Abstract.

In this working paper I discuss some issues, with the aim of suggesting research topics for the DIMACS special year "Combinatorial Optimization". I plan to focus on the unorthodox, hoping that at least one of the areas I describe below will gain new momentum as a result of this special year. A companion preprint, written by Martin Grötschel and myself as a chapter of the Handbook of Combinatorics, describes the state of the art, concentrating on the well-established methods and models like polyhedral combinatorics, and should be available soon.

0. Introduction

A typical problem in combinatorial optimization is the following: given a combinatorially defined set of 0-1 vectors (called *feasible solutions*), find one that maximizes a given linear objective function. A by now standard method (and often the method of choice) is to consider the convex hull of feasible solutions, find a description of it by linear inequalities, and then use linear programming to optimize the given objective function subject to these constraints.

However, combinatorial optimization cannot be restricted to polyhedral combinatorics. I would like to give examples of combinatorial optimization problems that are of a different kind and methods that do not work through polyhedral combinatorics.

My best examples concern the stable set problem, the stable set polyhedron STAB(G) of a graph, and perfect graphs. I hope that this is due to personal bias: the methods I describe extend naturally to quite general situations, and I hope that they turn out useful in the study of other combinatorial optimization problems.

1. Quadratic optimization

Polyhedral combinatorics can be viewed as a theory of linear inequalities valid for the incidence vectors of feasible solutions. It is natural to ask for quadratic inequalities (and, of course, higher degree inequalities) valid for these vectors.

As an example, consider the stable set problem. At the first sight it seems that we are getting too much too easily. Let $G = (V, E)$ be a graph and consider the following system of equations:

(1.1) $x_i^2 = x_i$ for every node $i \in V$,

(1.2) $x_i x_j = 0$ for every edge $ij \in E$.

Trivially, the solutions of (1.1) are precisely the 0-1 vectors, and so the solutions of (1.1)–(1.2) are precisely the incidence vectors of stable sets.

Unfortunately, little is known about the solutions of systems of quadratic equations. In fact, what (1.1) and (1.2) show is that even the solvability of such a simple system of quadratic equations (together with a linear equation $\sum_i x_i = t$) is NP-hard (whether we want to find solutions in the rational, real, complex, or any finite field). It follows that it is NP-hard to decide whether a system of quadratic inequalities implies another one. Therefore a complete P-time calculus of quadratic inequalities, analogous to linear programming, cannot be developed.

However, something is true. First note that it *can* be decided in polynomial time whether a single quadratic inequality has a solution. This can be reduced to the positibe definiteness of a matrix, which can be checked by Gaussian elimination. It is also known that a quadratic polynomial is identically non-negative iff it is a sum of squares of linear

polynomials, and is identically positive if and only if it is a sum of a positive constant and squares of linear polynomials. The positive definiteness check gives such a decomposition.

Second, it *can* be decided in polynomial time whether a given system of quadratic inequalities has a linear combination with non-negative coefficients that is identically false. In fact, one can work in the $\binom{n}{2} + n + 1$ dimensional space of quadratic inequalities, and check if the convex hull of the given inequalities intersects the convex cone of identically false inequalities. This can be done using the ellipsoid method (see Grötschel, Lovász and Schrijver (1990) [GLS] as general reference; lots of details must be filled in to make the above sketch precise).

Unfortunately, for quadratic inequalities there is no Farkas lemma: a given system may be unsolvable and yet have no identically false combination with non-negative coefficients. Let us say that the system is *strongly unsolvable* if it does. Replacing unsolvability by strong unsolvability leads to relaxations that are often of interest. For example, the smallest t for which the system (1.1), (1.2) and

$$\sum_i x_i \geq t$$

is not unsolvable is the stability number $\alpha(G)$; the smallest t for which this same system is not strongly unsolvable is the polynomial time computable upper bound $\vartheta(G)$ on the stability number (defined in Lovász 1989).

Among quadratic inequalities, there are other natural rules of inference besides linear combination with non-negative coefficients. One of these is the following: given two *linear* inequalities

$$\ell_1(x) \geq 0, \qquad \ell_2(x) \geq 0,$$

(these are special quadratic ones), we can form the quadratic inequality

$$\ell_1(x)\ell_2(x) \geq 0.$$

To illustrate that this method is quite powerful, we derive some known constraints for the stable set polytope. (1.1) implies that for every node i,

(1.3) $x_i = x_i^2 \geq 0, \quad 1 - x_i = (1 - x_i)^2 \geq 0,$

and using this (1.2) implies that for every edge ij,

(1.4) $1 - x_i - x_j = 1 - x_i - x_j + x_i x_j = (1 - x_i)(1 - x_j) \geq 0.$

So we can derive the edge constraints from (1.1)–(1.2) formally. We can go on and use these to derive the odd hole constraints. Consider e.g. a pentagon $(1, 2, 3, 4, 5)$. Then we have

$$1 - x_1 - x_2 - x_3 + x_1 x_3 = 1 - x_1 - x_2 - x_3 + x_1 x_2 + x_1 x_3 = (1 - x_1)(1 - x_2 - x_3) \geq 0$$

and similarly

$$1 - x_1 - x_4 - x_5 + x_1 x_4 \geq 0.$$

Furthermore,

$$x_1 - x_1x_3 - x_1x_4 = x_1(1 - x_3 - x_4) \geq 0.$$

Summing these inequalities, we get the odd hole constraint

(1.5) $2 - x_1 - x_2 - x_3 - x_4 - x_5 \geq 0.$

Making use of the fact that every square is non-negative, we can also derive the clique constraints. Assume that nodes $1, 2, 3, 4, 5$ induce a complete 5-graph. We start with the trivial inequality

$$(1 - x_1 - x_2 - x_3 - x_4 - x_5)^2 \geq 0.$$

Expanding, we get

$$1 + \sum_{i=1}^{5} x_i^2 - 2\sum_{i=1}^{5} x_i + 2\sum_{i \neq j} x_ix_j \geq 0.$$

Here the first sum is just $\sum_i x_i$ by (1.1) and the third sum is 0 by (1.2), so we get

(1.6) $\quad 1 - x_1 - x_2 - x_3 - x_4 - x_5 \geq 0.$

We can formalize these procedures: if we have a family of quadratic polynomials, we can apply to them the following rules of inference:

(R1) Form a non-negative linear combination of members of the family;

(R2) Take the product of two linear polynomials in the family;

(R3) Add a sum of squares of a linear polynomials to a member of the family.

It is usually convenient to assume in combinatorial situations that the inequalities $x_i \geq 0$ and $1 - x_i \geq 0$ are there to begin with. Then the following restricted version of (R2) turns out useful:

(R2') If ℓ is a linear polynomial in the family then $x_1\ell$ and $(1 - x_1)\ell$ are in the family.

One may also formulate the following converse to (R2):

(R4) If ℓ_1 and ℓ_2 are linear polynomials such that ℓ_1 and $\ell_1\ell_2 - 1$ are in the family, then we can add ℓ_2 to the family.

