OPTIMA 97

Mathematical Optimization Society Newsletter

Note from the Editors

Dear MOS members,
this year, we are celebrating the 50th anniversary of Jack Edmonds' seminal papers *Paths, trees, and flowers* [3] and *Maximum matching and a polyhedron with 0-1-vertices* [1], both published in 1965. These papers not only solved the matching problem both from an algorithmic as well as from a polyhedral point of view and laid out the plan for the field of Polyhedral Combinatorics. They also put on the mathematicians' agenda the question for the existence of efficient algorithms ("It is by no means obvious whether or not there exists an algorithm whose difficulty increases only algebraically with the size of the graph" [3]) and pointed out that "(...) in applying linear programming to a combinatorial problem, the number of relevant inequalities is not important but their combinatorial structure is" [1].

Just in time for the 50th anniversary, Thomas Rothvoss [5] recently constructed a brilliant proof demonstrating that the perfect matching polytopes of complete graphs do not admit polynomial size extended formulations. This settles a major open question raised by Mihalis Yannakakis more than two decades ago and extends Edmonds' conclusion cited above by showing that the existence of a polynomial (*algebraic)*) time algorithm for a problem does not imply the possibility of a polynomial size linear representation of the associated polytope. In this issue of our newsletter, Thomas explains his result and its proof in a way that is accessible for the general readership, informative for experts, and very enjoyable for all.

Furthermore, we are very glad to have an interview with Jack Edmonds, in which he talks about revolting against exponential time algorithms, reveals what a glimpse of heaven may be, and reminisces about the birth of the complexity classes P, NP, and coNP as well as the notorious conjectures on their relations that have been formalized following his fundamental contributions as expressed in statements like:

> The good characterization will describe certain information about the matrix which the supervisor can require his assistant to search out along with a minimum partition and which the supervisor can then use 'with ease' to verify with mathematical certainty that the partition is indeed minimum. Having a good characterization does not mean necessarily that there is a good algorithm. [2]
>
> I conjecture that there is no good algorithm for the traveling salesman problem. My reasons are the same as for any mathematical conjecture: (1) It is a legitimate mathematical possibility, and (2) I do not know. [4]

We are sure that reading what Jack has to say 50 years after he started it all will not only be a great pleasure for our readership, but also a very good source of inspiration.

Along with these two scientific highlights, the issue contains a note from Dan Bienstock, the new editor-in-chief of *Mathematical Programming Computation (MPC)*, that explains innovations he and his editorial team are planning to implement for the journal. Finally, the issue has calls for nominations for the upcoming elections within our own society and for the INFORMS John von Neumann Theory Prize, announcements of a summer school (with Jack Edmonds among the lecturers) on Polyhedral Combinatorics preceding the upcoming ISMP congress in Pittsburgh, and calls for papers for several special issues of *Mathematical Programing, Ser. B.*

Sam Burer, Co-Editor
Volker Kaibel, Editor
Jeff Linderoth, Co-Editor

References

[1] J. Edmonds. Maximum matching and a polyhedron with 0, 1-vertices. *J. Res. Nat. Bur. Standards Sect. B*, 69B:125–130, 1965.
[2] J. Edmonds. Minimum partition of a matroid into independent subsets. *J. Res. Nat. Bur. Standards Sect. B*, 69B:67–72, 1965.
[3] J. Edmonds. Paths, trees, and flowers. *Canad. J. Math.*, 17:449–467, 1965.
[4] J. Edmonds. Optimum branchings. *J. Res. Nat. Bur. Standards Sect. B*, 71B:233–240, 1967.
[5] T. Rothvoß. The matching polytope has exponential extension complexity. *CoRR*, abs/1311.2369, 2013.

Contents of Issue 97 / April 2015

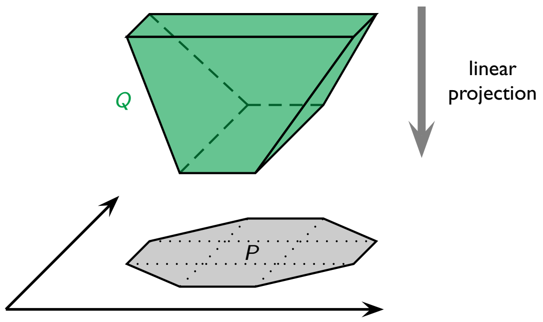
Thomas Rothvoss

The matching polytope has exponential extension complexity

1 Introduction

Linear programs are at the heart of combinatorial optimization as they allow to model a large class of polynomial time solvable problems such as flows, matchings and matroids. The concept of LP duality in many cases leads to structural insights that in turn leads to specialized polynomial time algorithms. In practice, general LP solvers turn out to be very competitive for many problems, even in cases in which specialized algorithms have the better theoretical running time. Hence it is particularly interesting to model problems with as few linear constraints as possible. For example if we consider the convex hull P_{ST} of the characteristic vectors of all spanning trees in a complete n-node graph, then this polytope has $2^{\Omega(n)}$ many facets [7]. However, one can write $P_{ST} = \{x \mid \exists y : (x, y) \in Q\}$ with a higher dimensional polytope Q with only $O(n^3)$ many inequalities [16]. Hence, instead of optimizing a linear function over P_{ST}, one can optimize over Q. In fact, Q is called a *linear extension* of P_{ST} and the minimum number of facets of any linear extension is called the *extension complexity* and it is denoted by $\mathrm{xc}(P_{ST})$; in this case $\mathrm{xc}(P_{ST}) \leq O(n^3)$. If $\mathrm{xc}(P)$ is bounded by a polynomial in n, then we say that $P \subseteq \mathbb{R}^n$ has a *compact formulation*.

Let us verify that this makes sense: in the example below, we have a 2-dimensional polygon P with 8 facets which is represented as a projection of a 3-dimensional polytope Q that has only 6 facets.



Other examples of non-trivial compact formulations contain the permutahedron [12], the parity polytope, the matching polytope in planar graphs [3] and more generally the matching polytope in graphs with bounded genus [11].

A natural question that emerges is which polytopes do *not* admit a compact formulation. The first progress was made by Yannakakis [23] who showed that any *symmetric* extended formulation for the matching polytope and the TSP polytope must have exponential size. Conveniently, this allowed to reject a sequence of flawed **P** = **NP** proofs, which claimed to have (complicated) polynomial size LPs for TSP. It was not clear a priori whether the symmetry condition would be essential, but Kaibel, Pashkovich and Theis [14] showed that for the convex hull of all $\log n$-size matchings, there is a compact asymmetric formulation, but no symmetric one.

Then the major breakthrough by Fiorini, Massar, Pokutta, Tiwary and de Wolf [10] showed unconditionally that several well studied polytopes, including the correlation polytope and the TSP polytope, have exponential extension complexity. More precisely, they show that the *rectangle covering lower bound* [23] for the correlation polytope is exponential, for which they use known tools from communication complexity such as Razborov's *rectangle corruption lemma* [19].

One insight that appeared already in [23, 10] is that if a "hard" polytope P is the linear projection of a face of another polytope P', then $\mathrm{xc}(P') \geq \mathrm{xc}(P)$. This way, the "hardness" of the correlation polytope can be translated to many other polytopes using a reduction (in fact, in many cases, the usual **NP**-hardness reduction can be used); see [18, 1] for some examples.

A completely independent line of research was given by Chan, Lee, Raghavendra and Steurer [5] who use techniques from Fourier analysis to show that for constraint satisfaction problems, known integrality gaps for the Sherali-Adams LP translate to lower bounds for any LPs of a certain size. For example they show that no LP of size $n^{O(\log n/\log\log n)}$ can approximate MaxCut better than $2-\varepsilon$. This is particularly interesting as in contrast the gap of the SDP relaxation is around 1.13 [13].

1.1 The matching polytope

However, all those polytopes model **NP**-hard problems and naturally, no complete description of their facets is known (and no efficiently separable description is possible if **NP** $\neq$ **P**). So what about nicely structured combinatorial polytopes that admit polynomial time algorithms to optimize linear functions? The most prominent example here is the *perfect matching polytope* P_{PM}, which is the convex hull of all characteristic vectors of perfect matchings in a complete n-node graph $G = (V, E)$. This year we can celebrate the 50th anniversary of the paper of Edmonds [6] which shows that apart from requiring non-negativity, the degree-constraints plus the *odd-set inequalities* are enough for an inequality description. In other words, we can write

$$P_{PM} = \mathrm{conv}\{\chi_M \in \mathbb{R}^E \mid M \subseteq E \text{ is a perfect matching}\}$$
$$= \left\{ x \in \mathbb{R}^E \mid \begin{array}{rcll} x(\delta(v)) & = & 1 & \forall v \in V \\ x(\delta(U)) & \geq & 1 & \forall U \subseteq V : |U| \text{ odd} \\ x_e & \geq & 0 & \forall e \in E \end{array} \right\}.$$

Here, χ_M is the characteristic vector of M. Note that there are only n degree constraints and $O(n^2)$ non-negativity constraints, but $2^{\Omega(n)}$ odd set inequalities. Any linear function can be optimized over P_{PM} in strongly polynomial time using Edmonds algorithm [6]. Moreover, given any point $x \notin P_{PM}$, a violating inequality can be found in polynomial time via the equivalence of optimization and separation or using Gomory-Hu trees, see Padberg and Rao [17]. There are compact formulations for P_{PM} for special graph classes [11] and every *active cone* of P_{PM} admits a compact formulation [22]. Moreover, the best known upper bound on the extension complexity in general graphs is $\mathrm{poly}(n) \cdot 2^{n/2}$ [8], which follows from the fact that $\mathrm{poly}(n) \cdot 2^{n/2}$ many randomly taken complete bipartite graphs cover all matchings and that the convex hull of the union of polytopes can be described with a few extra inequalities [2]. For a detailed discussion of the matching polytope we refer to the book of Schrijver [21].

