1 \cdots dx_N \tag{48}$$

with $\sum_K |C_K|^2 = \langle \psi, \psi \rangle = 1$.

A usual assumption consists in selecting a finite set of n spin-orbitals $\{\Phi_1, \ldots, \Phi_n\}$ so that the finite sum

$$\sum_{K \subseteq \{1,\ldots,n\}, |K| = N} C_K \Phi_K \tag{49}$$

constitutes a good approximation of the wavefunction ψ. From now on, we assume that ψ is, in fact, equal to the finite sum in (49). It can be shown [70] that the second-order density matrix $\Gamma_\psi^{(2)}$ can also be expanded in terms of the Slater determinants. Namely, if ψ is given by (49) where the coefficients C_K are given by (48), then

$$\Gamma_\psi^{(2)}(x_1' x_2' \,|\, x_1 x_2) = \sum_{1 \leqslant i < j \leqslant n, 1 \leqslant h < k \leqslant n} \gamma_\psi(ij \,|\, hk) \Phi_{\{i,j\}}^*(x_1', x_2') \Phi_{\{h,k\}}(x_1, x_2). \tag{50}$$

The coefficients $\gamma_\psi(ij \,|\, hk)$ are given by

$$\gamma_\psi(ij \,|\, hk) = \sum C_I^* C_K \varepsilon_{i,j,I-\{i,j\}}^{I} \varepsilon_{h,k,K-\{h,k\}}^{K}, \tag{51}$$

where the sum is over all subsets $I, K \subseteq \{1, \ldots, n\}$ of cardinality N such that $i, j \in I$, $h, k \in K$ and $I - \{i,j\} = K - \{h,k\}$, and we set $\varepsilon_{i_1 \ldots i_p}^{j_1 \ldots j_p} = \operatorname{sign}(\sigma)$ if there is a permutation σ mapping i_1 on $j_1, \ldots, i_p$ on j_p and $\varepsilon_{i_1 \ldots i_p}^{j_1 \ldots j_p} = 0$ otherwise. In particular, the diagonal terms are given by

$$\gamma_\psi(ij \,|\, ij) = \sum_{i,j \in K \subseteq \{1,\ldots,n\}, |K| = N} |C_K|^2. \tag{52}$$

They have the following physical meaning: $\binom{N}{2}^{-1} \gamma_\psi(ij \,|\, ij)$ is the probability of finding a particle in the ith spin-orbital and another one in the jth spin-orbital while all other particles occupy arbitrary spin-orbitals.

7.2. *The N-representability problem*

Given a complex-valued function Γ defined on $X^2 \times X^2$, Γ is said to be *N-representable* if there exists a wavefunction $\psi \in H(N)$ such that $\Gamma = \Gamma_\psi^{(2)}$. The *pure state representability problem* consists of finding the conditions that Γ must satisfy in order to be N-representable. This problem can be relaxed to the *ensemble representability problem* as follows. Instead of asking whether Γ is the second-order density matrix of a single wavefunction ψ, one may ask whether there exists a convex combination $\sum w_\psi \psi$ ($w_\psi \geqslant 0$, $\sum w_\psi = 1$) of wavefunctions such that $\Gamma = \sum w_\psi \Gamma_\psi^{(2)}$ is the convex combination of their second-order density matrices.

Note that, from the point of view of finding a state of minimum energy, it is equivalent to consider pure states or ensembles (mixtures) of states. Indeed, both $\langle \Omega \rangle_\psi$ and $\sum w_\psi \langle \Omega \rangle_\psi$ have the same minimum (equal to the minimum eigenvalue of the Hamiltonian Ω and attained at a corresponding eigenvector).

Let $\mathscr{P}_N^{(2)}$ denote the convex set consisting of the convex combinations $\sum_\psi w_\psi \Gamma_\psi^{(2)}$ ($w_\psi \geqslant 0$, $\sum_\psi w_\psi = 1$) of second-order density matrices of normalized wavefunctions $\psi \in H(N)$. The question of finding a characterization of $\mathscr{P}_N^{(2)}$ was formulated in [17] as the ensemble N-representability problem. The convex structure of $\mathscr{P}_N^{(2)}$ was studied, e.g., in [18, 21, 37].

The N-representability problem can be formulated similarly for density matrices of any order $p \geqslant 1$. The ensemble N-representability problem for density matrices of order $p = 1$ was solved in [17] (see also [64]). Namely, a matrix $\Gamma(x_1' \mid x_1)$ is of the form $\sum w_\psi \Gamma_\psi^{(1)}(x_1' \mid x_1)$ for $w_\psi \geqslant 0$, $\sum w_\psi = 1$, $\langle \psi, \psi \rangle = 1$ and $\psi \in H(N)$ if and only if $\mathrm{Tr}(\Gamma) = \int \Gamma(x_1 \mid x_1)\,dx_1 = N$ and the eigenvalues of Γ satisfy $0 \leqslant \lambda \leqslant 1$. However, the ensemble N-representability problem is already difficult for density matrices of order $p = 2$. In fact, as stated in Theorem 7.1, the representability problem for their diagonal elements is equivalent to the membership problem in the boolean quadric polytope and hence it is NP-hard. For $p \geqslant 2$, the representability problem involves not only conditions on the eigenvalues but also on the interrelations of the eigenvectors. On the other hand, no satisfactory solution exists for the pure N-representability problem even for the case $p = 1$.

Let $\mathrm{BQP}_n^{\mathscr{S}=2}(N)$ denote the polytope defined as the convex hull of the vectors $\pi^{\mathscr{S}=2}(K)$ for $K \subseteq \{1, \ldots, n\}$ of cardinality N. From relation (52), if $\psi = \Phi_K$ is a Slater determinant, then $\gamma_\psi(ij \mid hk) = 0$ except if $(i, j) = (h, k)$ and $i, j \in K$ in which case $\gamma_\psi(ij \mid ij) = 1$. Therefore, the diagonal terms of γ_{Φ_K} coincide with the vector $\pi^{\mathscr{S}=2}(K)$. For that reason, the polytope $\mathrm{BQP}_n^{\mathscr{S}=2}(N)$ is sometimes called the *N-Slater hull* (e.g., in [38, 40]).

From (50), the N-representability problem amounts to finding the boundary conditions on the coefficients $\gamma_\psi(ij \mid hk)$. In fact, the boundary conditions for the diagonal terms $\gamma_\psi(ij \mid ij)$ are precisely the valid inequalities for the N-Slater hull $\mathrm{BQP}_n^{\mathscr{S}=2}(N)$.