The examples above show that quite complicated inequalities can be derived this way in a few steps. In fact, it was shown by Lovász and Schrijver (1989, 1990) [LS] that if the starting family includes the polynomials x_i, $1 - x_i$ and $x_i^2 - x_i$ (so that all vectors on which these are non-negative are 0-1 vectors), then all linear polynomials that are non-negative whenever all of the original polynomials are, can be derived using only rules (R1) and (R2').

Problem 1. *Is a similar result valid for the derivation of quadratic (but non-linear) inequalities or for derivation from more general starting families?*

A deep and general result on inference among algebraic inequalities is Artin's Theorem in real algebraic geometry (see Jacobson 1980, Ch.11), but that seems to be too general to have direct algorithmic applications.

While many of the known facets of the stable set polytope can be derived simply in this framework, not all the important classes are clarified. If G is a line-graph, then STAB(G) is a matching polytope, and all of its facets are known from the classical work of Edmonds (1965).

Problem 2. *Is there a simple derivation of the Edmonds facets from (1.1)-(1.2) using (R1)-(R4), or (preferably) (R1)-(R3)?*

The relative strength of these rules is also unclear. It was shown in [LS] that it takes one round to get STAB(G) for a perfect graph using (R3) but $\omega(G)-2$ rounds if only (R1) and (R2') are allowed.

Problem 3. *Is there a similar gap when we allow (R4)? Is there any graph for which* STAB*(G) cannot be obtained in a bounded number of rounds using (R1) and (R3)? (This last assertion would follow from (1.7) below if $P \neq NP$.) Is a logarithmic number of rounds enough, if (say) (R1), (R2) and (R3) are used?*

Let $\mathcal{F}^k$ denote the set of quadratic inequalities obtained from (1.1), (1.2) and

$$x_i \geq 0, \qquad 1 - x_i - x_j \geq 0 \quad (\forall ij \in E)$$

in k rounds of application of (R1) and (R2'). (This means that we obtain $\mathcal{F}^{k+1}$ by applying (R2') to every linear inequality in $\mathcal{F}^k$, and take all non-negative linear combinations of the old and new quadratic inequalities.) Let $\mathcal{F}_1^k$ be the set of linear inequalities in $\mathcal{F}^k$. We have seen that the odd hole constraints belong to $\mathcal{F}_1^1$, and in fact it was shown in [LS] that $\mathcal{F}_1^1$ is generated by the odd hole contraints (1.4) and $\mathcal{F}_1^0$ via non-negative linear combinations.

Problem 4. *Can a generating set for $\mathcal{F}_0^k$ or $\mathcal{F}^k$ be described?*

In the special case when G is perfect, it follows that $\mathcal{F}_1^1$ is generated by the triangle constraints, edge constraints and non-negativity constraints. We also know that for $k = \omega(G) - 2$, $\mathcal{F}_1^k$ is generated by the clique constraints and the non-negativity constraints.

Problem 5. *Is it true that for a perfect graph G, $\mathcal{F}_1^k$ is generated by the constraints defined by cliques of size at most $k+2$ and the non-negativity constraints?*

More generally, consider an arbitrary graph G and a facet $a^T x \leq a_0$ of the solution set of $\mathcal{F}_1^k$. Is this a facet of STAB(G')*, where G' is the subgraph induced by those nodes of G with non-zero coefficient a_i?*

At this point, we have to address the question: what is the point in deriving inequalities algebraically, which can be proved true anyway by trivial combinatorial considerations? The following facts give the answer:

1.7 Fact. *Let $\mathcal{F}$ be a family of quadratic inequalities and let $\mathcal{F}^*$ be the family of linear inequalities obtained from $\mathcal{F}$ by using one round of inference rules (R2), (R3) and (R1). Then the optimization problem*

$$\text{maximize } c^T x$$

$$\text{subject to } \mathcal{F}^*$$

can be solved in polynomial time.

Problem 6. *Does this remain true if (R4) is also invoked?*

In particular, if G is a perfect graph then starting with (1.1) and (1.2), we obtain all facets of STAB(G), and thus we obtain that the stability number of a perfect graph can be computed in polynomial time. See [LS] for further versions of this fact.

Consider the system (1.1)-(1.2) and the family $\mathcal{F}$ of all linear inequalities that can be derived from it using rules (R1) and (R3) (it is easy to see that that no repeated application is needed: we take a sum of squares of linear polynomials and add to a a linear combination of the left hand sides of equations (1.1) and (1.2), so that all quadratic terms cancel). Let $TH(G)$ denote the solution set of $\mathcal{F}$ (see [GLS] for details). This set is a convex, but in general non-polyhedral body in $\mathbb{R}^n$; in fact, it is polyhedral if and only if G is perfect. (1.7) implies that any linear objective function can be optimized over it in polynomial time. In particular, one has

$$\vartheta(G) = \max\{\mathbf{1}^T x \ : \ x \in \mathrm{TH}(G)\}.$$

A further nice property of TH(G) is that for every graph G, the bodies TH(G) and TH($\overline{G}$) are antiblocking polytopes.

Problem 7. *Is there a more general duality theory of quadratic inequalities behind this fact?*

Perfect graphs are equivalent to antiblocking pairs of 0-1 polyhedra. Blocking pairs of polyhedra with 0-1 vertices correspond to clutters with the Max-Flow-Min-Cut property (see [GLS]).

A *clutter* is a pair $(V, \mathcal{H})$ where $\mathcal{H}$ is a family of mutually incomparable subsets of $\mathcal{H}$. The *blocker* of a clutter $(V, \mathcal{H})$ is the clutter $(V, \mathcal{H}^*)$, where $\mathcal{H}^*$ consists of all inclusion-minimal sets meeting every set in $\mathcal{H}$. It is easy to check that $\mathcal{H}^{**} = \mathcal{H}$. A clutter has the *Max-Flow-Min-Cut property* if the vertices of the polyhedron defined by

$$x_i \geq 0, \qquad \sum_{i \in B} x_i \geq 1 \qquad (\forall B \in \mathcal{H}^*)$$

are exactly the incidence vectors of members of $\mathcal{H}$.

Lehman proved that *if a clutter has the Max-Flow-Min-Cut property then so does its blocker.* It was proved in [GLS] that *if $(V, \mathcal{H})$ has the Max-Flow-Min-Cut property and for*

every weighting $w \in \mathbb{R}_+^V$ a minimum weight member of $\mathcal{H}$ can be found in polynomial time, then the same holds for its blocker.

However, perhaps the following much stronger assertion could be shown. Note that the analogous result for antiblocking pairs is true, and in fact follows e.g. by the quadratic inequality method.