1.2 Our contribution

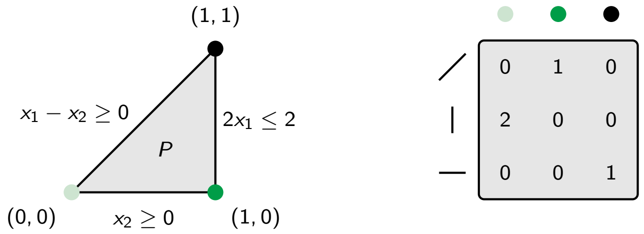
In this article, we want to discuss the following somewhat surprising theorem and its proof:

Theorem 1. *For all even n, the extension complexity of the perfect matching polytope in the complete n-node graph is $2^{\Omega(n)}$.*

Recall that the perfect matching polytope is a face of the matching polytope itself, hence the bound also holds for the convex hull of all (not necessarily perfect) matchings.

2 Our approach

Formally, the *extension complexity* $\mathrm{xc}(P)$ is the smallest number of facets of a (higher-dimensional) polyhedron Q such that there is a linear projection π with $\pi(Q) = P$. This definition seems to ignore the dimension, but one can always eliminate a non-trivial lineality space from Q and make Q full-dimensional, and then the dimension of Q is bounded by the number of inequalities anyway. Before we continue our discussion of the matching polytope, consider a general polytope P and let $x_1, \dots, x_v$ be a list of its vertices. Moreover, let $P = \mathrm{conv}\{x_1, \dots, x_v\} = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ be any inequality description, say with f inequalities. A crucial concept in extended formulations is the *slack matrix* $S \in \mathbb{R}^{f \times v}_{\geq 0}$ which is defined by $S_{ij} = b_i - A_i x_j$, where A_i is the ith row of A. A small example is as follows:



polytope slack matrix

Moreover, the *non-negative rank* of a matrix is

$$\text{rk}_+(S) = \min\{r \mid \exists U \in \mathbb{R}^{f \times r}_{\geq 0}, V \in \mathbb{R}^{r \times v}_{\geq 0} : S = UV\}.$$

Recall that if the non-negativity condition is dropped, we recover the usual rank from linear algebra.

2.1 Yannakakis' Factorization Theorem

The connection between extension complexity and non-negative rank is expressed by the following Theorem:

Theorem 2 (Yannakakis [23]). *Let P be a polytope*[1] *with vertices $\{x_1, \dots, x_v\}$, inequality description $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ and corresponding slack matrix S. Then $xc(P) = rk_+(S)$.*

Proof. Let A be the matrix consisting of rows $A_1, \dots, A_f$. We begin with showing that $r := \text{rk}_+(S) \Rightarrow \text{xc}(P) \leq r$.

So, suppose that we have a non-negative factorization $S = UV$ with $U \in \mathbb{R}^{f \times r}_{\geq 0}$ and $V \in \mathbb{R}^{r \times v}_{\geq 0}$. We claim that $Q := \{(x, y) \in \mathbb{R}^{n+r} \mid Ax + Uy = b;\ y \geq \mathbf{0}\}$ is a linear extension and the projection π with $\pi(x, y) = x$ satisfies that $\pi(Q) = P$; in other words, we claim that $P = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^r_{\geq 0} : Ax + Uy = b\}$. To see this, take a vertex x_j of P, then we can choose the witness $y := V^j$ and have $(x_j, y) \in Q$ as $A_i x_j + U_i V^j = A_i x_j + S_{ij} = b_i$. On the other hand, if $x \notin P$, then there is some constraint i with $A_i x > b_i$ and no matter what $y \geq \mathbf{0}$ is chosen, we always have $A_i x + U_i y \geq A_i x > b_i$.

For the second part, we have to prove that $r := \text{xc}(P) \Rightarrow \text{rk}_+(S) \leq r$. Hence, suppose that we have a linear extension $Q = \{(x, y) \mid Bx + Cy \leq d\}$ with r inequalities and a linear projection π so that $\pi(Q) = P$, see also Figure 1. After a linear transformation, we may assume that $\pi(x, y) = x$, that means π is just the projection on the x-variables. We need to come up with vectors $u_i, v_j \in \mathbb{R}^r_{\geq 0}$ so that for each constraint i and each vertex x_j one has $\langle u_i, v_j \rangle = S_{ij}$. For each point x_j, fix a *lift* $(x_j, y_j) \in Q$ and choose $v_j := d - Bx_j - Cy_j \in \mathbb{R}^r_{\geq 0}$ as the vector of slacks that the lift has w.r.t. Q.

By LP duality we know that each constraint $A_i x + \mathbf{0}y \leq b_i$ can be derived as a conic combination of the system $Bx + Cy \leq d$. In other words, there is a vector $u_i \in \mathbb{R}^r_{\geq 0}$ so that

$$u_i \begin{pmatrix} B \\ C \\ d \end{pmatrix} = \begin{pmatrix} A_i \\ \mathbf{0} \\ b_i \end{pmatrix}$$

Now multiplying gives that

$$\begin{aligned} \langle u_i, v_j \rangle &= \langle u_i, d - Bx_j - Cy_j \rangle \\ &= \underbrace{u_i d}_{=b_i} - \underbrace{u_i B}_{=A_i} x_j - \underbrace{u_i C}_{=\mathbf{0}} y_j = b_i - A_i x_j = S_{ij}. \end{aligned}$$

□

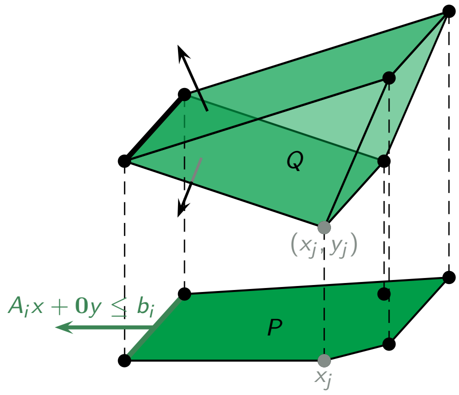


Figure 1. *Visualization of Yannakakis' Theorem*

2.2 The rectangle covering lower bound

A potential way of lower bounding $\text{rk}_+(S)$ was already pointed out in the classical paper of Yannakakis and is known as *rectangle covering lower bound*:

Lemma 3. *For any matrix S, $rk_+(S)$ is at least as large as the number of rectangles needed to cover exactly the support of S.*

The reason is simple: suppose we do have a factorization of $S = UV$ with non-negative matrices U and V. Then if we take the positive entries in the ith column of U and the positive entries in the ith row of V, then those induce a *combinatorial rectangle* where S will have strictly positive entries. Moreover, if U has only r columns, then this provides r rectangles and each entry in S has to be in at least one of them.

In the made-up example below, we can see a non-negative factorization $S = UV$ and the rectangle that is induced by the 1st row and column.

V

0	0	2	1	0
0	2	2	0	3

U

3	2
1	1
0	2
0	0
2	0

S

0	4	10	3	5
0	2	4	1	3
0	4	4	0	6
0	0	0	0	0
0	0	4	2	0

In fact, Fiorini et al. [10] show that the number of rectangles necessary for such a covering of the slack-matrix of the correlation polytope is exponential, which in turn lower bounds the extension complexity.

So, let us discuss the situation for the perfect matching polytope. Since the number of degree constraints and non-negativity inequalities is polynomial anyway, we consider the part of the slack matrix that is induced by the odd set inequalities. In other words, we consider the matrix S with

$$S_{UM} = |M \cap \delta(U)| - 1 \forall M \subseteq E \text{ perfect matching}$$
$$\forall U \subseteq V : |U| \text{ odd.}$$

The first natural approach would be to check whether the rectangle covering lower bound is super-polynomial. Unfortunately, this is not the case, as was already observed in [23]. To see this, take any pair $e_1, e_2 \in E$ of non-adjacent edges and choose

$$\mathcal{M}_{e_1,e_2} := \{M \mid e_1, e_2 \in M\} \quad \text{and} \quad \mathcal{U}_{e_1,e_2} := \{U \mid e_1, e_2 \in \delta(U)\},$$

then we obtain $O(n^4)$ many rectangles of the form $\mathcal{U}_{e_1,e_2} \times \mathcal{M}_{e_1,e_2}$. First of all, we have $S_{UM} \geq |\{e_1, e_2\}| - 1 \geq 1$ for each $U \in \mathcal{U}_{e_1,e_2}$ and $M \in \mathcal{M}_{e_1,e_2}$, hence the rectangles contain only entries (U, M) that have positive slack. But every entry (U, M) with $S_{UM} \geq 1$ is also contained in at least one such rectangle. To be precise, if $S_{UM} = k$ and $\delta(U) \cap M = \{e_1, \dots, e_{k+1}\}$, then the entry (U, M) lies in $\binom{k+1}{2}$ rectangles. So the approach with the rectangle covering bound does not work.

On the other hand, considering the rectangle covering as a sum of $O(n^4)$ many $0/1$ rank-1 matrices also does not provide a valid non-negative factorization of S. The reason is that an entry with $S_{UM} = k$ is contained in $\Theta(k^2)$ many rectangles instead of just k many, thus entries with large slack are *over-covered*. Moreover, we see no way of rescaling the rectangles in order to fix the problem. This raises the naive question: *Maybe every covering of S with polynomially many rectangles must over-cover entries with large slack*? Surprisingly, it turns out that the answer is "*yes*"!

2.3 The hyperplane separation bound

To make this more formal, we will use the *hyperplane separation lower bound* suggested by Fiorini [9]. For matrices $S, W \in \mathbb{R}^{f \times v}$, we will write $\langle S, W \rangle = \sum_{i=1}^{f} \sum_{j=1}^{v} W_{ij} \cdot S_{ij}$ as their *Frobenius inner product*. Intuitively, the hyperplane separation bound says that if we can find a linear function W that gives a large value for the slack-matrix S, but only small values on any rectangle, then the extension complexity is large.