Theorem 7.1. *Given $\gamma = (\gamma(ij))_{1 \leqslant i < j \leqslant n}$, the following assertions are equivalent.*

(i) *There exists a normalized wavefunction $\psi \in H(N)$ such that $\gamma(ij) = \gamma_\psi(ij \mid ij)$ for all $1 \leqslant i < j \leqslant n$.*

(ii) *There exists a convex combination $\sum w_\psi \psi$ ($w_\psi \geqslant 0, \sum w_\psi = 1$) of normalized wavefunctions $\psi \in H(N)$ such that $\gamma(ij) = \sum w_\psi \gamma_\psi(ij \mid ij)$ for $1 \leqslant i < j \leqslant n$.*

(iii) *The vector γ belongs to $\mathrm{BQP}_n^{\mathscr{S}=2}(N)$.*

Proof. (i) $\Rightarrow$ (ii) is clear.

(ii) $\Rightarrow$ (iii): Suppose first that $\gamma(ij) = \gamma_\psi(ij \mid ij)$ for some normalized $\psi \in H(N)$ given by (49). Then, from (52), $\gamma = \sum_{K \subseteq \{1, \ldots, n\}, |K| = N} |C_K|^2 \pi^{\mathscr{S}=2}(K)$ with $\sum |C_K|^2 = \langle \psi, \psi \rangle = 1$. Hence $\gamma \in \mathrm{BQP}_n^{\mathscr{S}=2}(N)$. Suppose now that $\gamma(ij) = \sum w_\psi \gamma_\psi(ij \mid ij)$ with $w_\psi \geqslant 0$, $\sum w_\psi = 1$, $\psi \in H(N)$ and $\langle \psi, \psi \rangle = 1$. Then, $\gamma = \sum_K t_K \pi^{\mathscr{S}=2}(K)$, where $t_K = \sum_\psi w_\psi |C_K^\psi|^2 \geqslant 0$ and $\sum_K t_K = 1$. Therefore, $\gamma \in \mathrm{BQP}_n^{\mathscr{S}=2}(N)$.

(iii) $\Rightarrow$ (i): Assume $\gamma = \sum_K t_K \pi^{\mathscr{S}=2}(K)$ for $t_K \geqslant 0$ and $\sum_K t_K = 1$. Set $C_K = \sqrt{t_K}$ and $\psi = \sum_K C_K \Phi_K$. Then, $\gamma = \gamma_\psi$ holds. □

Therefore, the pure and ensemble representability problems are the same when restricted to the diagonal terms. However, in their general form, they are distinct problems. For instance, $\mathscr{P}_N^{(2)}$ has additional extreme points besides the second-order density matrices of the Slater determinants

(even though those are the only extreme points when restricted to the diagonal terms). Other extreme points for $\mathscr{P}_N^{(2)}$ are given in [18,37].

We conclude with some additional remarks.

(1) The N-representability problem for variable N leads to the study of the boolean quadric polytope BQP_n.

(2) The polytopes $\mathrm{BQP}_n^{\mathscr{J}=2}(N)$ and $\mathrm{BQP}_n(N) = \mathrm{BQP}_n^{\mathscr{J}\leqslant 2}(N)$, lying respectively in $\mathbb{R}^{\binom{n}{2}}$ and $\mathbb{R}^{\binom{n+1}{2}}$, are in one-to-one correspondence. Indeed, each point $x \in \mathrm{BQP}_n(N)$ saisfies the equations:

$$\sum_{1\leqslant i<j\leqslant n} x_{ij} = \binom{N}{2},$$

$$\sum_{1\leqslant j\leqslant n, j\neq i} x_{ij} = (N-1)x_{ii} \quad \text{for } 1\leqslant i\leqslant n.$$

Hence both polytopes have the dimension $\binom{n}{2} - 1$.

(3) The combinatorial interpretation of the N-representability problem from Theorem 7.1 was given in [95]. Actually, this paper treats the general problem of N-representability for density matrices of arbitrary order $p \geqslant 1$. We have exposed only the case $p = 2$ for the sake of simplicity and because this is the case directly relevant to our problem of cuts. For arbitrary $p \geqslant 2$, the analogue of Theorem 7.1 leads to the study of the polytope $\mathrm{BQP}_n^{\mathscr{J}=p}(N)$ in $\mathbb{R}^{\binom{n}{p}}$, defined as the convex hull of the $\mathscr{J}_{=p}$-intersection vectors $\pi^{\mathscr{J}=p}(K)$, for $K \subseteq \{1,\ldots,n\}$, $|K| = N$.

The facial structure of the polytope $\mathrm{BQP}_n^{\mathscr{J}=p}(N)$ is studied in [94]; in particular, the full description of its facets in the cases: $p = 2$, $N = 3$, $n = 6,7$ and partial results in the case: $p = 2$, $N = 3$, $n = 8$ are given there.

(4) An additional alternative interpretation of the boolean quadric polytope BQP_n is given in [40], in terms of positive semidefinite two-body operators.

Let a_i denote the *annihilation operator* of the Fock space $\bigcup_{N\geqslant 1} H(N)$ and $a_i^\dagger$, its adjoint, the *creation operator* (see [50]). Both are defined by their action on the Slater determinants. Namely, for $K = \{k_1,\ldots,k_N\}$ with $1 \leqslant k_1 < \cdots < k_N$,

$$a_i(\Phi_K) = \begin{cases} 0 & \text{if } i\notin K,\\ (-1)^{j-1}\Phi_{K-\{i\}} & \text{if } i = k_j \in K,\end{cases}$$

$$a_i^\dagger(\Phi_K) = \begin{cases} 0 & \text{if } i\in K,\\ (-1)^{j-1}\Phi_{K\cup\{i\}} & \text{if } i\notin K \text{ and } k_{j-1} < i < k_j.\end{cases}$$

Hence, $a_i^\dagger a_i(\Phi_K) = |K\cap\{i\}|\Phi_K$ for each $K \subseteq \{1,\ldots,n\}$. Therefore, the Slater determinants Φ_K are common eigenvectors for the operators $a_i^\dagger a_i$ and thus for any two-body operator of the form

$$B = b_0 + \sum_{1\leqslant i\leqslant n} b_i a_i^\dagger a_i + \sum_{1\leqslant i\leqslant j\leqslant n} b_{ij} a_i^\dagger a_i a_j^\dagger a_j. \tag{53}$$

The cone $Q^+(I^n)$, consisting of the two-body operators B of the form (53) which are positive semidefinite, is considered in [40]. Since any such operator has the same eigenvectors Φ_K associated with the eigenvalues $b_0 + \sum_{i\in K} b_i + \sum_{i,j\in K} b_{ij}$, the cone $Q^+(I^n)$ can be equivalently defined as the cone of the vectors $b := (b_0, b_i,\ 1\leqslant i\leqslant n, b_{ij},\ 1\leqslant i\leqslant j\leqslant n)$ for which $b(x) := b_0 + \sum_{1\leqslant i\leqslant n} b_i x_i + \sum_{1\leqslant i\leqslant j\leqslant n} b_{ij}x_i x_j \geqslant 0$ for each $x \in \{0,1\}^n$. Therefore, $Q^+(I^n)$ is the dual cone to BQP_n, i.e., $b \in Q^+(I^n)$ if and only if the inequality $b(x) \geqslant 0$ is valid for BQP_n.