Problem 8. *Prove of disprove: Let $(V, \mathcal{H})$ be a clutter with the Max-Flow-Min-Cut property. Assume that we have an oracle to decide if a set $X \subseteq V$ contains a member of calH. Then for every weighting $w \in \mathbb{R}_+^V$ a minimum weight member of $\mathcal{H}$ can be found in polynomial time.*

An application of the quadratic method to the theory of interactive proofs was proposed by Feige and Lovász (1991). Let us formulate a special case which involves only pure graph theory. Let $G_1 = (V_1, E_1)$ and $G_2 = (V - 2, E_2)$ be two graphs. We write that $G_1 \longrightarrow G_2$ if there is a homomorphism of G_1 into G_2, i.e. a mapping of V_1 into V_2 which maps adjacent nodes to adjacent nodes. Special cases of this notion are: G_1 is k-colorable (choose G_2 the complete k-graph), or G_2 has a k-clique (choose G_1 the complete k-graph).

Each mapping ϕ from V_1 to V_2 can be represented by a 0-1 matrix (u_{si}) ($s \in V_1$, $i \in V_2$) such that

$$u_{si} = \begin{cases} 1, & \text{if } \phi(s) = i, \\ 0, & \text{otherwise.} \end{cases}$$

The fact that this represents a mapping can be expressed by the equations

$$u_{si}^2 - u_{si} = 0, \qquad \sum_i u_{si} = 1.$$

The fact that we have a homomorphism is equivalent to the equations

$$u_{si} u_{tk} = 0 \qquad \text{if } si \in E_1 \text{ but } tk \notin E_2.$$

(One may use other equations like

$$u_{sj} u_{sk} = 0 \qquad \text{if } j \neq k,$$

but these would not bring anything new, at least as long as we are on this theoretical level.)

This system of quadratic equations is solable if and only if $G_1 \longrightarrow G_2$. Let us write $G_1 \Longrightarrow G_2$ if the system is not strongly unsolvable. We can think of this as "G_1 almost maps into G_2". The following lemma is easily verified.

1.8 Lemma. *If $G_1 \Longrightarrow G_2$ and $G_2 \Longrightarrow G_3$ then $G_1 \Longrightarrow G_3$.*

Every pair of graphs G_1 and G_2 with $G_1 \not\longrightarrow G_2$ defines two classes of finite graphs:

$$\mathcal{K} = \{H : \ G_1 \longrightarrow H\},$$

and

$$\mathcal{K}' = \{H : H \longrightarrow G_2\},$$

which are trivially disjoint, and both are in NP. An outstanding open problem in complexity theory asks if any two such classes can be separated in P, i.e., if there is a polynomial time decidable class $\mathcal{K}''$ of graphs such that

$$\mathcal{K} \subseteq \mathcal{K}'', \qquad \mathcal{K}' \cap \mathcal{K}'' = \emptyset.$$

While we do not know the answer to this general question, we can show that it is in the affirmative if we make the stronger assumption that $G_1 \not\Longrightarrow G_2$.

1.9 Theorem. *Let G_1 and G_2 be two graphs such that $G_1 \not\Longrightarrow G_2$. Then the class*

$$\mathcal{K}'' = \{H : G_1 \Longrightarrow H\}$$

is in P and separates $\mathcal{K}$ and $\mathcal{K}'$.

To apply this theorem, we have to find interesting pairs of graphs with $G_1 \not\Longrightarrow G_2$. One easy example is given by the following:

$$K_k \not\Longrightarrow K_{k-1}.$$

(This is not entirely obvious to see.) It follows that the classes of graphs containing a k-clique can be separated from the class of $(k-1)$-colorable graphs by a class in P. This has been known, and in fact the methods obtaining it served as a motivation for these results.

We can obtain further interesting properties expressible as the existence of a homomorphism if we allow the graphs to have directed edges and colored edges (which must be mapped onto edges with the same color). It is also easy to formulate the method when the mapping is injective, surjective or bijective, thus including the hamilton cycle problem and the isomorphism problem. For the complexity problem formulated above, the following question is crucial:

Problem 9. *Find other interesting pairs of graphs with $G_1 \not\Longrightarrow G_2$.*

Expressing the isomorphism problem as the solvability of a system of quadratic equations, is there a pair of non-isomorphic graphs for which the system is not strongly unsolvable?

The above ideas are closely related (in fact, equivalent) to *semidefinite programming.* Let $A(x)$ be a matrix in which each entry is a linear form of the variables $x \in \mathbb{R}^n$, let $c \in \mathbb{R}^n$, and consider the optimization problem

$$\text{maximize } \{c^T x \mid A(x) \text{ positive semidefinite.}\}$$

If A is a diagonal matrix, then this is just a linear program. In general, the constraint that $A(x)$ is positive semidefinite is equivalent to infinitely many linear constraints

$$v^T A(x) v \geq 0 \text{ for all } v \in \mathbb{R}^n,$$

so the set of feasible vectors for this problem is convex (but in general not polyhedral). The question whether a quadratic constraint can be derived from a given set of quadratic contraints using (R1) and (R3) can easily be formulated as a semidefinite program.

Much of the machinary of linear programming extends to semidefinite programs; in particular, duality theory, the ellipsoid method, and Karmarkar's method. See e.g. Fletcher (1985), Overton (1988), Alizadeh (1992). We do not go into the details, but remark that this approach promises the practically most efficient implementations of the ideas in this section.

We conclude this section with formulating some problems of a more algebraic nature. Let $G = (V, E)$ be a graph and let $I(G)$ denote the polynomial ideal generated by the polynomials $x_i^2 - x_i$ $(i \in V)$ and by the polynomials $x_i x_j$ $(ij \in E)$. Consider the quotient ring $R = \mathbb{R}[x_1, \ldots, x_n]/I(G)$ (this is related to the so-called *Reisner–Stanley ring* of the graph; cf. Stanley 1983). We can introduce an order on polynomials by writing $f \geq 0 \pmod{I(G)}$ iff $f(x) \geq 0$ for every root of the ideal $I(G)$; clearly, this induces an order on the quotient ring. It is not difficult to prove the following:

1.10 Lemma. *For any polynomial f, we have $f \geq 0 \pmod{I(G)}$ iff there exist polynomials $g_1, \ldots, g_N$ such that $f \equiv g_1^2 + \ldots + g_N^2 \pmod{I(G)}$.*

This is not true for all ideals I in place of $I(G)$; for $I = \{0\}$ Artin proved, solving Hilbert's 17^{th} problem, that every non-negative polynomial is a sum of squares of rational functions.

Problem 10. *How to to characterize those ideals that have this property?*

The following theorem shows an interesting connection between this property and the perfectness of a graph:

1.11 Theorem. *A graph G is perfect if and only if the following holds:*

$(*)$ *For any linear polynomial f, we have $f \geq 0 \pmod{I(G)}$ iff there exist linear polynomials $g_1, \ldots, g_N$ such that $f \equiv g_1^2 + \ldots + g_N^2 \pmod{I(G)}$.*

The proof of this theorem follows from the characterization of perfectness in terms of the body $\mathrm{TH}(G)$ (see [GLS]). One could formulate analogous characterizations of t-perfect graphs using Corollary 1.10.

Problem 11. *Which polynomial ideals in general have property $(*)$ (and so can be considered as an algebraic generalization of perfect graphs)?*

The method sketched here can be extended, at least in some sense, ¿from quadratic to higher-order inequalities. For these extensions, see Lovász and Schrijver (1989, 1990) and Sherali and Adams (1990a-b). But in fact very little is known about these possible applications of real algebraic geometry to combinatorics.