Lemma 4 (Hyperplane separation lower bound [9]). *Let $S \in \mathbb{R}^{f \times v}_{\geq 0}$ be the slack-matrix of any polytope P and let $W \in \mathbb{R}^{f \times v}$ be any matrix. Then*

$$\mathrm{xc}(P) \geq \frac{\langle W, S \rangle}{\|S\|_\infty \cdot \alpha}$$

with $\alpha := \max\{\langle W, R \rangle \mid R \in \{0,1\}^{f \times v}$ rank-1 matrix$\}$.

Proof. First, note that the assumption provides that even for any *fractional* rank-1 matrix $R \in [0,1]^{f \times v}$ one has $\langle W, R \rangle \leq \alpha$. To see this, take an arbitrary rank-1 matrix $R \in [0,1]^{f \times v}$ and write it as $R = uv^T$ with vectors u and v. After scaling one can assume that $u \in [0,1]^f$ and $v \in [0,1]^v$. If we now take independent random $0/1$ vectors x, y so that $\Pr[x_i = 1] = u_i$ and $\Pr[y_j = 1] = v_j$, then $\mathbb{E}[x_i y_j] = u_i v_j$ and hence $\mathbb{E}[xy^T] = R$. Hence R lies in the convex hull of all $0/1$ matrices of rank-1.

Now abbreviate $r = \mathrm{xc}(P) = \mathrm{rk}_+(S)$, then there are r rank-1 matrices $R_1, \dots, R_r$ with $S = \sum_{i=1}^{r} R_i$. We obtain

$$\langle W, S \rangle = \sum_{i=1}^{r} \|R_i\|_\infty \cdot \underbrace{\left\langle W, \frac{R_i}{\|R_i\|_\infty} \right\rangle}_{\leq \alpha} \leq \alpha \cdot \sum_{i=1}^{r} \underbrace{\|R_i\|_\infty}_{\leq \|S\|_\infty} \leq \alpha \cdot r \cdot \|S\|_\infty.$$

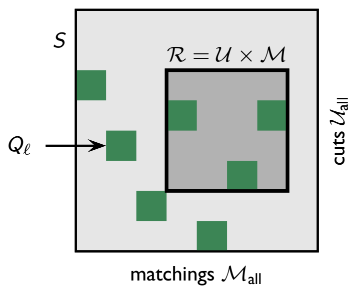
Rearranging gives the claim. □

Now, let us go back to the matching polytope and see how we can make use of this bound. Let $k \geq 3$ be an odd integer constant that we choose later. We consider only complete graphs $G = (V, E)$ that have $|V| = n = 3m(k-3) + 2k$ many vertices, for some odd integer m. In particular, $m = \Theta(n)$.

We fix the set $\mathcal{M}_{\mathrm{all}} := \{M \subseteq E \mid M \text{ is a perfect matching}\}$ of all perfect matchings in G. In contrast, we will only consider cuts that have all the same size t, where $t = \Theta(n)$ will be odd. Formally, we define $t := \frac{m+1}{2}(k-3)+3$, and abbreviate $\mathcal{U}_{\mathrm{all}} := \{U \subseteq V \mid |U| = t\}$ as all t-node cuts in G. Let

$$Q_\ell := \{(U, M) \in \mathcal{U}_{\mathrm{all}} \times \mathcal{M}_{\mathrm{all}} \mid |\delta(U) \cap M| = \ell\}$$

be the set of pairs of cuts and matchings intersecting in ℓ edges and let μ_ℓ be the *uniform measure* on Q_ℓ. In other words, each (U, M) with $|\delta(U) \cap M| = \ell$ carries the same probability measure of $\frac{1}{|Q_\ell|}$. In the following, a *rectangle* is of the form $\mathcal{R} = \mathcal{U} \times \mathcal{M}$ with $\mathcal{M} \subseteq \mathcal{M}_{\mathrm{all}}$ and $\mathcal{U} \subseteq \mathcal{U}_{\mathrm{all}}$. Note that for parity reasons $\mu_{2i}(\mathcal{R}) = 0$ for all $i \in \mathbb{Z}_{\geq 0}$. One could try a schematic picture of the situation which is as follows:



Now we want to choose a matrix $W \in \mathbb{R}^{\mathcal{U}_{\mathrm{all}} \times \mathcal{M}_{\mathrm{all}}}$ for which the hyperplane separation bound provides an exponential lower bound. We choose

$$W_{U,M} = \begin{cases} -\infty & |\delta(U) \cap M| = 1, \\ \frac{1}{|Q_3|} & |\delta(U) \cap M| = 3, \\ -\frac{1}{k-1} \cdot \frac{1}{|Q_k|} & |\delta(U) \cap M| = k, \\ 0 & \text{otherwise}. \end{cases}$$

The intuition is that we reward a rectangle for covering an entry in Q_3, punish it for covering entries in Q_k and completely forbid to cover any entry in Q_1. First, it is not difficult to see that

$$\langle W, S \rangle = 0 + (3-1) \cdot |Q_3| \cdot \frac{1}{|Q_3|} - (k-1) \cdot |Q_k| \cdot \frac{1}{k-1} \cdot \frac{1}{|Q_k|} = 1. \quad (1)$$

Our hope is that any large rectangle $\mathcal{R}$ must over-cover entries with $|\delta(U) \cap M| = k$ and hence $\langle W, \mathcal{R} \rangle$ is small. In fact, we can prove

Lemma 5. *For any large enough odd constant k ($k := 501$ suffices) and any rectangle $\mathcal{R}$ with $\mathcal{R} = \mathcal{U} \times \mathcal{M}$ with $\mathcal{U} \subseteq \mathcal{U}_{\mathrm{all}}$ and $\mathcal{M} \subseteq \mathcal{M}_{\mathrm{all}}$ one has $\langle W, \mathcal{R} \rangle \leq 2^{-\delta \cdot n}$ where $\delta := \delta(k) > 0$ is a constant.*

The proof of this lemma takes the most part of the work. From the technical point of view, our proof is a substantial modification of Razborov's original rectangle corruption lemma [19].

Assuming the bound from Lemma 5 we can then apply Lemma 4, and infer that the perfect matching polytope satisfies

$$\mathrm{xc}(P_{PM}) \geq \frac{\langle W, S \rangle}{\|S\|_\infty \cdot \max\{\langle W, \mathcal{R} \rangle \mid \mathcal{R} \text{ rectangle}\}}$$

$$\geq \frac{1}{n \cdot 2^{-\delta n}} \geq 2^{\Omega(n)}.$$

Here we use that $\langle W, S \rangle = 1$, $\|S\|_\infty \leq n$ and that $\langle W, \mathcal{R} \rangle \leq 2^{-\delta n}$ for all rectangles $\mathcal{R}$.

3 The pseudo-random behavior of large sets

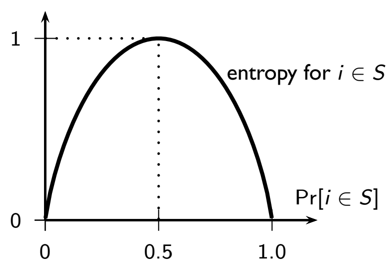
Before we go on with the discussion of the matching polytope, we want to discuss a crucial tool for analyzing the behauvior of combinatorial rectangles. We want to keep things general, so we will not talk about cuts and matchings in this section. Instead we will show that the following claim is true: Suppose you have a set family $Y \subseteq 2^{[m]}$ with $|Y| \geq 2^{(1-\varepsilon)m}$ for a small enough constant $\varepsilon > 0$. Then 99 % of indices $i \in \{1, \dots, m\}$ will be in 50 % ± 1 % of sets in the family Y.

We can prove such claims using an *entropy counting argument*. Recall that for a random variable y over $\{1, \dots, k\}$, the *entropy* is defined by $H(y) := \sum_{j=1}^{k} \Pr[y=j] \cdot \log_2 \frac{1}{\Pr[y=j]}$. Moreover, the entropy is maximized if y is the *uniform distribution*; in that case we have $H(y) = \log_2(k)$. A useful property is that entropy is *subadditive*. For example if $y = (y_1, \dots, y_m)$ is a random vector, then $H(y) \leq \sum_{i=1}^{m} H(y_i)$. In the following, if we write $y \sim Y$, then y is a uniformly drawn random element from Y. If Y is a set family over ground-set $1, \dots, m$, then we say that index $i \in [m]$ is *ε-unbiased* if

$$\frac{1}{2} \cdot (1-\varepsilon) \leq \Pr_{S \sim Y}[i \in S] \leq \frac{1}{2} \cdot (1+\varepsilon).$$

Lemma 6. *Let $Y \subseteq 2^{[m]}$ be a set family with $|Y| \geq 2^{(1-\Theta(\varepsilon^3))m}$. Then at least a $1-\varepsilon$ fraction of indices $i \in \{1, \dots, m\}$ is ε-unbiased.*

Proof. Let us choose a uniform random set $S \in Y$ from our family. First, fix an index i and consider the binary random variable "$i \in S$". If i is perfectly unbiased, then $H(i \in S) = \frac{1}{2}$. On the other hand, if i is ε-biased, then it is not hard to see that the entropy will be a little less than $\frac{1}{2}$, simply because the function $p \log_2(p) + (1-p) \log_2 \frac{1}{1-p}$ has a unique maximum at $p = \frac{1}{2}$:



In fact, one can see that in this case $H(i \in S) \leq \frac{1}{2} - \Theta(\varepsilon^2)$.