The cone $Q^+(\mathbb{Z}^n) := \{b: b(x) \geqslant 0 \text{ for all } x \in \mathbb{Z}^n\}$, which corresponds to the case of a system of bosons (when several particles may occupy the same spin-orbital) while $Q^+(I^n)$ corresponds to a system of fermions (with at most one particle per spin-orbital), is also considered in [40].

Let us finally mention a connection between the hypermetric cone HYP_{n+1} and the cone $Q^+(\mathbb{Z}^n)$. It can be established via the covariance map φ_{c_0}. Namely,

$$\varphi_{c_0}(\mathrm{HYP}_{n+1}) = \left\{a = (a_{ij})_{1 \leqslant i \leqslant j \leqslant n}: \sum_{1 \leqslant i,j \leqslant n} a_{ij}x_ix_j - \sum_{1 \leqslant i \leqslant n} a_{ii}x_i \geqslant 0 \text{ for } x \in \mathbb{Z}^n\right\}$$

and, therefore,

$$\varphi_{c_0}(\mathrm{HYP}_{n+1}) = Q^+(\mathbb{Z}^n) \cap \{b: b_0 = 0,\ b_i = -b_{ii} \text{ for } 1 \leqslant i \leqslant n\}$$

is a section of the cone $Q^+(\mathbb{Z}^n)$.

7.3. *The quantum correlation polytope*

We address in this section a connection between the boolean quadric polytope BQP_n and the quantum correlation polytope, considered in [77, 78].

Recall that the boolean quadric polytope BQP_n arises naturaly in the theory of probability. Namely, from [31, Theorem 3.2], given $p = (p_{ij}, 1 \leqslant i \leqslant j \leqslant n) \in \mathbb{R}^{\binom{n+1}{2}}$, then $p \in \mathrm{BQP}_n$ if and only if there exist a probability space $(\Omega, \mathcal{A}, \mu)$ and n events $A_1, \ldots, A_n \in \mathcal{A}$ such that

$$p_{ij} = \mu(A_i \cap A_j) \quad \text{for all } 1 \leqslant i \leqslant j \leqslant n.$$

For that reason, the polytope BQP_n is also called the *correlation polytope* in [77–79]. For $n = 3$, BQP_n is also called the Bell–Wigner polytope.

As an extension, [77] introduces the quantum correlation polytope whose points represent the probability that a quantum mechanical system has the properties associated with two projection operators in a given state. We fix some notation.

As we saw before, the state of a quantum mechanical system is represented by a unit vector ψ of a Hilbert space H ($H = H(N)$ if the system has N particles). Let E_ψ denote the projection operator from H to the line spanned by ψ, i.e., $E_\psi(\phi) = \langle \psi, \phi \rangle \psi$ for $\phi \in H$. Equivalently, a state of the system is given by such a projection operator E_ψ; such a state is called a *pure state*. More generally, we consider also nonpure states, namely convex combinations of pure states: $W = \sum_\psi \lambda_\psi E_\psi$ ($\lambda_\psi \geqslant 0$, $\sum_\psi \lambda_\psi = 1$, $\psi \in H$ with $\langle \psi, \psi \rangle = 1$). Such states W are called *ensemble states*, or *mixtures*. Pure and ensemble states were already considered in Section 7.2. Alternatively, a state of the system is a bounded linear operator W of H which is Hermitian, positive semidefinite and has trace one.

Given $p = (p_{ij}, 1 \leqslant i \leqslant j \leqslant n) \in \mathbb{R}^{\binom{n+1}{2}}$, we say that p has a *quantum mechanical representation* if there exists a Hilbert space H, a state W, n projections $E_1, \ldots, E_n$ (not necessarily distinct, nor commuting) such that

$$p_{ij} = \mathrm{trace}(WE_i \wedge E_j) \quad \text{for } 1 \leqslant i \leqslant j \leqslant n,$$

where $E_i \wedge E_j$ denotes the projection from H to the subspace $E_i(H) \cap E_j(H)$. So p_{ij} represents the probability that the system has the properties associated with the projections E_i and E_j when it is in the state W. Let QCP_n denote the polytope in $\mathbb{R}^{\binom{n+1}{2}}$ consisting of those p which admit a quantum mechanical representation; QCP_n is called the *quantum correlation polytope*.

Finally let T_n denote the set of the vectors $p \in \mathbb{R}^{\binom{n+1}{2}}$ satisfying

$$0 \leqslant p_{ij} \leqslant \min(p_{ii}, p_{jj}) \leqslant \max(p_{ii}, p_{jj}) \leqslant 1$$

for $1 \leqslant i \leqslant j \leqslant n$. It is easy to see that the extreme points of T_n are exactly the vectors $p \in T_n$ with 0–1 coordinates.

Theorem 7.2. (i) $\mathrm{BQP}_n \subseteq \mathrm{QCP}_n \subseteq T_n$.

(ii) QCP_n *is a convex set which contains the interior of* T_n.

(iii) *The subset of* QCP_n *consisting of those* p *admitting a quantum mechanical representation in which the state* $W = E_\psi$ *is pure is also convex and contains the interior of* T_n.

For clarity, we give the proof of the statement (i) of Theorem 7.2.