2. Projection representation

If we project a polytope to a subspace then the number of vertices can, of course not increase; but we have little control over the the number of facets, which may increase substantially. This seemingly negative fact can be turned around: a polytope having an inconveniently large number of facets may be represented as the projection of a polytope with a much smaller number of facets.

If we find a representation of a polytope P as the projection of another polytope P' with only a polynomial number of facets, then this can be used to optimize linear objective functions over P in polynomial time, by optimizing over P' instead.

This technique in polyhedral optimization is relatively new and promising; Ball, Liu and Pulleyblank (1987), Maculan (1987), Balas and Pulleyblank (1983, 1987), Barahona and Mahjoub (1987), Cameron and Edmonds (1988) provide examples of interesting applications of this idea. At the same time, many of the basic questions are unsettled.

As an example, consider the stable set polytope of a comparability graph G, obtained by taking a partially ordered set $(V, \leq)$, and connecting two elements of V iff they are comparable in this partial order. This graph is perfect and so its stable set polytope is described by the non-negativity and clique constraints:

$x_i \geq 0$ for every $i \in V$,

$\sum_{i \in B} x_i \leq 1$ for every chain $B \subseteq V$.

This letter family of constraints is typically exponentially large. However STAB(G) can be represented as the projection of a polytope with $O(|V|^2)$ facets:

2.1 Lemma. *Let $(V, \leq)$ be a poset. Assign two variables x_i, y_i to each $i \in V$ and consider the polytope P defined by the inequalities*

$0 \leq x_i \leq y_i \leq 1$ *for all* $i \in V$,

$y_i + x_j \leq y_j$ *for all pairs* $i < j$.

Then projecting P on the x-coordinates, we obtain STAB(G).

This example raises the following interesting question:

Problem 12. *For which graphs G can* STAB(G) *be represented as the projection of a polytope with a polynomial number of facets?*

The answer is in the affirmative for comparability graphs and their complements, chordal graphs and their complements, and several other classes of perfect graphs, but is not known for all perfect graphs. The answer is positive for t-perfect graphs; this can be shown using the quadratic constraint method described in section 1 [LS]. One suspects that the answer will be negative for a general graph; however, the only negative result to date is that of Yannakakis (1988), who proves that if we also require that the automorphism group of G "lifts up" to isometries of the polytope then the stable set polytope of the line-graph of K_n (i.e., the matching polytope of K_n) cannot be represented as the projection of a polytope with a polynomial number of facets.

Yannakakis also formulated the following combinatorial necessary condition. Let $\{U_1, W_1\}$, $\{U_2, W_2\}$, ..., $\{U_N, W_N\}$ be partitions of $V(G)$ into two classes. We say that this family *separates cliques and stable sets* if for every stable set A and clique B such that $A \cap B = \emptyset$ there exists an i with $A \subseteq U_i$ and $B \subseteq W_i$.

2.2 Lemma. *If* STAB(G) *can be represented as the projection of a polytope with N facets then there exist N partitions of $V(G)$ separating cliques from stable sets.*

So this leads to the problem:

Problem 13. *Can the cliques in every graph (or even in every perfect graph) be separated from stable sets by a polynomial number of partitions?*

It follows from the theory of communication complexity that $O(n^{\log n})$ partitions suffice in every graph. This was improved by A. Hajnal (unpublished) to $O(n^{(1/2)\log n})$.

Another interesting class of polytopes from this point of view is the class of basis polytopes of matroids. It follows from the work of Maculan (1987) that for a graphic matroid on m elements, the basis polytope can be represented as the projection of a polytope with only $O(m)$ facets. This is trivially also valid for cographic matroids (which have essentially the same basis polytopes).

Problem 14. *For which matroids can the basis polytope be represented as the projection of a polytope with polynomially many facets?*

Probably it can be shown, using Seymour's decomposition, that the answer is positive for regular matroids. On the other hand, I expect that the answer is negative for general matroids. A heuristic argument can be given by counting. There are doubly exponentially many matroids on n elements, so if the basis polytope of each of them had a representation as the projection of a polytope with (say) n^{100} vertices, then one of these representing polytopes would necessarily have a facet with either irrational or exponentially long coefficients. Perhaps with some number theoretic argument one can show that so ugly coefficients are not needed.

Let $P \subseteq \mathbb{R}^n$ be a polyhedron, and let $\psi(P)$ denote the minimum number of facets of a polyhedron P' that can be projected so as to obtain P. It is easy to see that P' need never have higher dimension than $\psi(P)-1$, so we don't have to worry about the number of variables introduced. This number ψ appears to have some nice properties. For example, it is not difficult to prove:

2.3 Fact *If P contains the origin in its interior and P^* is its polar, then $\psi(P^*) = \psi(P)$.*

The ψ values of blocking and anti-blocking pairs of polyhedra are not the same in general, but differ by at most n.

On the other hand, ψ appears to be difficult to compute. The following question should be easier than the two more specific problems stated above.

Problem 15. *Is there a polytope in $\mathbb{R}^n$ with 0-1 vertices for which $\psi(P)$ is not polynomial in n?*

The answer is surely negative; but I don't have the slightest idea how to prove it. This is my favorite question in this paper. Can the answer be combinatorial, based on some property of the f-vector (the sequence of numbers of faces in various dimensions)? Can the interesting theory of fiber polytopes (Billera and Strumfels 1992), related to many other aspects of projections of polyhedra, be used?

3. Entropy

Graph entropy, introduced by Körner (1973), is an information theoretic functional on a graph with a probability distribution given on its vertex set; it may also be considered as a probabilistic refinement of the notion of chromatic number.

Definition. Let $G = (V, E)$ be a graph and let p be a probability distribution on V. Let $G^{(n)} = (V^n, E^{(n)})$ denote the n-th conormal power of G, i.e., V^n is the set of sequences of length n from V, and

$$E^{(n)} = \{(x, y) \in V^n \times V^n : \exists i : (x_i, y_i) \in E\}.$$

Define the probability distribution p^n on V^n by $p^n(x) = \prod_{i=1}^n p(x_i)$. For $U \subseteq V^n$, let $G^{(n)}(U)$ denote the subgraph induced by U in $G^{(n)}$, and let $\chi(G^{(n)}(U))$ denote its chromatic number. Then for every $0 < \varepsilon < 1$, the limit

$$H(G,p) = \lim_{n\to\infty} \frac{1}{n} \min_{p^n(U)\geq 1-\varepsilon} \log \chi(G^{(n)}(U))$$

exists and is independent of ε. $H(G,p)$ is called the *graph entropy* of the graph G with respect to the probability distribution p.