Now, we assume for the sake of contradiction that there are more than εm indices that are ε-biased. Then we can bound the entropy of the random set S by

$$\log_2(|Y|) = H(S) \leq \sum_{i=1}^{m} H(i \in S) < m - \varepsilon m \cdot \Theta(\varepsilon^2)$$

using the subadditivity of H. Rearranging yields $|Y| < 2^{(1-\Theta(\varepsilon^3))m}$. □

The reason why unbiased indices are useful is the following observation that follows from Bayes' Theorem:

Corollary 7. *Let Y be a family and i be an ε-unbiased index. Then*

$$\Pr_{S \subseteq [m]}[S \in Y \mid i \in S] \in (1 \pm \varepsilon) \cdot \Pr_{S \subseteq [m]}[S \in Y]$$

where S is a uniform random set that contains each element j independently with probability $\frac{1}{2}$.

For example, in our application, if we want to test the "density" of the set of cuts $\mathcal{U}$, then for most nodes $i \in V$, the tests $\Pr_U[U \in \mathcal{U}]$ and $\Pr_U[U \in \mathcal{U} \mid i \in U]$ will yield roughly the same probability (one can imagine that U here is a uniform random cut).

4 The quadratic measure increase

In this section, we provide the proof of the main technical ingredient, Lemma 5. Formally, we will prove the following statement:

Lemma 8. *For each odd $k \geq 3$ and for any rectangle $\mathcal{R}$ with $\mu_1(\mathcal{R}) = 0$, one has $\mu_3(\mathcal{R}) \leq \frac{400}{k^2} \cdot \mu_k(\mathcal{R}) + 2^{-\delta m}$ where $\delta := \delta(k) > 0$.*

We verify that this indeed implies Lemma 5. Consider a rectangle $\mathcal{R}$ and assume that $\mu_1(\mathcal{R}) = 0$ since otherwise $\langle W, \mathcal{R} \rangle = -\infty$. Then

$$\langle W, \mathcal{R} \rangle = \mu_3(\mathcal{R}) - \frac{1}{k-1} \mu_k(\mathcal{R}) \stackrel{\text{Lem. 8}}{\leq} \underbrace{\left(\frac{400}{k^2} - \frac{1}{k-1}\right)}_{\leq 0} \mu_k(\mathcal{R}) + 2^{-\delta m} \leq 2^{-\delta m}$$

where we choose k as a large enough constant (e.g., $k = 501$).

4.1 The concept of partitions

For the remainder of this work, we fix a rectangle $\mathcal{R} = \mathcal{U} \times \mathcal{M}$ with $\mu_1(\mathcal{R}) = 0$. If we want to compare the fractions $\mu_3(\mathcal{R})$ and $\mu_k(\mathcal{R})$, then we should start with answering the following question: how does one actually sample from Q_3 or Q_k? We cannot just independently sample a cut and a matching because most likely those would intersect in $\Theta(n)$ edges. The trick is to sample pairs $(U, M) \sim Q_3$ and $(U, M) \sim Q_k$ in two stages. In the first stage, we partition the graph into a certain block structure that is depicted in Figure 2. In the second phase we then sample U and M w.r.t. those blocks; in particular there will only be few edges where U and M might intersect.

Formally, a *partition* is a tuple $T = (A = A_1 \dot\cup \dots \dot\cup A_m, C, D, B = B_1 \dot\cup \dots \dot\cup B_m)$ with $V = A \dot\cup C \dot\cup D \dot\cup B$ and the following properties:

- $A \subseteq V$ is a set of $|A| = m(k-3)$ nodes that is partitioned into blocks $A = A_1 \dot\cup \dots \dot\cup A_m$ with $|A_i| = k-3$ nodes each.
- $C \subseteq V$ is a set of k nodes.
- $D \subseteq V$ is a set of k nodes.
- $B = B_1 \dot\cup \dots \dot\cup B_m$ with $B \subseteq V$ is a partition of the remaining nodes so that $|B_i| = 2(k-3)$.

For a node-set U, let $E(U) := \{(u, v) \in E \mid u, v \in U\}$ be the edges lying inside of U. We abbreviate $E(T) := \bigcup_{i=1}^{m} E(A_i) \cup E(C \cup D) \cup \bigcup_{i=1}^{m} E(B_i)$ as the edges associated with the partition T, see again Figure 2.

We say that the matchings

$$\mathcal{M}(T) := \{M \in \mathcal{M} \mid M \subseteq E(T)\}$$

respect the partition T. In other words, the matchings in $\mathcal{M}(T)$ have only edges running inside A_i or B_i or inside $C \cup D$. Similarly, we say that the cuts

$$\mathcal{U}(T) := \{U \in \mathcal{U} \mid U \subseteq A \cup C \text{ with } |U \cap A_i| \in \{0, |A_i|\} \ \forall i \in [m]\}$$

respect the partition (see Figure 3). In other words, those cuts contain either all or none of the nodes in each A_i. Let $\mathcal{M}_{\text{all}}(T) := \{M \in \mathcal{M}_{\text{all}} \mid M \subseteq E(T)\}$ and $\mathcal{U}_{\text{all}}(T) := \{U \in \mathcal{U}_{\text{all}} \mid U \subseteq A \cup C;\ |U \cap A_i| \in \{0, |A_i|\}\ \forall i \in [m]\}$ be the supersets of $\mathcal{M}(T)$ and $\mathcal{U}(T)$ containing all possible matchings and cuts that would respect the partition. The advantage of such a partition is that if we take a random matching $M \sim \mathcal{M}_{\text{all}}(T)$ and a random cut $U \sim \mathcal{U}_{\text{all}}(T)$, then the intersection $\delta(U) \cap M$ can only contain edges in $E(C \cup D)$ (in fact, it contains an odd number between 1 and k edges).

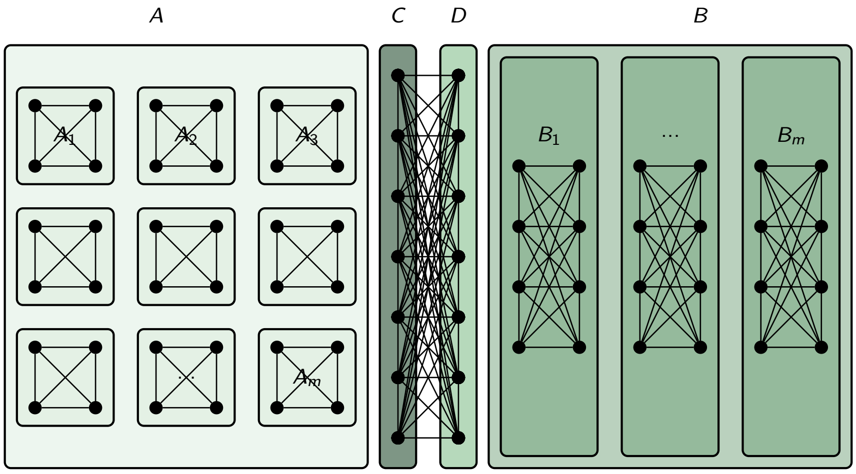


Figure 2. *Visualization of a partition T with all edges $E(T)$*

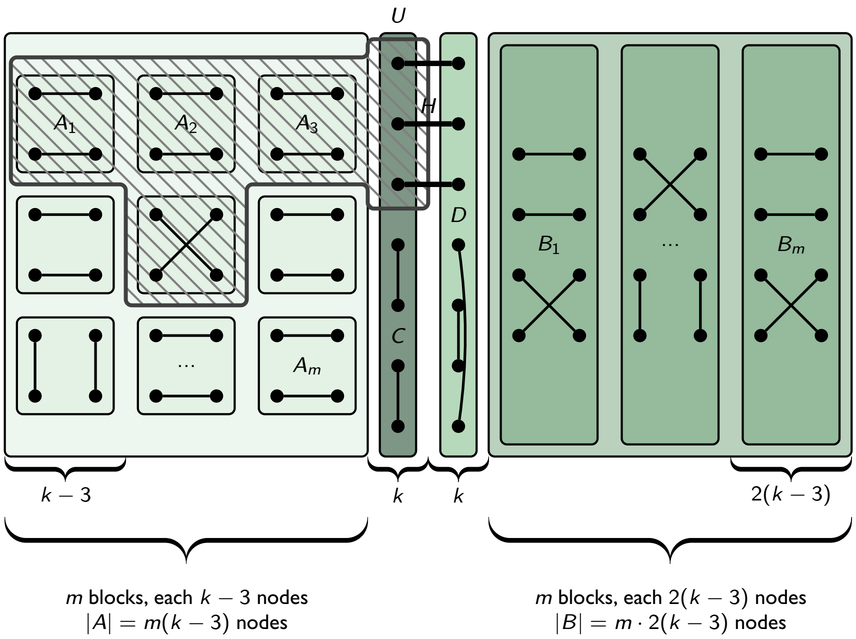


Figure 3. *Visualization of a partition T together with one matching $M \in \mathcal{M}(T)$ and one cut $U \in \mathcal{U}(T)$*

4.2 Generating the distributions μ_3 and μ_k

The key trick is that the measures $\mu_3(\mathcal{R})$ and $\mu_k(\mathcal{R})$ can be nicely compared for the rectangles that are induced by partitions. To fix some notation, we say that H is an *ℓ-matching* if H is a matching with exactly ℓ edges. The nodes incident to edges H are denoted by $V(H)$.