Proof. The inclusion $\mathrm{QCP}_n \subseteq T_n$ follows from the fact that each state W is positive semidefinite with trace 1. We check the inclusion $\mathrm{BQP}_n \subseteq \mathrm{QCP}_n$. Let $p \in \mathrm{BQP}_n$. Hence $p = \sum_{K \subseteq \{1,\dots,n\}} \lambda_K \pi(K)$ where $\lambda_K \geqslant 0$ and $\sum_K \lambda_K = 1$. Let H be a Hilbert space of dimension 2^n and let $(\psi_K, K \subseteq \{1,\dots,n\})$ be an orthonormal basis of H indexed by the subsets of $\{1,\dots,n\}$. Let W be the operator of H defined by $W(\psi_K) = \lambda_K \psi_K$ for all K. Let E_i denote the projection from H to the subspace H_i spanned by the vectors ψ_K with $i \in K$; then $E_i \wedge E_j$ is the projection on the subspace spanned by ψ_K for $i, j \in K$. Note that the trace of the operator $WE_i \wedge E_j$ is equal to $\sum_{i,j \in K} \lambda_K = p_{ij}$. This shows that p belongs to QCP_n. □

Note that, if $p \in \mathrm{QCP}_n$ has a quantum mechanical representation in which the operators E_i commute, then, in fact, $p \in \mathrm{BQP}_n$.

Note also that every $p \in L_n$ with $0 < p_{ij} < 1$ for all i, j belongs to QCP_n. Therefore, except for some boundary cases, every $p \in T_n$ has a quantum mechanical representation, i.e. the only requirements for joint probabilities in the quantum case are that probabilities be numbers between 0 and 1 and that the probability of the joint be less than or equal to the probability of each event. Hence the probabilities of quantum mechanical events do not obey the laws of classical probability theory. New theories of quantum probability and quantum logic have been developed; see, for instance, [77, 78].

The region $\mathrm{QCP}_n - \mathrm{BQP}_n$ is called the *interference region*. Several examples of physical experiments are described in [77, 78] that yield some pair distributions p lying in the interference region. For example, the classical Einstein–Podolsky–Rosen experiment [36] yields $p \in \mathrm{QCP}_3 - \mathrm{BQP}_3$.

We conclude this section with a concrete example in the simplest case $n = 2$. Consider the vector $p = (p_{11} = p_{22} = (\cos\theta)^2, p_{12} = 0)$. Then, $p \notin \mathrm{BQP}_2$ if $1 > (\cos\theta)^2 > \frac{1}{2}$, since it violates the inequality $p_{11} + p_{22} - p_{12} \leqslant 1$. On the other hand, $p \in \mathrm{QCP}_2$. Indeed, let $H = \mathbb{R}^3$ be a Hilbert space with canonical basis (e_1, e_2, e_3), W be the projection on the line spanned by e_3 and let E_i be the projection on the line spanned by u_i, for $i = 1, 2$, where $u_1 = (\sin\theta, 0, \cos\theta)$ and $u_2 = (-\sin\theta, 0, \cos\theta)$. Then, $\mathrm{trace}(WE_i) = (\cos\theta)^2 = p_{ii}$ for $i = 1, 2$ and $E_1 \wedge E_2 = 0$.

The vector p has the following physical interpretation. Consider a source of photons all polarized in the e_3 direction in the space. Let $\psi = e_3$ be the quantum mechanical wavefunction associated with these photons, so $W = E_\psi$ is the state of the system. The projection E_i corresponds to the property "the photon is polarized in the direction u_i"; this corresponds to the experiment where

a polarizer is located in front of the source, oriented in the direction u_i and p_{ii} counts the frequency of the photons which pass through the polarizer. The relation $p_{12} = 0$ should be understood as follows. There may be some photons having both properties E_1 and E_2, but no experiment exists which could detect the simultaneous existence of the properties E_1 and E_2.

Note that BQP_2 has the following extreme points: (0,0,0), (1,0,0), (0,1,0), and (1,1,1), while T_2 has one more extreme point (1,1,0). In fact, $QCP_2 = T_2 - \{(1,1,0)\}$.

8. Other applications

8.1. The L_1-metric in probability theory

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $X: \Omega \to \mathbb{R}$ be a random variable with finite expected value $E(X) = \int_\Omega |X(\omega)| \mu(d\omega) < \infty$, i.e., $X \in L_1(\Omega, \mathcal{A}, \mu)$. Let F_X denote the distribution function of X, i.e., $F_X(x) = \mu(\{\omega \in \Omega: X(\omega) = x\})$ for $x \in \mathbb{R}$; when it exists, its derivative F'_X is called the *density* of X. A great variety of metrics on random variables are studied in the monograph [84]; among them, the following are based on the L_1-metric:

- the usual L_1-metric between the random variables,

$$L_1(X, Y) = E(|X - Y|) = \int_\Omega |X(\omega) - Y(\omega)| \mu(d\omega);$$

- the Monge–Kantorovich–Wasserstein metric (i.e., the L_1-metric between the distribution functions), $k(X, Y) = \int_{\mathbb{R}} |F_X(x) - F_Y(x)| \, dx$;
- the total valuation metric (i.e., the L_1-metric between the densities when they exist), $\sigma(X, Y) = \frac{1}{2} \int_{\mathbb{R}} |F'_X(x) - F'_Y(x)| \, dx$;
- the engineer metric (i.e., the L_1-metric between the expected values), $EN(X, Y) = |E(X) - E(Y)|$;
- the indicator metric,

$$i(X, Y) = E(1_{X \neq Y}) = \mu(\{\omega \in \Omega: X(\omega) \neq Y(\omega)\}).$$

In fact, the L_p-analogues ($1 \leqslant p \leqslant \infty$) of the above metrics, especially of the first two, are also used in probability theory.

Several results are known, establishing links among the above metrics. One of the main such results is the Monge–Kantorovich mass-transportation theorem which shows that the second metric $k(X, Y)$ can be viewed as a minimum of the first metric $L_1(X, Y)$ over all joint distributions of X and Y with fixed marginal. A relationship between the $L_1(X, Y)$ and the engineer metric $EN(X, Y)$ is given by [84] as solution of a moment problem. Similarly, a connection between the total valuation metric $\sigma(X, Y)$ and the indicator metric $i(X, Y)$ is given in Dobrushin's theorem on the existence and uniqueness of Gibbs fields in statistical physics. See [84] for a detailed account of the above topics.

We mention another example of use of the L_1-metric in probability theory, namely for Gaussian random fields. We refer to [73, 74] for a detailed account. Let $B = (B(x); x \in M)$ be a centered

Gaussian system with parameter space M, $0 \in M$. The variance of the increment is denoted by:

$$d(x,y) := E((B(x) - B(y))^2) \quad \text{for } x, y \in M.$$

When (M,d) is a metric space which is L_1-embeddable, the Gaussian system is called a Lévy's Brownian motion with parameter space (M,d). The case $M = \mathbb{R}^n$ and $d(x,y) = \|x - y\|_2$ gives the usual Brownian motion with n-dimensional parameter. By [31, Lemma 3.5], (M,d) is L_1-embeddable if and only if there exist a nonnegative measure space (H,ν) and a map $x \mapsto A_x \subseteq H$ with $\nu(A_x) < \infty$ for $x \in M$, such that $d(x,y) = \nu(A_x \triangle A_y)$ for $x, y \in M$. Hence, a Gaussian system admits a representation called of *Chentsov type*

$$B(x) = \int_{A_x} W(\mathrm{d}h) \quad \text{for } x \in M$$

in terms of a Gaussian random measure based on the measure space (H,ν) with $d(x,y) = \nu(A_x \triangle A_y)$ if and only if d is L_1-embeddable.