We may view the elements of V as an alphabet, two letters being connected by an edge iff they are "distinguishable". The elements of V^n are words of length k. Two such words are connected in the n-th conormal power iff they are distinguishable (i.e., they are distinguishable in at least one position). We want to encode these words by 0-1 words of length as smal;l as possible, so that distinguishable words get different codes (two words that are indistinguishable anyway may get the same code). Such an encoding corresponds to a coloring of the n-th conormal power. However, we only want to encode the "majority" of words, i.e., a fraction of ε is allowed not to get codes. Then for large n, the optimum encoding uses words with length about $H(G,p) \cdot n$.

The *Shannon entropy* of a probability distribution $p \in \mathbb{R}^V$ is defined by $H(p) = -\sum_{i\in V} p_i \log p_i$. The entropy of the complete graph with respect to the probability distribution p is just $H(p)$. The following result of Csiszár, Körner, Lovász, Marton and Simonyi (1990) [CKLMS] shows that $H(G,p)$ can be defined by an optimization problem over the stable set polytope STAB(G).

3.1 Theorem.

$$H(G,p) = \min\left\{ \sum_{i\in V} -p_i \log x_i \;\middle|\; x \in \mathrm{STAB}(G) \right\}.$$

Since the objective function is concave, this optimization problem is well behaved; in particular, $H(G,p)$ can be computed in polynomial time using the ellipsoid method.

Graph entropy can be used to obtain lower bounds for the minimum number of graphs of a given type needed to cover the edge set of a fixed graph, and to obtain lower bounds on the complexity of certain algorithms; cf. Körner (1986), Körner and Marton (1988), Boppana (1989), Newman, Ragde and Wigderson (1990), Radhakhrishnan (1991), Kahn and Kim (1992). These applications are based on the following sub-additivity property of graph entropy.

Let $F = (V, E_1)$, $G = (V, E_2)$ be graphs on the same vertex set V. Körner (1986) proved that

$$H(F \cup G, p) \leq H(F, p) + H(G, p)$$

for every probability distribution p on V. (This properety is the key to most applications of graph entropy in complexity theory.) In particular,

$$H(G,p) + H(\overline{G},p) \geq H(K_n,p) = H(p). \qquad (*)$$

Graph entropy is also closely related to perfectness. Körner and Longo (1973) raised the question of characterizing those graphs for which equality holds in $(*)$ for all p. This was answered in [CKLMS]:

3.2 Theorem. *The equality*

$$H(G,p) + H(\overline{G},p) = H(p).$$

holds for every probability distribution p iff G is perfect.

Studying the structure of optimum solutions of the above optimization problem leads to the following (somewhat unexplored) property of perfect graphs not involving entropy:

3.4 Theorem. *Let G be a perfect graph on n vertices. Then G contains a family $\mathcal{A}$ of independent sets and a family $\mathcal{B}$ of cliques with the following properties:*

(a) $|\mathcal{A}| + |\mathcal{B}| = n + 1$;

(b) *the sets in $\mathcal{A}$ ($\mathcal{B}$) cover all points;*

(c) *the incidence vectors of sets in $\mathcal{A}$ ($\mathcal{B}$) are linearly independent;*

(d) *every $A \in \mathcal{A}$ intersects every $B \in \mathcal{B}$.*

This close connection between graph entropy and quite classical combinatorial properties of graphs suggests that it might be worth while to study the "entropy" of other combinatorial polytopes. Some general facts were established in [CKLMS], but there are very few concrete examples. The entropy of a relaxation of STAB was considered by Marton (1990). In a quite different setting, Pulleyblank (1992) considers an entropy-type objective function as the dual of a certain non-linear optimization problem.

One may also wonder if there is any connection between these occurences of entropy and the entropy-type barrier functions occurring e.g. in Karmarkar's linear programming algorithm.

4. Rubber bands, rods, cables, and such

Let us replace each edge of a graph G by a rubber band, and let it find its equilibrium. This is of course the position when all nodes are in the same place; let us make some assumption to prevent this, e.g. nail down some of the nodes. Then the equilibrium configuration will be quite closely related to a number of graph-theoretic properties.

Example 1. Let us choose the Hooke constants for the rubber bands as algebraically independent transcendentals. Take k nodes $u_1, \dots, u_k$ of G and nail them to k affine independent points in $\mathbb{R}^{k-1}$. Let $v_1, \dots, v_k$ be k further nodes. Then there exist k node-disjoint paths from $u_1, \dots, u_k$ to $v_1, \dots, v_k$ iff $v_1, \dots, v_k$ are in affine independent positions in the equilibrium.

Various versions of this result, along with algorithmic applications, were given by Linial, Lovász and Wigderson (1988). The construction in Example 1 also relates to random walks, as the following simple (folklore) result shows.

Example 2. Let us choose the same Hooke constant for every rubber band. Take k nodes $u_0, \dots, u_{k-1}$ and nail them down at the vertices $e_0 = 0, e_1, \dots, e_{k-1}$ of the standard simplex in $\mathbb{R}^{k-1}$. Consider a vertex a and $\alpha_0, \dots, \alpha_{k-1}$ be its coordinates in the equilibrium position. Then for every $0 \le i \le k-1$, α_i is equal to the probability that a random walk on graph G starting at a hits v_i before any other v_j. If $k = 2$, then the energy of the equilibrium state is m/t, where m is the number of edges and t is the expected time for a random walk starting at v_0 to reach v_1 and then return to v_0.

A promising fact about these constructions is that they can be related to minor-preserving properties. Here is a simple example.

Example 3. For a fixed $d \ge 1$, consider the class $\mathcal{H}^d$ of graphs that can be represented on the sphere S^d so that not all nodes are in the same position, and replacing the edges by rubber bands with appropriate Hooke constants, we get a locally stable state.

For $d = 1$, we get graphs containing a circuit. For $d = 2$, we get graphs containing a K_4 minor. In general, the complement of $\mathcal{H}^d$ is closed under minors. To see this, assume that the deletion $G \setminus e$ can be represented in S^d. Then we can add e with an almost infinitesimally weak Hooke constant, to get a representation of G. Similarly, assume that the contraction G/e has a representation in S^d. Then we can view this as a representation of G where e is an infinitely strong rubber band, and then relax this rubber band a little to get a representation of G. The local stability implies that this can be done.

More constructions could be mentioned. Tutte (1963) uses rubber bands to show that every 3-connected planar graph has an embedding in the plane with straight segments as edges and convex polygons as faces.

Replacing the rubber bands by rigid bars and/or cables, we find ourselves in the classical field of statics, which has a huge literature (see e.g. Recski (1989), White and Whiteley (1983) for some of the combinatorially most interesting aspects of this theory.)

Let G be a graph on n nodes. A *generalized adjacency matrix* of G is a symmetric $n \times n$ matrix A, whose rows and columns are indexed by the nodes of G, having $A_{ij} = 0$ if $i \neq j$ and $ij \notin E(G)$. (So the diagonal entries are arbitrary.) A d-dimensional *orthogonal representation* of the graph G is an assignment of a vector $u_i \in \mathbb{R}^d$ to each node i such that $u_i^T u_j = 0$ whenever $i \neq j$ and $ij \notin E$.