For a matching $H \subseteq E(C \cup D)$, we define

$$p_{M,T}(H) := \Pr_{M \in \mathcal{M}_{\text{all}}(T)}[M \in \mathcal{M} \mid H \subseteq M]$$

as the chance that a random matching from this block structure lies in our rectangle. For $c \subseteq C$, let

$$p_{U,T}(c) := \Pr_{U \in \mathcal{U}_{\text{all}}(T)}[U \in \mathcal{U} \mid U \cap C = c]$$

be the chance that a random extension of c to a cut U lies in the rectangle. By a slight abuse of notation, we denote $p_{U,T}(H) := p_{U,T}(V(H) \cap C)$ for a matching $H \subseteq C \times D$. We should remark that we only consider cuts U of size $|U| = t = \frac{m+1}{2}(k-3)+3$, that means $|U|-3$ is a multiple of $k-3$, and hence $p_{U,T}(c) > 0$ only if $|c| \in \{3,k\}$.

Now we can use the block structure to generate entries from Q_3 and Q_k:

- *Generating a uniform random entry $(U,M) \in Q_k$*: Pick a random partition T. Pick a random k-matching F in the bipartite graph $C \times D$ and randomly extend F. Hence

$$\mu_k(\mathcal{R}) = \mathop{\mathbb{E}}_{T}\Big[\mathop{\mathbb{E}}_{|F|=k}[p_{M,T}(F) \cdot p_{U,T}(F)]\Big]$$

- *Generating a uniform random entry $(U,M) \in Q_3$*: Pick a random partition T. Pick a random k-matching $F \subseteq C \times D$. Pick 3 edges H out of F and extend H:

$$\mu_3(\mathcal{R}) = \mathop{\mathbb{E}}_{T}\Big[\mathop{\mathbb{E}}_{|F|=k}\Big[\mathop{\mathbb{E}}_{H \in \binom{F}{3}}[p_{M,T}(H) \cdot p_{U,T}(H)]\Big]\Big]$$

4.3 The notion of good partitions

Now is the time to remember our insights from Section 3. We can expect that for most partitions T with k edges F certain local conditioning will not change the outcome of a “density test”. Extending those insights that we discussed only for set families to cuts and matchings, here is the property that we would expect for a $1-\varepsilon$ fraction of partitions:

Definition 1. A pair (T, F) with a partition T and a k-matching F is called *good* if for all $H \in \binom{F}{3}$ at least one of the conditions is satisfied:

- *smallness*: One has $p_{M,T}(H) \leq 2^{-\delta m}$ or $p_{U,T}(H) \leq 2^{-\delta m}$ (or both).
- *unbiasedness*: One has

$$p_{U,T}(H) = (1 \pm \varepsilon) \cdot p_{U,T}(F) > 0.$$

Moreover, for every matching $H' : H \subseteq H' \subseteq E(C \cup D)$ one has

$$p_{M,T}(H) = (1 \pm \varepsilon) \cdot p_{M,T}(H') > 0.$$

Let $\texttt{GOOD}(T, F)$ be the indicator variable telling whether the pair (T, F) is good. In Section 4.5, we will give an argument, how one can apply Theorem 6 to derive that a $1 - \varepsilon$ fraction of pairs will be good.

But for now, we want to get to the proof that the contribution of those "good" pairs cannot be large. This contains the key arguments, *why* the matching polytope has no compact LP representation. It is also the part where we make use of the combinatorial properties of matchings and cuts.

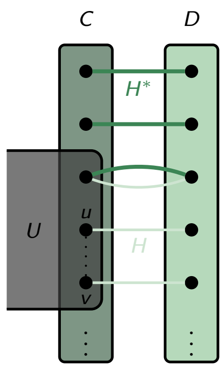
4.4 Contribution of good partitions

Recall that the definition of a good pair (T, F) means that for each triple $H \in \binom{F}{3}$ we have either *smallness* or *unbiasedness* satisfied. Let us use $\texttt{SMALL}(T, F, H)$ and $\texttt{UNBIASED}(T, F, H)$ as indicator variables and we omit T, F if they are clear from the context. Now, the key argument is that only a $\Theta(\frac{1}{k^2})$-fraction of triples H can be unbiased.

Lemma 9. *Fix any good pair (T, F). Then one has*

$$\Pr_{H \sim \binom{F}{3}}[\texttt{UNBIASED}(H) = 1] \leq \frac{100}{k^2}.$$

Proof. Consider triples $H, H^* \in \binom{F}{3}$ that are both unbiased. We claim that then $|H \cap H^*| \geq 2$. For the sake of contradiction suppose that $|H \cap H^*| \leq 1$. Then there are distinct nodes $u, v \in V(H \backslash H^*) \cap C$.



By assumption, H^* is unbiased, hence $p_{M,T}(H^* \cup \{(u, v)\}) > 0$. In other words, there exists a matching $M \in \mathcal{M}(T)$ with $H^* \cup \{(u, v)\} \subseteq M$. But also (T, H) is unbiased and hence $p_{U,T}(H) > 0$, which implies that there is a cut $U \in \mathcal{U}(T)$ so that $U \cap C = V(H) \cap C$. Then (u, v) runs inside of U and hence $|\delta(U) \cap M| = 1$, which is a contradiction to $\mu_1(\mathcal{R}) = 0$. Thus unbiased pairs must indeed overlap in at least 2 edges.

Now fix a triple H^* that is unbiased (if there is none, there is nothing to show). Then

$$\Pr_{H \sim \binom{F}{3}}\left[\texttt{UNBIASED}(H) = 1\right] \leq \Pr_{H \sim \binom{F}{3}}\left[|H \cap H^*| \geq 2\right] \leq \frac{3k}{\binom{k}{3}} \leq \frac{100}{k^2}.$$

This settles the claim. □

Now we can easily relate the contribution of the good pairs to $\mu_3(\mathcal{R})$. In particular, we use that if a triple $H \subseteq F$ is unbiased, then $p_{M,T}(H) \leq (1 + \varepsilon) p_{M,T}(F)$ (same for cuts). Formally, we have

$$\begin{aligned}
&\mathop{\mathbb{E}}_{T}\Big[\mathop{\mathbb{E}}_{|F|=k}\Big[\texttt{GOOD}(T, F) \cdot \mathop{\mathbb{E}}_{H \in \binom{F}{3}}\big[p_{M,T}(H) \cdot p_{U,T}(H)\big]\Big]\Big] \\
&\leq \mathop{\mathbb{E}}_{T}\Big[\mathop{\mathbb{E}}_{|F|=k}\Big[\mathop{\mathbb{E}}_{H \in \binom{F}{3}}\Big[\underbrace{\texttt{SMALL}(T, F, H) \cdot p_{M,T}(H) \cdot p_{U,T}(H)}_{\leq 2^{-\delta m}}\Big]\Big]\Big] \\
&+ \mathop{\mathbb{E}}_{T,F}\Big[\underbrace{\mathop{\mathbb{E}}_{H \in \binom{F}{3}}\big[\texttt{UNBIASED}(T, F, H) \cdot}_{\leq \frac{100}{k^2}} \underbrace{p_{M,T}(H)}_{\leq (1+\varepsilon) p_{M,T}(F)} \cdot \underbrace{p_{U,T}(H)}_{\leq (1+\varepsilon) p_{U,T}(F)}\big]\Big] \\
&\leq \frac{100(1+\varepsilon)^2}{k^2} \mathop{\mathbb{E}}_{T}\Big[\mathop{\mathbb{E}}_{|F|=k}\big[p_{M,T}(F) \cdot p_{U,T}(F)\big]\Big] + 2^{-\delta m} \\
&\leq \frac{200}{k^2} \mu_k(\mathcal{R}) + 2^{-\delta m}
\end{aligned}$$

In the last inequality, we assume that $\varepsilon \leq \frac{1}{8}$.

4.5 Contribution of bad partitions

Unfortunately, not all pairs (T, F) will be good, so we need to bound the contribution of those that are *bad*. It can be proven that the following is true:

$$\mathop{\mathbb{E}}_{T,F}\Big[\texttt{BAD}(T, F) \cdot \mathop{\mathbb{E}}_{H}\big[p_{M,T}(H) p_{U,T}(H)\big]\Big] \leq \varepsilon \mathop{\mathbb{E}}_{T,F}\Big[\mathop{\mathbb{E}}_{H}\big[p_{M,T}(H) p_{U,T}(H)\big]\Big]$$

This then implies that the contribution of good pairs already determines the claim of Lemma 8.

To prove this remaining inequality, we want to give the argument why one would expect that a random partition is "good". Let us only show half of the claim by focusing on the cut part (the arguments for the matching part are analogous). In fact, it suffices to show that for a random partition T and a 3-matching H, either the smallness part or the unbiasedness part is satisfied with high probability. Then we can take the union bound over $\binom{k}{3}$ choices of $H \in \binom{F}{3}$.

Let us make this more formal. For a partition T and a corresponding 3-matching $H \subseteq C \times D$, let us say that (T, H) is *U-good* if at least one of the conditions is true:

- the pair is U-small: $p_{U,T}(H) \leq 2^{-\delta m}$
- the pair is U-unbiased: $p_{U,T}(H) \in (1 \pm \varepsilon) \cdot p_{U,T}(C)$

Now we want to prove:

Lemma 10. *For any $\varepsilon > 0$ there is a constant $\delta > 0$ so that*

$$\Pr_{T,H}[U\text{–}\texttt{GOOD}(T, H)] \geq 1 - \varepsilon.$$

We will choose the random pair (T, H) in a "reverse" manner. First, take another random partition $\tilde{T}$ and 3-matching $H \subseteq C \times D$ as in Figure 3. Now, imagine that we "forget" which nodes are $C/V(H)$; instead we have disjoint blocks $\tilde{A}_1, \ldots, \tilde{A}_{m+1}$, all of size $k - 3$, see Figure 4. Now, we choose an index $i^* \in [m + 1]$ at random and declare the block $\tilde{A}_{i^*}$ as the missing nodes $C/V(H)$; the other blocks then constitute $A_1, \ldots, A_m$. This finally gives us a partition T, still with a corresponding 3-matching H. We claim that with probability $1 - \varepsilon$, the pair (T, H) will be U-good. In fact, a stronger claim holds, namely we could even adversarily fix the choice of $\tilde{T}$ – we need only the tiny bit of randomness that lies in the choice of the index i^* for that claim.