This Chentsov-type representation can be compared with the Crofton formula for projective metrics from [31, Theorem 4.12]. Actually both come naturally together in [1] (see parts A.8 and A.9 of Appendix A there).

8.2. *The ℓ_1-metric in statistical data analysis*

A *data structure* is a pair (I,d), where I is a finite set, called *population*, and $d: I \times I \to \mathbb{R}_+$ is a symmetric map with $d_{ii} = 0$ for $i \in I$, called *dissimilarity index*. The typical problem in statistical data analysis is to choose a "good representation" of a data structure; usually, "good" means a representation allowing to represent the data structure visually by a graphic display. Each sort of visual display corresponds, in fact, to a special choice of the dissimilarity index as a distance and the problem turns out to be the classical isometric embedding problem in special classes of metrics.

For instance, in hierarchical classification, the case when d is ultrametric corresponds to the possibility of a so-called indexed hierarchy (see [57]). A natural extension is the case when d is the path metric of a weighted tree, i.e., d satisfies the four-point condition (see [31, Section 4.1]); then the data structure is called an *additive tree*. Also, data structures (I,d) for which d is ℓ_2-embeddable are considered in factor analysis and multidimensional scaling. These two cases together with cluster analysis are the main three techniques for studying data structures. The case when d is ℓ_1-embeddable is a natural extension of the ultrametric and ℓ_2 cases.

An ℓ_p-approximation consists of minimizing the estimator $\|e\|_p$, where e is a vector or a random variable (representing an error, deviation, etc.). The following criteria are used in statistical data analysis:

- the ℓ_2-norm, in the least-square method or its square;
- the ℓ_∞-norm, in the minimax or Chebychev method;
- the ℓ_1-norm, in the least absolute values (LAV) method.

In fact, the ℓ_1 criterion has been increasingly used. Its importance can be seen, for instance, from the volume [34] of proceedings of a conference entitled "Statistical data analysis based on the L_1-norm and related methods"; we refer, in particular, to [15, 33, 44, 91].

8.3. Hypercube embeddings and designs

In this section, we describe how some questions about the existence of special classes of designs are connected with questions about $\mathbb{Z}_+$-realizations of the equidistant metric $2td(K_n)$ and, in particular, about its minimum h-size.

We recall some definitions.

Given integers $n, t \geqslant 1$, $d(K_n)$ denotes the path metric of the complete graph K_n and $2td(K_n)$ is the equidistant metric with components all equal to $2t$. The metric $2td(K_n)$ is clearly h-embeddable, since $2td(K_n) = \sum_{1 \leqslant i \leqslant n} t\delta(\{i\})$, called its *starcut realization*. Any decomposition of $2td(K_n)$ as $\sum_{S \in \mathcal{B}} \delta(S)$, where $\mathcal{B}$ is a collection of (nonnecessarily distinct) subsets of $V_n = \{1, \ldots, n\}$, is called a *$\mathbb{Z}_+$-realization* of $2td(K_n)$ and $|\mathcal{B}|$ (counting the multiplicities) is its *size*. The $\mathbb{Z}_+$-realization is called *k-uniform* if $|S| = k$ holds for all $S \in \mathcal{B}$. Let z_n^t denote the minimum size of a $\mathbb{Z}_+$-realization of $2td(K_n)$. The metric $2td(K_n)$ is h-rigid if the starcut realization is its only $\mathbb{Z}_+$-realization, i.e., $z_n^t = nt$.

In fact, the set families $\mathcal{B}$ giving $\mathbb{Z}_+$-realizations of $2td(K_n)$, i.e., for which $2td(K_n) = \sum_{S \in \mathcal{B}} \delta(S)$, correspond to some designs. Let us first recall some notions about designs; for details about designs, see, e.g., [85].

Let $\mathcal{B}$ be a collection of (nonnecessarily distinct) subsets of V_n, the sets $B \in \mathcal{B}$ are called *blocks*. Let r, k, λ be integers.

Then, $\mathcal{B}$ is called a $(r, \lambda; n)$-*design* if each point $i \in V_n$ belongs to r blocks and any two distinct points $i, j \in V_n$ belong to λ common blocks.

$\mathcal{B}$ is called an (n, k, λ)-BIBD (BIBD standing for balanced incomplete block design) if any two distinct points $i, j \in V_n$ belong to λ common blocks and each block has cardinality k. This implies that each point $i \in V_n$ belong to $r = \lambda(n-1)/(k-1)$ blocks and the total number of blocks is $b := |\mathcal{B}| = rn/k$. It is well known that $b \geqslant n$ holds. The BIBD is called *symmetric* if $b = n$ or, equivalently, $r = k$ holds. Two important cases of symmetric BIBD are

- the projective plane PG$(2, t)$, i.e. $(t^2 + t + 1, t + 1, 1)$-BIBD,
- the Hadamard design of order $4t - 1$, i.e. $(4t - 1, 2t, t)$-BIBD.

It is well known that a Hadamard design of order $4t - 1$ corresponds to a Hadamard matrix of order $4t$ (i.e., a matrix with ± 1 entries whose rows are pairwise orthogonal).

We have the following links between the $\mathbb{Z}_+$-realizations of $2td(K_n)$ and designs [30]:

(i) There is a one-to-one correspondence between the $\mathbb{Z}_+$-realizations of $2td(K_n)$ and the $(2t, t; n-1)$-designs.

(ii) There is a one-to-one correspondence between the k-uniform $\mathbb{Z}_+$-realizations of $2td(K_n)$ and the (n, k, λ)-BIBD, where the parameters satisfy:

$$r = \frac{t(n-1)}{n-k}, \qquad \lambda = r - t = \frac{t(k-1)}{n-k}.$$

(iii) If there exists a symmetric $(n, \lambda + t, t)$-BIBD with $n \neq 4t$, $n = 2t + \lambda + t(t-1)/\lambda$, then $z_n^t = n$ [85].