The following connections between these constructions are obvious. The Gram matrix of an orthogonal representation is a generalized adjacency matrix. Conversely, given any generalized adjacency matrix, we can add a constant in the diagonal to make it positive semidefinite, and then it can be written as the Gram matrix of an orthogonal representation.

Assume that A is a generalized adjacency matrix of G of rank d, and also assume that the off-diagonal entries of A are non-negative. Then there are $n - d$ linearly independent vectors $v_1, \ldots, v_{n-d}$ such that $Av_k = 0$ for $k = 1, \ldots, n - d$. Write $v_k = (v_{k1}, \ldots, v_{kn})^T$, and let $u_i = (v_{1i}, \ldots, v_{n-d,i})^T \in \mathbb{R}^{n-d}$. Then the definition of the v_k can be written as

$$\sum_{i \in V} a_{ij} u_j = 0 \qquad (j \in V).$$

Set $b_i = -\sum_j a_{ij}$, then this last equation is equivalent to

$$\sum_{i \in V} a_{ij}(u_j - u_i) + b_i(0 - u_i) = 0.$$

This eqution can be interpreted as follows. Place node i of G in $\mathbb{R}^{n-d}$ in position u_i. Connect u_i to u_j by a rubber band with Hooke constant a_{ij} for every edge ij. Connect node i to the origin by a bar. Then the system will be in equilibrium. We may replace the bars by rubber bands with Hooke-constants b_i; since b_i may be negative, these "rubber bands" may push their endpoints apart (and the farther these endpoints are, the stronger they push).

In the case of our first example, there seems to be a complete parallelism between characterizations of connectivity in terms of rubber bands, generalized adjacency matrices, and orthogonal representations. Lovász, Saks and Schrijver (1989) [LSS] proved: *A graph G on n nodes is k-connected if and only if it has an orthogonal representation in $\mathbb{R}^{n-k}$ such that any $n - k$ representing vectors are linearly independent.* The following result is folklore. *A graph G on n nodes is k-connected if and only if it has a generalized adjacency matrix in which every $(n - k) \times (n - k)$ minor is non-singular.*

In [LSS] the algebraic variety $\mathcal{V}^d(G)$ of all orthogonal representations of a graph G in $\mathbb{R}^d$ was studied. It was shown that this variety is reducible if G is not $(n - d)$-connected. Assume that G is $(n - d)$-connected; then the variety is irreducible if $d \leq 3$ but may be

reducible for $d = 4$, e.g. when G is the complement of the cube. While some partial results are given, the following question remains open:

Problem 16. *For which graphs is $\mathcal{V}^d(G)$ irreducible?*

It may also be interesting to study the structure of the variety of generalized adjacency matrices with given rank, or with the property that every $(n-k) \times (n-k)$ minor is non-singular.

Orthogonal representations are also closely related to the quadratic constraint method described in section 1, in particular to the body TH(G). See [GLS] for details.

Example 3 was motivated by a more involved, but related construction due to Colin de Verdière (1986, 1992). He considers a generalized adjacency matrix A of G having negative entries corresponding to the edges, and having a single negative eigenvalue. He also stipulates the following condition, called the *Strong Arnold Hypothesis*: for every symmetric matrix $B \in \mathbb{R}^{V \times V}$ there exists a generalized adjacency matrix B' such that $v^T B v = v^T B' v$ for every v with $Av = 0$. Let $n - \mu(G)$ be the minimum rank of such a matrix A. Then he proves that $k(G)$ decreases (non-strictly) under taking minors, and that $\mu(G) \leq 2$ and 3 characterizes outerplanar and planar graphs, respectively.

Using the general construction as described above, Colin de Verdiére's parameter can be interpreted as the maximum dimension in which a certain rubber-band representation exists. The fact that there is a single negative eigenvalue corresponds to the stability of the configuration.

Problem 17. *Translate the Strong Arnold Hypothesis in terms of this representation in any transparent way.*

5. Randomized algorithms; simulated annealing

There is a very general (and somewhat controversial) scheme for solving combinatorial optimization problems. Suppose that we want to maximize an objective function c defined on an (exponentially large) set Ω of feasible solutions. There are several combinatorial algorithms that utilize a local structure of Ω; more exactly, they assume that there is a graph defined on Ω, and that having a feasible solution we can somehow check whether it has a neighbor that gives better objective value.

The most important algorithm in this category is the simplex method, but there are many heuristics, like 2-opt and 3-opt for the Travelling Salesman problem, that use local improvement.

In a randomized version of local search, a random neighbor of the current feasible solution is selected. If this improves the objective function, the current feasible solution is replaced by this neighbor. If not, it may still be replaced, but only with some probability less than 1, depending on how much worse the objective value at the neighbor is.

A more general way of looking at this method is to consider it as a method of generating a random element in Ω, from some given probability distribution Q. Let $f : \Omega \to \mathbb{R}_+$ be the objective function; then maximizing f over Ω is just the extreme case when we want

to generate a random element from a distribution concentrated on the set of optimum solutions. If, instead, we generate a random point w from the distibution Q in which $Q(v)$ is proportional to $\exp(f(v)/T)$, where T is a very small positive number, then with large probability w will maximize f. In fact, a randomized algorithm that finds a solution that is (nearly) optimal with large probability is equivalent to a procedure of generating a random element from a distribution that is heavily concentrated on the (nearly) optimal solutions.

To generate a random element from a distribution over a (large and complicated) set is of course a much more general question, and is a major step in various algorithms for enumeration, integration, volume computation, simulation, statistical sampling, etc. (see Jerrum, Vazirani and Vazirani (1983), Dyer and Frieze (1990), Sinclair and Jerrum (1988), Dyer, Frieze and Kannan (1989), Lovász and Simonovits (1990, 1992) [LS] for some of the applications with combinatorial flavor). An efficient general technique here is *random walks* or *Markov chains*. Let $G = (\Omega, E)$ be a connected graph on Ω, and assume, for simplicity of presentation, that G is non-bipartite and d-regular. If we start a random walk on G and follow it long enough, then the current point will be almost uniformly distributed over Ω. How many steps does "long enough" mean depends on the spectrum, or in combinatorial terms on global connectivity properties called expansion rate or conductance, of the graph; see Diaconis (1988), Diaconis and Stroock (1991), and the papers quoted above.

In optimization, we are interested in very non-uniform, rather than uniform, distributions. Fortunately, there is an elegant way, called the *Metropolis filter* (Metropolis, Rosenbluth, Rosenbluth, Teller and Teller 1953), to modify the random walk, so that it gives any arbitrary prescribed probability distribution. Let $F : \Omega \to \mathbb{R}_+$. Assume that we are at node v. We choose a random neighbor u. If $F(u) > F(v)$ then we move to u; else, we flip a biased coin and move to u only with probability $F(u)/F(v)$, and stay at v with probability $1 - F(u)/F(v)$.

The miraculous property of this Metropolis-filtered random walk is that its stationary distribution Q_F (the distribution to which the distribution of the t^{th} point of the random walk tends as $t \to \infty$) has the property that $Q_F(v)$ is proportional to $F(v)$. So choosing $F(v) = \exp(f(v)/T)$, we get a randomized optimization algorithm.