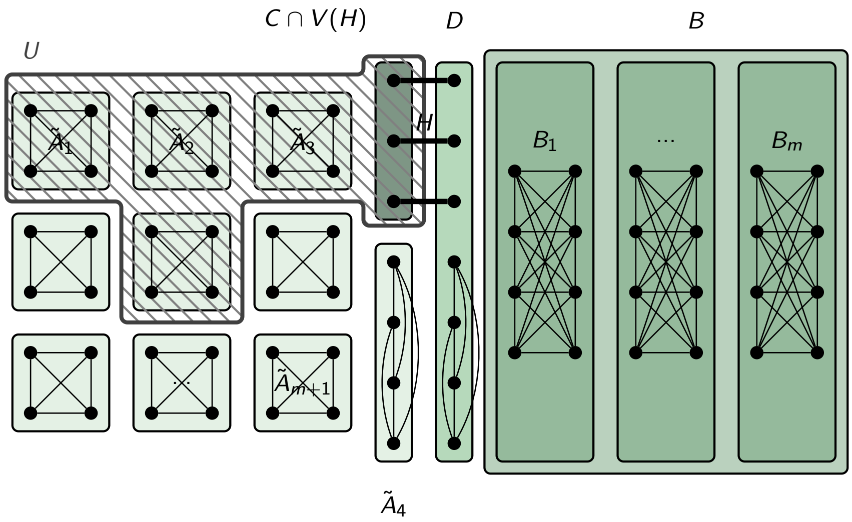


Figure 4. *Situation after we decided for* $\tilde{A}_1, \dots, \tilde{A}_{m+1}, H, B_1, \dots, B_m$

Now, consider the function

$$f(I) := \big(C \cap V(H)\big) \cup \bigcup_{i \in I} \tilde{A}_i$$

which gives a cut for each index set $I \subseteq [m+1]$. Moreover, consider the family of such cuts

$$Y := \{ f(I) \in \mathcal{U} \mid I \subseteq [m+1] \}$$

that lie in our rectangle. One condition that we need before we can apply Lemma 6 is that Y is large enough, say $|Y| \geq 2^{(1-\delta)m}$. But if that is not the case, then all indices i^* will lead to U-small pairs and we are done. So suppose that $|Y| \geq 2^{(1-\delta)m}$ is indeed satisfied. Then Lemma 6 provides that if we chose $\delta > 0$ small enough, then for a $(1-\varepsilon)$-fraction of indices $i^* \in [m+1]$ the fraction of cuts in Y that contain the block $\tilde{A}_{i^*}$ is $\frac{1}{2} \cdot (1 \pm \varepsilon)$. As we learned in Cor. 7, this implies that $p_{U,T}(H) = (1 \pm \varepsilon) p_{U,T}(C)$ for all unbiased indices.

Remark 1. Here, we showed that a random pair (T, H) will be good with probability $1 - \varepsilon$. For a formal proof one needs a slightly more complicated statement: fix an entry $(U^*, M^*) \in Q_3$ and then take a random pair (T, H) *containing* that entry (U^*, M^*); then that pair (T, H) will be good with high probability. The details can be found in [20] – but they do not contain any more crucial ideas.

5 Conclusion

The result can be modified to show that any $(1+\varepsilon)$-approximate linear program for the matching polytope (which contains also non-perfect matching) must have size $2^{\Omega(1/\varepsilon)}$, see Pokutta and Braun [4] and the remarks in [20].

After a sequence of papers showed lower bounds on the size of linear programs (including [10, 5] and this result), the natural next challenge was whether one could also prove lower bounds on the size of *semidefinite programs*. A recent breakthrough of Lee, Raghavendra and Steurer [15] answers this affirmatively for the correlation polytope, the cut polytope and approximate versions of constraint satisfaction problems. However, it is still unknown whether there is a polynomial size SDP for the perfect matching polytope.

Thomas Rothvoss, Department of Mathematics & Department of Computer Science, University of Washington, Seattle, WA, USA.
Email: rothvoss@uw.edu

Note

1. For technical reasons we will always assume that the dimension of P is at least 1.

References

[1] D. Avis and H. R. Tiwary. On the extension complexity of combinatorial polytopes. In *ICALP (1)*, pages 57–68, 2013.
[2] E. Balas. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM J. Algebraic Discrete Methods*, 6(3):466–486, 1985.
[3] F. Barahona. On cuts and matchings in planar graphs. *Mathematical Programming*, 60:53–68, 1993. 10.1007/BF01580600.
[4] G. Braun and S. Pokutta. The matching polytope does not admit fully-polynomial size relaxation schemes. *CoRR*, abs/1403.6710, 2014.
[5] S. O. Chan, J. R. Lee, P. Raghavendra, and D. Steurer. Approximate constraint satisfaction requires large lp relaxations. *FOCS'13*, 2013.
[6] J. Edmonds. Maximum matching and a polyhedron with 0, 1-vertices. *J. Res. Nat. Bur. Standards Sect. B*, 69B:125–130, 1965.
[7] J. Edmonds. Matroids and the greedy algorithm. *Mathematical Programming*, 1(1):127–136, 1971.
[8] Y. Faenza, S. Fiorini, R. Grappe, and H. R. Tiwary. Extended formulations, nonnegative factorizations, and randomized communication protocols. In *ISCO*, pages 129–140, 2012.
[9] S. Fiorini. Personal communication, 2013.
[10] S. Fiorini, S. Massar, S. Pokutta, H. Tiwary, and R. de Wolf. Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In *STOC*, pages 95–106, 2012.
[11] A. M. H. Gerards. Compact systems for t-join and perfect matching polyhedra of graphs with bounded genus. *Operations Research Letters*, 10(7):377 – 382, 1991.
[12] M. Goemans. Smallest compact formulation for the permutahedron. Working paper. http://math.mit.edu/ goemans/PAPERS/permutahedron.pdf, 2010.
[13] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.
[14] V. Kaibel, K. Pashkovich, and D. O. Theis. Symmetry matters for the sizes of extended formulations. In *IPCO*, pages 135–148, 2010.
[15] J. R. Lee, P. Raghavendra, and D. Steurer. Lower bounds on the size of semidefinite programming relaxations. *CoRR*, abs/1411.6317, 2014.
[16] R. K. Martin. Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters*, 10(3):119 – 128, 1991.
[17] M. Padberg and M. Rao. Odd minimum cut-sets and b-matchings. *Mathematics of Operations Research*, pages 7:67–80, 1982.
[18] S. Pokutta and M. V. Vyve. A note on the extension complexity of the knapsack polytope. *Oper. Res. Lett.*, 41(4):347–350, 2013.
[19] A. Razborov. On the distributional complexity of disjointness. In *ICALP*, pages 249–253, 1990.
[20] T. Rothvoß. The matching polytope has exponential extension complexity. *CoRR*, abs/1311.2369, 2013.
[21] A. Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. A,B,C*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.

[22] P. Ventura and F. Eisenbrand. A compact linear program for testing optimality of perfect matchings. *Oper. Res. Lett.*, 31(3):429–434, 2003.
[23] M. Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441 – 466, 1991.

A glimpse of heaven is feeling you understand something beautiful

An interview with Jack Edmonds

This year we are celebrating the 50th anniversary of your groundbreaking papers about matching, which opened the gates to the field of Polyhedral Combinatorics and developed the most fundamental concept of a good characterization, which then led to the definition of the complexity classes NP and coNP and to the question of whether their intersection equals P. It seems that you did your work in an almost empty environment compared to the situation discrete optimizers find themselves in today.

It wasn't in a vacuum. As part of my job at NIST (U.S. National Institute of Standards and Technology, formerly the National Bureau of Standards) I saw papers by IBM'ers about how they solved combinatorial problems like simplifying switching circuits or cutting planes to solve ILP's. That's not a vacuum. I happened to be in the perfect time and place. Almost anyone there would have reacted similarly.

Also, you might have heard that my personality tends to have a chip on its shoulder, and that was kind of like how the sixties were. Either you were a hippie or a protester or both, and I was both – in mathematics, and publishing, and academia. A few years later students in Paris made riots.

Jobs were easy to get then. I dropped out of graduate school and got a wonderful job at NIST that was way better than graduate school. It became my graduate school. Kids in their teens and twenties could get whatever they went for. Demonstrations for civil rights, peace, disarmament, women's rights, free love, gay rights, were popular. My reaction to combinatorial optimization was in that spirit. I was, of course, kind of a geek, and so looking for nice polytopes was my way of correcting the establishment.

So your reaction was to search for something better than general purpose cutting planes?

I figured that we would find a polynomial time algorithm for traveling salesmen and for integer programming. I never read the details of the paper by Dantzig, Fulkerson, Johnson, but I got the message that they used an exponentially large set of inequalities, the subtour elimination inequalities which are easily recognizable. It just happened that they weren't enough. This made me think, ok, they just don't have all the inequalities, but presumably there is a set of, what we now call NP, polytime recognizable inequalities, whose solution-set is the hull of the tours. (It is not important that the set be irredundant. In fact a description of just the irredundant inequalities might be more complicated than for some larger valid set.)

I felt sure that the only reason the travelling salesman problem was not well-solved was that nobody had tried to find a determining set of polytime-recognizable inequalities. George, Ray, and Selmer, used obvious inequalites, an exponentially large set, and they weren't enough, but I thought I would find an NP description of a set which is enough. Man, if you have an easy description of a set of points like the incidence vectors of the matchings in a graph, or the incidence vectors of the traveling salesman tours in a graph, why shouldn't you have an easy description of inequalities defining the hull of those points. I was sure I was going to figure it out for tours. Matching was a technicality on the way. My only worry was that someone else would do it before me as soon as they saw it done for matchings. After a few years of frustrating unsuccess I conjectured that there is no good way to recognize when a tour is optimum – hence, NP is not coNP, and hence NP is not P.