In the cases $\lambda = 1, t$, the implication of (iii) is, in fact, an equivalence. Namely, we have:

(iv)

$$\text{PG}(2, t) \text{ exists} \overset{[54] \text{ and } [85]}{\Leftrightarrow} z_{t^2+t+1}^t = t^2 + t + 1$$

$$\overset{[24]}{\Leftrightarrow} \quad 2td(K_{t^2+t+2}) \text{ is not } h\text{-rigid},$$

$$\text{i.e., } z^t_{t^2+t+2} < t(t^2+t+2)$$

$$\overset{[30]}{\Leftrightarrow} \quad z^t_{t^2+t+2} = t^2 + 2t \text{ if } t \geqslant 3$$

$$= t^2 + t + 1 \text{ if } t = 1, 2.$$

(v) There exists a Hadamard matrix of order $4t \Leftrightarrow z^t_{4t-1} = 4t - 1 \Leftrightarrow z^t_{4t} = 4t - 1$ [85].

The following bounds hold for z^t_n:

(vi) by [31, (13)], $z^t_n \leqslant nt$, with equality if and only if $2td(K_n)$ is h-rigid;

(vii) $z^t_n \geqslant n - 1$, with equality if and only if $n = 4t$ and there exists a Hadamard matrix of order $4t$ [85];

(viii) $z^t_n \geqslant n$ if we are not in the case of equality of (vii);

(ix) by [31, (13)],

$$z^t_n \geqslant a^t_n := \left\lceil \frac{n(n-1)t}{\lfloor \frac{1}{2}n \rfloor \lceil \frac{1}{2}n \rceil} \right\rceil = 4t - \left\lfloor \frac{2t}{\lceil \frac{1}{2}n \rceil} \right\rfloor.$$

Observe that $a^t_{4t} = a^t_{4t-1} = 4t - 1$, and $a^t_{t^2+t+1} = a^t_{t^2+t+2} = 4t$ if $t \geqslant 3$.

From (iv), there exists a projective plane $PG(2, t)$ if and only if equality holds in the bound (viii) for $n = t^2 + t + 1$ or, equivalently, there is a strict inequality in the bound (vii) for $n = t^2 + t + 2$. From (v), there exists a Hadamard matrix of order $4t$ if and only if equality holds in the bounds (vii) and (ix) for $n = 4t$ or, equivalently, equality holds in the bounds (viii) and (ix) for $n = 4t - 1$.

Therefore, the $\mathbb{Z}_+$-realizations of minimum size of $2td(K_n)$ provide a common generalization of the two most interesting cases of symmetric BIBD, namely finite projective planes and Hadamard designs.

Finally, we mention a conjecture which generalizes the implication (iii) in the case $\lambda = t$; it is stated and partially proved in [30].

Conjecture 8.1. (a) *For $n \leqslant 4t$, if there exists a Hadamard matrix of order $4t$, then $z^t_n = a^t_n$.*

(b) *If $\lceil \frac{1}{2}n \rceil$ divides $2t$ and there exists a Hadamard matrix of order $4t$, then $z^t_n = a^t_n$.*

8.4. Miscellaneous

The variety of uses of the ℓ_1-metric is very vast as we already saw in Sections 8.1 and 8.2. We group here several other examples where ℓ_1-embeddable metrics are useful.

On the integers, besides the usual ℓ_1-metric $|a - b|$, we have, for instance, the well-known Hamming distance between the binary expansions of a, b, and $\log(\text{l.c.m.}(a, b)/\text{g.c.d.}(a, b))$ (mentioned after [31, Theorem 4.13]) which are both ℓ_1-embeddable.

Two examples of ℓ_1-embeddable metrics are used in biology:

- The *Prevosti's genetic distance*: $(1/2r)\sum_{1 \leqslant j \leqslant r}\sum_{1 \leqslant i \leqslant k_j} |p_{ij} - q_{ij}|$ between two populations P and Q, where r is the number of loci or chromosomes, p_{ij} (resp. q_{ij}) is the frequency of the chromosomal ordering i in the locus or chromosome j within the population P (resp. Q); the literature on this distance started in [83].

- The *biotope distance*: $|A \triangle B|/|A \cup B|$ between biotopes A, B (sets of species in, say, forests); it was introduced in [71] and it is shown in [2] to be ℓ_1-embeddable.

The so-called *chemical distance* between two graphs G_1, G_2 on n nodes: $\min \| A_1 - P^T A_2 P \|$, where the minimum is taken over all $n \times n$ permutation matrices P, and A_i denotes the adjacency matrix of G_i, for $i = 1, 2$ [65].

The Hamming distance $|\{(a, b) \in G^2 : a \cdot b \neq a * b\}|$ between the multiplication tables of two groups $A = (G, \cdot)$ and $B = (G, *)$ on the same underlying set G is used in [35].

Given compact subsets A, B of the plane $\mathbb{R}^2$, the ℓ_1-distance aire$(A \triangle B)$ is used in the treatment of images; see, for instance, [58].

References

[1] R.V. Ambarzumian, *Combinatorial Integral Geometry* (Wiley, New York, 1982).

[2] P. Assouad, Produit tensoriel, distances extrémales et réalisation de covariances, *C. R. Acad. Sci. Paris* **288**(A) (1979) 649–652 and 675–677.

[3] P. Assouad, Sous espaces de L^1 et inégalités hypermétriques, *C. R. Acad. Sci. Paris* **294**(A) (1982) 439–442.

[4] D. Avis and M. Deza, The cut cone, L^1-embeddability, complexity and multicommodity flows, *Networks* **21** (1991) 595–617.

[5] D. Avis and V.P. Grishukhin, A bound on the k-gonality of the hypermetric cone and related complexity problems, *Comput. Geom.* **2** (1993) 241–254.

[6] F. Barahona, The max-cut problem on graphs not contractible to K_5, *Oper. Res. Lett.* **2** (1983) 107–111.

[7] F. Barahona, M. Grötschel, M. Jünger and G. Reinelt, An application of combinatorial optimization to statistical physics and circuit layout design, *Oper. Res.* **36**(3) (1988) 493–513.

[8] F. Barahona, M. Jünger and G. Reinelt, Experiments in quadratic 0–1 programming, *Math. Programming* **44** (1989) 127–137.

[9] F. Barahona and A.R. Mahjoub, On the cut polytope, *Math. Programming* **36** (1986) 157–173.

[10] C.E. Bonferroni, Il calcolo delle assicurazioni su grouppi di teste, *Studi in Onore del Professor S.O. Carboni (Roma)* (1936).