Unfortunately, the issue of how long one has to walk gets rather messy. The techniques to estimate the conductance of a Metropolis-filtered walk are not general enough, although Applegate and Kannan (1991) have applied this technique to volume computation, and Diaconis and Hanlon (1992) extended certain eigenvalue techniques to this case. Roughly speaking, in both cases the underlying graph has to have good conductance, and the function has to be "smooth" and "log-concave".

Problem 18. *Find a lower bound on the conductance of a Metropolis-filtered random walk, assuming the underlying graph has good conductance and the filtering function is "nice".*

Coming back to optimization, let us follow Kirckpatrick, Gelatt and Vecchl (1983), and call the elements of Ω *states* (of some physical system), $1/F(v)$ the *energy* of the state, and T the *temperature*. In this language, we want to find a state with minimum (or almost

minimum) energy. A random walk means letting the system get into a stationary state at the given temperature.

The main, and not quite understood, issue is the choice of the temperature. If we choose T large, the quality of the solution is poor, i.e. the probability that it is close to being optimal is small. If we choose T small, then there will be barriers of very small probability (or, equivalently, with very large energy) between local optima, and it will take extremely long to get away from a local, but not global optimum.

The technique of *simulated annealing* suggests to start with the temperature T sufficiently large, so that the random walk with this parameter mixes fast. Then we decrease T gradually. In each phase, the random walk starts from a distribution which is already close to the limiting distribution, so there is hope that the walk will mix fast. (A similar trick works quite well in integration and volume computation; see [LS].)

There are many empirical studies with this method; see Johnson, Aragon, McGeoch and Schevon (1989, 1991) or Johnson (1990). There are also some general estimates on its performance (Holley and Stroock 1988, Holley, Kusuoka and Stroock 1989). Examples of problems, in particular of the matching problem, are known where simulated annealing performs badly (Sasaki and Hajek 1988, Sasaki 1991), and some positive results in the case of the matching problem are also known (Jerrum and Sinclair 1989). But I feel it is justified to state as an important research problem:

Problem 19. *Find reasonably interesting combinatorial optimization problems for which it can be proved that simulated annealing gives, in polynomial time and with probability at least 1/2, a solution that is within a factor of 2 to the optimum (and of course to get such an approximate solution by other means should be at least non-trivial).* Of course, instead of the factor of 2, other measures of approximation can be considered.

Random walk techniques may be useful in other types of questions. Here is a problem, which I feel is quite basic:

POLYTOPE-POLYHEDRON PROBLEM. Let $v_1, \ldots, v_p \in \mathbb{Q}^n$ be a list of vectors and $a_1^T x \leq b_1$, ..., $a_m^T x \leq b_m$, a system of linear inequalities. Is the convex hull P of the given vectors the same as the solution set P' of the given linear ineqialities?

It is clear that this problem is in co-NP. One may of course check that each v_i satisfies each $a_j^T x \leq b_j$. One may also assume that each v_i is a vertex of P, and that each $a_j^T x \leq b_j$ is a facet of P'. Then one may also check that each v_i is a vertex of P' and that each $a_j^T x \leq b_j$ is a facet of P. Unfortunately, these necessary conditions are not enough.

Problem 20. *Is the Polytope-polyhedron Problem in NP or in P?*

Here is an idea for a randomized algorithm: if we can generate a random facet of P that is reasonably uniformly distributed among all facets, then in the case $P \neq P'$, this facet would be different from the facets of P' with probability $\Omega(1/m)$, so repeating this procedure $O(m)$ times we discover that P and P' are not the same.

The following easy procedure generates a random facet F of P: Choose a v_i at random. Choose a random edge e of P incident with v_i. Choose a random 2-face of P incident with e, etc. Eventually you obtain a facet F of P.

It is easy to implement this procedure in polynomial time. Several modifications are possible: one may try to "bias" the choices on the way up so that the distribution of f should be close to uniform on all facets of P'. One may also allow steps "down" as in the Metropolis algorithm.

Problem 21. *How close to uniform is the distribution of F? What is the probability that this procedure finds a facet of P that is not a facet of P', assuming that $P \neq P'$? Is any of the modifications mentioned above better?*

The Polytope-polyhedron Problem was motivated by the following problem raised by Gottlob (1990).

BLOCKING CLUTTER PROBLEM. Given two clutters (V, E_1) and (V, E_2) on the same underlying set, are they blockers of each other?

This problem is in co-NP, but is not known to be in P or NP. Gottlob gives several equivalent problems, and conjectures that problems reducible to the Blocking Clutter Problem form a class properly between P and co-NP.

Problem 22. *Is there any reduciblity between the Polytope-polyhedron Problem and the Blocking Clutter Problem? Is there a naturally defined complexity class that explains the structural similarity between these problems?*

Problem 23. *Find a randomized polynomial-time algorithm for the Blocking Clutter Problem.*

References

F. Alizadeh (1992): Optimization over the positive semidefinite cone: interior point methods and combinatorial applications, *Proc. IPCO '92.*

N. Alon (1986): Eigenvalues and expanders, *Combinatorica* **6**, pp. 83–96.

D. Applegate and R. Kannan (1990): Sampling and integration of near log-concave functions, *Proc. 23th ACM STOC*, 156–163.

E. Balas and W. R. Pulleyblank (1983): The perfect matchable subgraph polytope of a bipartite graph, *Networks* **13**, 495-516.

E. Balas and W. R. Pulleyblank (1989), The perfect matchable subgraph polytope of an arbitrary graph, *Combinatorica* **9**, 321–337.

M. O. Ball, W. Liu, W. R. Pulleyblank (1987): Two terminal Steiner tree polyhedra, in: *Contributione to Operations Research and Economics*, MIT Press, Cambridge, MA, 251–284.

F. Barahona and A.R.Mahjoub (1987): Compositions of graphs and polyhedra II: stable sets, Res. Rep. 87464-OR, Univ. Bonn.

L. J. Billera and B. Sturmfels (1992): Fiber polytopes, *Annals of Math.* **135**, 527–549.

R. B. Boppana (1989): Optimal separations between concurrent–write parallel machines, *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, 320–326.

K. Cameron and J. Edmonds (1989): Coflow polyhedra (preprint, Univ. of Waterloo)

I. Csiszár, J. Körner, L. Lovász, K. Marton and G. Simonyi (1990): Entropy splitting for antiblocking pairs and perfect graphs, *Cominatorica* **10**, 27–40.

Y. Colin de Verdier (1986): Spectres de variétés riemannienes et spectres de graphes, *Proc. Int. Congr. Math.* Berkeley, 1986, 522–530.

Y. Colin de Verdier (1992): On a new graph invariant and a criterion for planarity (preprint)

P. Diaconis (1988) *Group representations in probability and statistics,* Inst. for Math. Statistics, Hayward, California.

P. Diaconis, D. Stroock (1991): Geometric bounds for eigenvalues of Markov chains, *Annals of Appl. Probability* **1**, 36–61.