People at that time did not show any interest in the idea of polynomial time. A thrill of my life was in the mid 1960s when Dantzig invited me to lecture to his linear programming class, and I presented a sequence of non-degenerate shortest path problems where the number of pivots of the simplex method grows exponentially. My only pivot rule was that it improve the objective. Later, exponentiallity was proved for various specific pivot rules by Klee, Minty, and others. All of this depended on caring.

When working on the matching problem, were you more interested in investigating just some problem for which it was both conceivable that there could be an efficient algorithm or not, or were you specifically interested in matchings?

The second thing.

Did you fail on other problems before you turned to matching or somehow you smelled that matching was the right problem to attack?

I looked at the cuts separating node s from node t in the network flow problem. That's too easy. I looked at tours. I thought that maybe sufficient inequalites for "0/1 valued 2-matching" polytopes together with subtour elimination inequalities would be enough for the traveling saleman problem. 1-matching provides a simple extended formulation of a 0/1 valued 2-matching in a graph G (a subset of edges in G such that at most 2 of them hit any node).

Tutte had a theorem which characterizes when a graph has a perfect 1-matching (a subset of edges which hits each node exactly once) with a proof which was not helpful to me algorithmically for optimum matching, or even for finding a perfect matching. I could see that Berge's augmenting path theorem did not achieve polytime. I have an enormous debt to both of them, and they both became closest friends. However the biggest jump for me was a min-max formula, which is the LP duality theorem applied to a certain TDI system of inequalities. In the case of the max cardinality 1-matching problem, that min-max formula is: max size of a 1-matching equals min capacity-sum of an odd-node-set covering where, for $k = 1, 2, \dots$, a node-set Q of size $2k + 1$ has capacity k and covers the edges with both ends in Q, and a node-set $\{v\}$ has capacity 1 and covers the edges which hit node v.

Obviously, the size of any 1-matching is at most the capacity-sum of any "odd-node-set covering" of all the edges. That was the 1st breakthrough for me, and so all I then needed was a polynomial time algorithm to find a 1-matching and a covering which achieved equality. From that it was easy to do the same for a general objective function. From that a corollary is an NP inequality-set whose solution-set is the hull of the 1-matchings.

Curiously, no one had ever used a min-max equality to trivially conclude the convex-hull sufficiency of an LP relaxation. It was popular to go in the other direction. It seems to be still not noticed that Birkhoff's 1946 theorem on the hull of doubly stochastic matrices is a corollary of Egervary's 1931 optimum bipartite matching min-max.

Berge's min-max formula for largest 1-matching didn't make sense to me. It still doesn't. But this simple idea does: when you're optimizing over 'feasible' subsets of objects, if you can get an easy upper bound on the size of the intersection of any feasible set with a certain kind of subset, say Q, of the objects, then take that upper bound as the capacity of Q. If you are lucky, the largest size of a feasible set equals the minimum capacity-sum of a cover by sets Q. If you are really lucky these sets Q give you adequate inequalities for the hull of the feasible sets. This same idea works for some matroidal optimization problems and for stable sets in perfect graphs.

What else did you come across when starting to think about combinatorial optimization problems?

Computing theory consisted of minimizing switching circuits, minimizing conjunctive normal formulas, and Gomory's cutting planes. Nice polytopes meant polytopes with few enough inequalities that you can explicitly record them. I took some pride in revolting against exponential time approaches in the same way my buddies were against the American draft and the abuse of black people. We were all protesting something.

There was Land and Doig, branch-and-bound. Did you actually burn Land and Doig's paper?

Hahaha. Actually I didn't look at their work and lots of other work. But I liked Gomory's cutting plane methods. I felt we might get some more polytime methods with it if we got it away from exponential simplex methods. I felt we could, and tried, to get LP away from exponentiallity since the LP duality theorem is a good (i.e., $NP \cap coNP$) characterization of LP optima. The inequalities for matchings are obtained by Gomory cuts. Later work by Chvàtal and others showed some intrinsic limitations of Gomory cuts. However it is certainly a beautiful way to derive inequalities. Gomory implicitly proved that the hull of integer solutions is obtained by iterated cutting, but if I remember correctly the only theorem in the ILP book by Hu was that the algorithm is finite, which for most problems is not so interesting. Classes of sets of integer points, besides matchings, for which polynomially iterated cuts yield hulls remain to be found. Anyway, the technicalities of matchings anyone could have done in the environment. I'm not agile, and that might have been an advantage.

What's important at the time was protesting against the establishment, academia, black lists, mistreatment of black brothers, mistreatment of South-East Asians, foolish exclusion of Jews and others by the Allied empire, ...

And the mistreatment of computing?

Exactly.

Let's come to another technicality, but which is important at the beginning of what you did. What you wanted to do was to investigate whether there is an efficient algorithm for a problem or not. Was it clear to you right away that polynomiality would be the right quantification of efficiency?

Good question. No. To me what was important was having some way to regard 'good' as mathematics – in particular 'good' certification – because I hoped that having a mathematically 'good' way to certify outputs, e.g., 'good' theorems, would lead to 'good' algorithms. Looking at theorems it occurred to me that what mathematicians regard as beautiful structural theorems are existentially polytime – that is, for any input there exists something, any instance of which is easy to recognize. This seems a simple formalization of mathematically beautiful, regardless of a good algorithm to find whatever exists. Thinking first about 'NP' required a meaning for 'P'. However, I was finding some algorithmic results which could not even be in a math publication without a mathematical definition of 'good'. Polynomial time seemed easy to work with.

In fact the 'always-polytime' definition of 'good' is actually not so good, but since few known methods qualify it's a good place to start. A main reason 'always-polytime' is not such a wonderful meaning for 'good' is that bad running time might be so rare that only theory cares – like for the simplex method. In fact, though I do not know an example, in theory an algorithm might be not polytime, and yet never be bad for inputs smaller than the number of elementary particles. Good algorithms in practice have been developed for the traveling salesman and other 'hard' problems. Polytime is a theoretically handy practical heuristic, and a toy for nerds. Topologists found that topological spheres and balls are not necessarily nice. 'Good characterization' means it is easy to prove if the answer is yes and it is easy to prove if the answer is no. Which is called $NP \cap coNP$. The main point is that if a question is in $NP \cap coNP$, i.e., proving the answer is yes is easy and proving the answer is no is easy, it ought to be easy to decide which – what is now called the conjecture $NP \cap coNP = P$. What the hell does 'easy' mean? Well, it's easy what one might mean by easy.

But still you were aware that you need a formal definition of what easy means. If you look at it in hindsight, polynomiality seems to be the only concept that works. You need something that is preserved under reductions. You didn't talk about reductions at that point in time, but you came up with this notion of efficiency. Does this mean that you had in mind already somethig like reductions?

Of course. I reduced shortest paths with negative costs to optimum assignment, reduced coloring to packing, this to that. How could one look for methods, and discard possibilities, without reductions? Scott Lockhart, the Yale student who programmed for Ellis Johnson and me, suggested we write up the reductions of one unsolved problem to another. It seemed like fun when there was time after solving problems. Would it have made us famous? No. What never occurred to us, darn it, is that we were hoping to solve some problems which are as hard as any NP problem. What a strange development. Is there an $NP \cap coNP$ problem which is as hard as any?

In A Glimpse of Heaven [1], some 25 years ago, you indicated that you still hadn't accepted that NP intersected coNP was not P. You phrased it in terms of integer programming, I think. But some 25 years ago, you apparently still hoped for NP intersected coNP equal to P. How do you think about this today?

In fact, I was naive. My feelings have changed only recently, and I tell you why. I was not aware of number theoretic stuff. I never studied very much the does-there-exist-an-integer-factor question, the is-this-a-prime question. I've learned a little about it lately. To be specific, there is now a famous polytime algorithm for, given an integer, is it prime or not? But given a number, if there is a factor, find one. There is no good algorithm known for this. I was unaware of this difference between decision problem and search problem. To me the only way to solve a decision problem was to solve the search problem. I was an idiot about that for most of my life. This is really exciting for cryptography – that there is a good algorithm for deciding if a number has a factor, but there's no good algorithm known for finding one. And I do not know who discovered it, but friends finally made me aware that finding a factor can be reduced to an $NP \cap coNP$ question. Being taught this kind of thing has made me a little less naive in my fantasy that $P = NP \cap coNP$.

And are you now disappointed? Did you really hope in your heart for NP intersected coNP being equal to P?

Yes. In fact, I still do. It has turned out to be true for a lot of problems. It may turn out that finding a factor is polytime (a cryptographic tragedy). All right, I will now hedge my bet, since I am so naive except for a tiny bit of combinatorial optimization. It is technically interesting, how I hedge the bet. The $NP \cap coNP$ things I have studied have been mostly polyhedral. So, I weaken my conjecture to the subclass of problems, possibly proper subclass, that are given by good descriptions of polyhedra. That is, if you have an NP-description of a set of points and you have an NP-description of a set of inequalities which gives the hull of those points, then applying LP-duality shows that recognizing an optimum over the set of points is in NP ∩ coNP. That was the idea which I was using 50 years ago. I didn't know anything about NP ∩ coNP except the idea of applying linear programming duality to those descriptions. I still carry the same fantasy that NP ∩ coNP = P. However I realize what I was thinking about was polyhedral.

So here's a complexity class, call it JC: NP ∩ coNP predicates that are corollaries of having an NP description of a set of points and an NP description of a set of inequalities which gives you the hull of the points. The weakened conjecture is that JC is contained in P. The ellipsoid method has made it almost half true. Of course it might be that NP ∩ coNP reduces to JC.