[11] G. Boole, *An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities* (Dover, New York, 1854).

[12] E. Boros and P.L. Hammer, The max-cut poblem and quadratic 0–1 optimization; polyhedral aspects, relaxations and bounds, *Ann. Oper. Res.* **33** (1991) 151–180.

[13] E. Boros and A. Prékopa, Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur, *Math. Oper. Res.* **14**(2) (1989) 317–342.

[14] F.C. Bussemaker, D.M. Cvetkovic and J.J. Seidel, Graphs related to exceptional root systems, TH-report 76-WSK-05, Technical Univ. Eindhoven, 1976.

[15] G. Le Calve, L_1-embeddings of a data structure (I, D), in: Y. Dodge, Ed, *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Elsevier, Amsterdam, 1987) 195–202.

[16] K.L. Chung, On the probability of the occurrence of at least m events among n arbitrary events, *Ann. Math. Statist.* **12** (1941) 328–338.

[17] A.J. Coleman, The structure of fermion density matrices – I, *Rev. Modern Phys.* **35** (1963) 668.

[18] A.J. Coleman, Necessary conditions for N-representability of reduced density matrices, *J. Math. Phys.* **13**(2) (1972) 214–222.

[19] A.J. Coleman and R.M. Erdahl, Eds., *Reduced Density Operators with Applications to Physical and Chemical Systems I*, Queen's Papers in Pure and Appl. Math. **11** (Queen's Univ., Kingston, Ont., 1968).

[20] J.V. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups, Grundlehren Math. Wiss.* **290** (Springer, Berlin, 1987).

[21] E. Davidson, Linear inequalities for density matrices, *J. Math. Phys.* **10**(4) (1969) 725–734.

[22] D.A. Dawson and D. Sankoff, An inequality for probabilities, *Proc. Amer. Math. Soc.* **18** (1967) 504–507.
[23] C. Delorme and S. Poljak, Laplacian eigenvalues and the maximum cut problem, *Math. Programming* **62** (1993) 557–574.
[24] M. Deza, Une propriété extrémale des plans projectifs finis dans une classe de codes equidistants, *Discrete Math.* **6** (1973) 343–352.
[25] M. Deza and V.P. Grishukhin, L-polytopes and equiangular lines, *Discrete Appl. Math.* **56** (1995) 181–214.
[26] M. Deza and V.P. Grishukhin, Hypermetric graphs, *Quart. J. Math. Oxford* **2** (1993) 399–433.
[27] M. Deza, V.P. Grishukhin and M. Laurent, Extreme hypermetrics and L-polytopes, in: *Sets, Graphs and Numbers*, Budapest, 1991, Colloq. Math. Soc. János Bolyai **60** (North-Holland, Amsterdam, 1992) 157–209.
[28] M. Deza, V.P. Grishukhin and M. Laurent, The hypermetric cone is polyhedral, *Combinatorica* **13**(4) (1993) 1–15.
[29] M. Deza and M. Laurent, A survey of the known facets of the cut cone, Research report 91722-OR, Instit. Diskrete Math., Univ. Bonn, 1991.
[30] M. Deza and M. Laurent, Variety of hypercube embeddings of the equidistant metric and designs, *J. Combin. Inform. System Sci.* **18**(3–4) (1993) 293–320.
[31] M. Deza and M. Laurent, Applications of cut polyhedra – I, *J. Comput. Appl. Math.* **55**(2) (1994) 191–216 (this issue).
[32] P.A.M. Dirac, Note on exchange phenomena in the Thomas atom, *Proc. Cambridge Philos. Soc.* **26** (1930) 376–385.
[33] Y. Dodge, An introduction to statistical data analysis L_1-norm based, in: Y. Dodge, Ed., *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Elsevier, Amsterdam, 1987) 1–22.
[34] Y. Dodge, Ed., *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Elsevier, Amsterdam, 1987).
[35] A. Drápal, How far apart can the group multiplication tables be?, *European J. Combin.* **13**(5) (1992) 335–344.
[36] A. Einstein, B. Podolsky and W. Rosen, Can quantum mechanical description of physical reality be considered complete?, *Phys. Rev.* **47** (1935) 777 (reprinted in Wheeler and Zurek (1983).
[37] R.M. Erdahl, The convex structure of the set of N-representable reduced 2-matrices, *J. Math. Phys.* **13**(10) (1972) 1608–1621.
[38] R.M. Erdahl, A convex set of second order inhomogeneous polynomials with applications to reduced density matrices, in: R.M. Erdahl, Ed., *Reduced Density Operators with Applications to Physical and Chemical Systems II*, Queen's Papers in Pure and App. Math. **40** (Queen's Univ., Kingston, Ont., 1974) 28–35.
[39] R.M. Erdahl, Ed., *Reduced Density Operators with Applications to Physical and Chemical Systems II*, Queen's papers in Pure and Appl. mat. **40** (Queen's Univ., Kingston, Ont. 1974).
[40] R.M. Erdahl, Representability conditions, in: R.M. Erdahl and V.H. Smith, Eds., *Density Matrices and Density Functionals* (Reidel, Dordrecht, 1987) 51–75.
[41] R.M. Erdahl, A cone of inhomogeneous second-order polynomials, *Discrete Comput. Geom.* **8**(4) (1992) 387–416.
[42] R.M. Erdahl and S.S. Ryshkov, The empty sphere, *Canad. J. Math.* **39**(4) (1987) 794–824.
[43] R.M. Erdahl and V.H. Smith, Eds., *Density Matrices and Density Functionals* (Reidel, Dordrecht, 1987).
[44] B. Fichet, The role played by L_1 in data analysis, in: Y. Dodge, Ed., *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Elsevier, Amsterdam, 1987) 185–193.
[45] J. Fonlupt, A.R. Mahjoub and J.P. Uhry, Compositions in the bipartite subgraph polytope, *Discrete Math.* **105** (1992) 73–91.
[46] L.R. Ford and D.R. Fulkerson, *Flows in Networks* (Princeton Univ. Press, Princeton, NJ, 1962).
[47] J. Galambos, Bonferroni inequalities, *Ann. Probab.* **5** (1977) 577–581.
[48] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, CA, 1979).
[49] M.X. Goemans and D.P. Williamson, 0.787-approximation algorithms for MAX CUT and MAX 2SAT, in: *Proc. 26th Ann. ACM Symp. on the Theory of Computing* (ACM, New York, 1994).
[50] E.K.U. Gross, E. Runge and O. Heinonen, *Many-Particle Theory* (IOP Publishing, Bristol, 1991).
[51] M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization* (Springer, Berlin, 1988).
[52] M. Grötschel and W.R. Pulleybank, Weakly bipartite graphs and the max-cut problem, *Oper. Res. Lett.* **1** (1981) 23–27.