P. Diaconis and P. Hanlon (1992): Eigen analysis for some examples of the Metropolis algorithm (preprint)

M. Dyer and A. Frieze (1991): Computing the volume of convex bodies: a case where randomness provably helps, in: *Probabilistic Combinatorics and its Applications* (ed. B. Bollobás), (to appear)

M. Dyer, A. Frieze and R. Kannan (1989): A Random Polynomial Time Algorithm for Approximating the Volume of Convex Bodies, *Proc. of the 21st Annual ACM Symposium on Theory of Computing*, 375-381.

J. Edmonds (1965): Maximum matching and a polyhedron with (0,1) vertices, *J. Res. Nat. Bur. Standards Sect. B*, 69B, 125–130.

U. Feige and L. Lovász (1992): Two-prover one-round proof systems: their power and their problems, *Proc. 24th Ann. ACM Symp. on the Theory of Computing*, 733–744.

R. Fletcher (1985): Semi-definite matrix constraints in optimization, *SIAM J. on Control and Optim.* **23**, 493–513.

G. Gottlob (1990): (personal communication)

M. Grötschel, L. Lovász and A. Schrijver (1988): *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag.

R. Holley and D. Stroock (1988): Simulated annealing via sobolev inequalities, *Comm. Math. Phys.* **115**, 553–569.

R. Holley, S. Kusuoka and D. Stroock (1989): Asymptotics pf the spectral gap with applications to the theory of simulated annealing, *J. Func. Analysis* **83**, 333–347.

N. Jacobson (1980): *Basic Algebra II*, W. H. Freeman, San Francisco.

M. R. Jerrum and A. J. Sinclair (1989), Approximating the permanent, *SIAM Journal on Computing* **18**, 1149–1178.

M. R. Jerrum, L. G. Valiant and V. V. Vazirani (1986): Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* **43**, 169–188.

D. S. Johnson (1990): Local optimization and the travelling salesman problem, in: *Proc. 17th Coll. on Automata, Languages and Programming,* Springer-Verlag, 446–461.

D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon (1989): Optimization by simulated annealing: an exxperimental evaluation; Part I, graph partitioning, *Operations Research* **37**, 865–892.

D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon (1991): Optimization by simulated annealing: an exxperimental evaluation; Part II, graph coloring and number partitioning, *Operations Research* **39**, 378–406.

J. Kahn, J. H. Kim (1992): Entropy and sorting, *Proc. 24th Ann. ACM Symp. on the Theory of Computing*, 178–187.

S. Kirckpatrick, C. D. Gelatt and M. P. Vecchl (1983): Optimization by simulated annealing, *Science* **230**, 671–680.

J. Körner (1973): Coding of an information source having ambiguous alphabet and the entropy of graphs. In: *Transactions of the 6th Prague Conference on Information Theory, etc.*, Academia, Prague, 1973, pp. 411–425.

J. Körner (1986): Fredman–Komlós bounds and information theory. *SIAM J. of Algebraic and Discrete Methods* **7**, 560–570.

J. Körner and G. Longo (1973): Two-step encoding of finite memoryless sources. *IEEE Trans. on Inform. Theory* **19**, 778–782.

J. Körner, K. Marton: New bounds for perfect hashing via information theory. *European J. of Combinatorics*, **9** (1988) 523–530.

J. Körner, G. Simonyi, Z. Tuza (1992): Perfect couples, *Combinatorica* (to appear)

N. Linial, L. Lovász, A. Wigderson (1988): Rubber bands, convex embeddings, and graph connectivity, *Combinatorica* **8**, 91–102.

L. Lovász (1979): On the Shannon capacity of graphs, *IEEE Trans. Inform. Theory* **25**, 1-7.

L. Lovász, M. Saks, A. Schrijver (1989): Orthogonal representations and connectivity of graphs, *Linear Alg. Appl.* **114/115**, 439–454.

L. Lovász and A. Schrijver (1989): Matrix cones, projection representations, and stable set polyhedra, in: *Polyhedral Combinatorics*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science I, 1–17.

L. Lovász and A. Schrijver (1990): Cones of matrices and setfunctions, and 0-1 optimization, *SIAM J. Optim.* **1**, 166–190.

L. Lovász and M. Simonovits (1990): Mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. *Proc. 31st Annual Symp. on Found. of Computer Science,* IEEE Computer Soc., 346–355.

L. Lovász and M. Simonovits (1992): Random walks in a convex body and an improved volume algorithm (preprint, Math. Inst. Hung. Acad. Sci. No. 1/1992).

N. Maculan (1987): The Steiner problem in graphs, *Annals of Discr. Math.* **31**, 185-222.

K. Marton, On the Shannon capacity of probabilistic graphs, *J. Comb. Theory*, to appear.

N. Metropolis, A. Rosenblut, M. Rosenbluth, A. Teller amd E. Teller (1953): Equation of state calculation by fast computing machines, *J. Chem. Physics* **21**, 1087–1092.

I. Newman, P. Ragde, A. Wigderson (1990): Perfect hashing, graph entropy and circuit complexity, *Proceedings of the 5th Annual Conference on Structure in Complexity Theory*, 91–100,

M. Overton (1988): On minimizing the maximum eigenvalue of a symmetric matrix, *SIAM J. on Matrix Analysis and Applications* **9**, 256–268.

W. R. Pulleyblank (1992): Lecture at the Combinatorics Conference at Waterloo

J. Radhakhrishnan (1991): Better bounds for threshold formulas, Proc. of the 32nd IEEE FOCS, 314–323,

A. Recski (1989): *Matroid Theory and its Applications in Electric Network Theory and Statics*, Algorithms and Combinatorics **4**, Springer-Verlag – Akadémiai Kiadó, Budapest.

G. Sasaki (1991), The effect of the density of states on the Metropolis algorithm, *Information Processing Letters* **37**, 159–163.

G. H. Sasaki and B. Hajek (1988), The time complexity of maximum matching by simulated annealing, *Journal of the ACM* **35**, 387–403.

H. D. Sherali and W. P. Adams (1990a): A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, *SIAM J. Discr. Math.* **3**, 411–430.

H. D. Sherali and W. P. Adams (1990b): Linearization strategies for a class of zero-one mixed integer programming problems, *Oper. Res.* **38**, 217–236.

A. Sinclair and M. Jerrum (1988): Conductance and the rapid mixing property for Markov chains: the approximation of the permanent resolved, *Proc. 20th ACM STOC*, pp. 235–244.

R.P. Stanley (1983): *Combinatorics and commutative algebra*, Birkhäuser, Boston–Basel–Stuttgart.

W.T. Tutte (1963): How to draw a graph, *Proc. London Math. Soc.* **13**, 743-768.

N. White and W. Whiteley (1983): The algebraic geometry of stresses in frameworks, *SIAM J. Alg. Discr. Methods* **4**, 481–511.

M. Yannakakis (1988): Expressing combinatorial optimization problems by linear programs, *Proc. 29th IEEE FOCS*, 223-228.