[53] F.O. Hadlock, Finding a maximum cut of planar graphs in polynomial time, *SIAM J. Comput.* **4** (1975) 221–225.
[54] J.I. Hall, Bounds for equidistant codes and projective planes, *Discrete Math.* **17**(1) (1977) 85–94.
[55] K. Husimi, *Proc. Physico-Math. Soc. Japan* **22** (1940) 264.
[56] M. Iri, On an extension of the maximum-flow minimum-cut theorem to multicommodity flows, *J. Oper. Res. Soc. Japan* **13** (1970–1971) 129–135.
[57] S.C. Johnson, Hierarchical clustering schemes, *Psychometrika* **32**(3) (1967) 241–254.
[58] M. Jourlin, Métriques de formes et applications, in: *Proc. Conf. Distancia*, Rennes (1992) 279–282.
[59] A.V. Karzanov, Metrics and undirected cuts, *Math. Programming* **32** (1985) 183–198.
[60] A.V. Karzanov, Half-integral five-terminus flows, *Discrete Appl. Math.* **18** (1987) 263–278.
[61] A.V. Karzanov, Polyhedra related to undirected multicommodity flows, *Linear Algebra Appl.* **114–115** (1989) 293–328.
[62] A.V. Karzanov, Sums of cuts and bipartite metrics, *European J. Combin.* **11** (1990) 473–484.
[63] S. Kounias and J. Marin, Best linear Bonferroni bounds, *SIAM J. Appl. Math.* **30**(2) (1976) 307–323.
[64] H.W. Kuhn, Linear inequalities and the Pauli principle, in: Proc. Sympos. Appl. Math. **10** (Amer. Mathematical Soc., Providence, RI, 1960) 141–147.
[65] V. Kvasnička and J. Pospichal, Two metrics for a graph-theoretical model of organic chemistry, *J. Math. Chem.* **3** (1989) 161–191.
[66] M. Laurent and S. Poljak, One-third-integrality in the metric polytope, Report BS-R9209, Centrum voor Wiskunde en Informatica, Amsterdam, 1992.
[67] P.W.H. Lemmens and J.J. Seidel, Equiangular lines, *J. Algebra* **24** (1973) 494–512.
[68] M.V. Lomonossov, On a system of flows in a network, *Problemy Peredachi Informatsii* **14** (1978) 60–73 (in Russian).
[69] M.V. Lomonossov, Combinatorial approaches to multiflow problems, *Discrete Appl. Math.* **11** (1985) 1–93.
[70] P.O. Löwdin, Quantum theory of many particles systems I, Physical interpretation by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction, *Phys. Rev.* **97** (1955) 1474–1489.
[71] F. Marczewski and H. Steinhaus, On a certain distance of sets and the corresponding distance of functions, *Colloq. Math.* **6** (1958) 319–327.
[72] W.B. McRae and E.R. Davidson, Linear inequalities for density matrices II, *J. Math. Phys.* **13** (1972) 1527–1538.
[73] A. Noda, Generalized Radon transform and Lévy's Brownian motion I and II, *Nagoya Math. J.* **105** (1987) 71–87 and 89–107.
[74] A. Noda, White noise representation of some Gausian random fields, *Bull. Aichi Univ. Education* **38** (1989) 35–45.
[75] K. Onaga and O. Kakusho, On feasibility conditions of multicommodity flows in networks, *IEEE Trans. Circuit Theory* **18** (1971) 425–429.
[76] B.A. Papernov, On existence of multicommodity flows, *Studies in Discrete Optimization* (Nauka, Moscow, 1976) 230–261 (in Russian).
[77] I. Pitowsky, The range of quantum probability, *J. Math. Phys.* **27** (1986) 1556–1565.
[78] I. Pitowsky, *Quantum probability — Quantum logic*, Lecture Notes in Phys. **321** (Springer, Berlin, 1989).
[79] I. Pitowsky, Correlation polytopes: their geometry and complexity, *Math. Programming* **50** (1991) 395–414
[80] S. Poljak, Polyhedral and eigenvalue approximations of the max-cut problem, in: *Sets, Graphs and Numbers*, Budapest, 1991, Colloq. Math. Soc. János Bolyai **60** (North-Holland, Amsterdam, 1992).
[81] S. Poljak and F. Rendl, Solving the max-cut problem using eigenvalues, Research Report No. 91735-OR, Inst. Diskrete Math., Univ. Bonn, 1991.
[82] S. Poljak and F. Rendl, Nonpolyhedral relaxations of graph-bisection problems, Tech. Report 92-55, DIMACS, 1992.
[83] A. Prevosti, J. Ocana and G. Alonso, Distances between populations of drosophila subobscura based on chromosome arrangement frequencies, *Theoret. Appl. Genetics* **45** (1975) 231–241.
[84] S.T. Rachev, *Probability Metrics and the Stability of Stochastic Models* (Wiley, New York, 1991).
[85] H.J. Ryser, *Combinatorial Mathematics, Carus Math. Monographs* **14** (Math. Assoc. America, New York, 1963).
[86] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, New York, 1986).
[87] A. Schrijver, private communication, 1992.

[88] W. Schwärzler and A. Sebö, A generalized cut condition for multiflows in matroids, Research Report 869-M, Artemis, IMAG, Univ. Joseph Fourier, Grenoble, 1992.
[89] P.D. Seymour, Matroids and multicommodity flows, *European J. Combin.* **2** (1981) 257–290.
[90] P. Terwiliger and M. Deza, Classification of finite connected hypermetric spaces, *Graphs Combin.* **3** (1987) 293–298.
[91] I. Vajda, L_1-distances in statistical inference: comparison of topological functional and statistical properties, in: Y. Dodge, Ed., *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (Elsevier, Amsterdam, 1987) 177–184.
[92] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, NJ, 1955) (original German ed. in 1933).
[93] G.F. Voronoi, Nouvelles applications des paramètres continus à la théorie des formes quadratiques, Deuxième mémoire, *J. Reine Angew. Math.* **134** (1908) 198–287; **136** (1909) 67–181.
[94] M.L. Yoseloff, A combinatorial approach to the diagonal N-representability problem, Ph.D. Thesis, Princeton Univ., 1970.
[95] M.L. Yoseloff and H.W. Kuhn, Combinatorial approach to the N-representability of p-density matrices, *J. Math. Phys.* **10** (1969) 703–706.