If all that is true, is that heaven? Is that your current definition of heaven?

Hahaha! You are pulling my leg.

No, I'm serious. You talked about a glimpse of heaven. What was heaven at that time and what is it now? I assume at that time NP ∩ coNP was heaven. So what is this the new heaven?

No, a glimpse of heaven is feeling you understand something beautiful. Heaven is understanding it all. There is an old country gospel song called "Farther Along". I have about 50 recordings of it.[From extensive experience during editing the text, *Optima* recommends to listen to the recording by Monte Ingersoll while reading the rest of the interview.] The refrain:

> Farther along we'll know more about it,
> Farther along we'll understand why;
> Cheer up, my brother, live in the sunshine,
> We'll understand it, all by and by.

Corinthians 13 says:

> For now we see through a glass darkly;
> I know in part;
> but then shall I know even as also I am known.

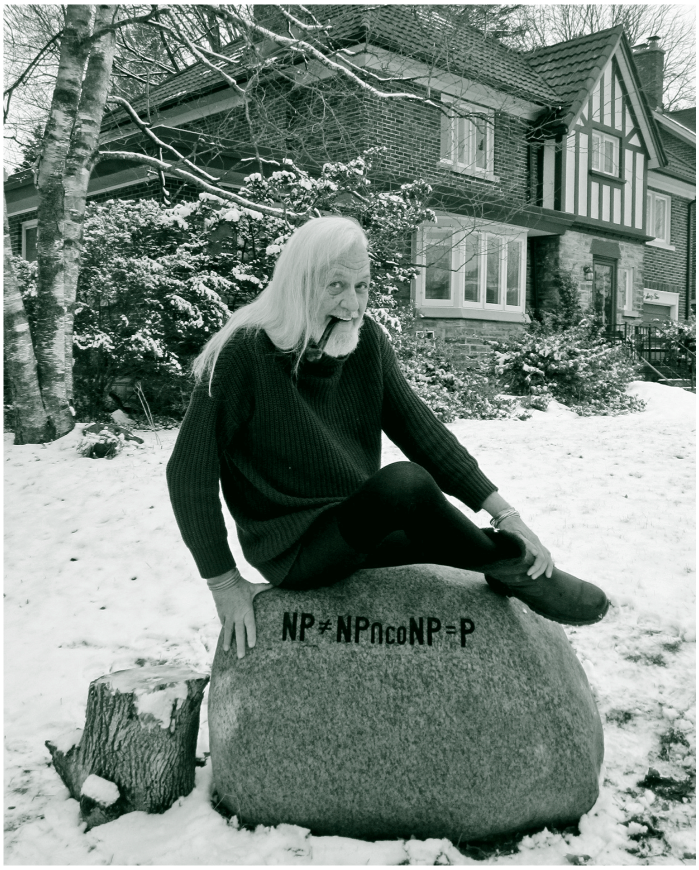
I have a son, Jeff, who is a professor of computing theory. And I have a son, Alex, who just got his Bachelor's degree in Zermelo-Fraenkel's set theory with the axiom of choice and the Gödel incompleteness Theorems.

So he's a mathematician?

Yes, but now he thinks he'd rather be a plumber or something. Jeff and I have been trying to turn Alex on to Goodstein's Theorem and Hydra battles. Have you ever heard of the Goodstein sequence? This is nice.

You know the East-Indians taught us how to express a number as a sum of powers of a base-n. Instead, we not only express our number in base-n, but we recursively do the same thing for the exponents. This is called hereditary base-n representation. So in case of base-2 you have all these towers like two to the two to the two to the two with some additions along the way. A Goodstein sequence is a sequence of numbers, starting with any positive integer. To obtain the *(k+1)*th number of the Goodstein sequence, we write the *(k)*th number with hereditary base-k representation. Then replace every appearance of exponent *k* with *k+1* and subtract 1 from the resulting number.

What do you think happens? Goodstein says that no matter what number you start with that sequence eventually terminates with zero even though the numbers grow horrendously quickly for horrendously many terms. Now, what's interesting is that this cannot be proved, like most truths we know about numbers can be proved, by



Jack Edmonds in front of his house in Kitchener, Ontario

using normal Peano arithmetic axioms. Expositors say it is proved using set theory, which to me is a fairy tale.

You don't like set theory.

I don't like the power set axiom. I like Cantor proving that from any infinite list of real numbers you can describe a real number not in the list. This is kind of like for any integer you can describe a bigger one. People tend to say 'you can't prove such and such using Peano arithmetic axioms, but you can prove it using set theory', which includes using an axiom of choice saying that for any set of subsets of the set of all real numbers there exists a meaningful selection of one element from each of the subsets. Some rather ridiculous things can be proved using this set theory, ZFC, but you surely do not need much of it to prove termination of Goodstein's sequence. There are believable things in between. So that's what Jeff, Alex, and I have been studying. Maybe heaven is understanding it all by and by.

So this would be more important for you than settling NP ∩ coNP?

Oh no. But I have followed the attempts to prove NP not equal to P enough to know that I am not capable of proving coNP is not NP. It might well be that there is no proof from more believable assumptions. In the meantime it is a good axiom. It amazes me that proofs are eventually found to exist for so many mathematical truths.

Is it true that you are mainly interested in mathematical questions which are also rather philosophical?

Or beautiful or easy. I'm slow. I've never understood anything that takes more than a few enjoyable pages.

Let's come back to matching once more. Thomas Rothvoss has this brilliant proof that shows that there is no polynomial size extended formulation for the perfect matching polytope. He will explain it in the same issue of Optima in which this interview is going to appear. If it would have turned out to be the other way, so if somebody suddenly would have come *up with an extended formulation for the matching polytope of size, say, n^5. Would you have been surprised, would you have liked it, or would it have had disturbed your view on the matching problem?*

I would have not liked it, I would have been disturbed and greatly disappointed. I can recognize Rothvoss' achievement as really great. However, the reason it's a great achievement is because my department mate Ted Swart insisted on a flawed proof of a good characterization of the traveling saleman polytope, and that prompted Yannakakis to brilliantly refute Swart's proof. And because of limitations Yannakakis described in his own related result about matching polytopes, there was a great challenge which was met by Rothvoss. This is the reason that I admire it as important work.

Would you say that among the systems over which you can separate in polynomial time those that are of polynomial size do not play a particularly important role?

Yes.

If every problem that has a good description would have had a compact extended LP description it would mean that Karmarkar's algorithm could have solved these problems and not only the ellipsoid algorithm.

That's a good point. But first let's clarify that the ellipsoid algorithm only provides a polytime algorithm for optimizing when there is a polytime algorithm for separation, and vice versa. A main example: For a so called polymatroid, P, given by an oracally NP set of inequalities, specified by an oracally given submodular set function, there is an especially simple polytime algorithm, called the greedy algorithm, which maximizes any linear objective over P, and hence provides an oracally NP description of the vertices of P, and hence an oracally NP description of any point x in P as a convex combination of points given by the greedy algorithm. The ellipsoid method does more. It immediately provides a polytime algorithm for separation – that is for, given any point x, determining that x is in P by describing x as a convex combination of points given by the greedy algorithm, or else determining that x is not in P by specifying one of the oracally given inequalities which x violates. It took some 30 years and many papers to finally develop direct algorithms for separation of the point x which are polytime without using the ellipsoid method.

Given several polymatroids in the same space by submodular-function systems of inequalities, the ellipsoid method gives us immediately a polytime algorithm for optimizing a linear function over the intersection of these polymatroids by using the polytime algorithm for possible separation of any point x by this intersection.

And so we here have main examples of where certain JC problems have polytime algorithms. It is only my fantasy conjecture that JC in general implies good algorithms for separation and optimizing. It happens that if the set of inequalities is 'compact' then there is a trivial good algorithm for separation, and if the set of points is 'compact' there is a trivial good algorithm for optimizing over them. Yes, the value of having a system of inequalities which is small enough to explicitly list is that you can apply algorithms that work well for explicitly listed LPs.

It may seem that it is more complicated to optimize over matching polytopes than over spanning tree polytopes, and the latter have polynomial extension complexity, but the first don't have. Is this only a mere coincidence?

Yes. A mere coincidence. Optimizing over the polytope whose vertices are the incidence vectors of the linearly independent subsets of columns of a matrix is as simple as optimizing over a spanning tree polytope, namely by the greedy algorithm, but I doubt if the former has a compact extended formulation. This might have an interesting but easier proof than Rothvoss' proof for matching polytopes.

I agree. Also, if you intersect two graphic matroids you still have small extension complexity, while optimization is not so simple anymore. So you would say that it is not conceivable that small extended formulations have an important theoretical meaning.

I would say that. On the other hand impossibility proofs seem to be what mathematicians have always most admired. And theories are mainly for fun. Like the impossibility of proving Goodstein's Theorem about integers from the usual axioms about integers. What mathematician studies how to solve a quartic polynomial, or even a cubic polynomial, using radicals? But every pure math student studies a proof that it is not possible to solve a fifth degree polynomial by radicals. I'm a possibility person without much experience of proofs that something is impossible. But complexity theorists are obsessed with proving things are not possible, and I know of great complexity theorists who are interested in work on extended formulation complexity.

Are extended formulations relevant in any sense?

Oh yes. Extended formulations are extremely natural and extremely useful. Indeed, one excuse for somewhat neglecting b-matching polytopes is that they have extended formulations to 1-matching polytopes. And a very easy extended formulation of the Chinese Postman polytope is a b-matching problem where you have a loop at each of the nodes, so all that is important is parity of the number of edges hitting each of the nodes.

As another example, some 35 years ago I suggested to K. Cameron that we take a dual network flow problem and play around with eliminating variables to come up with something nice. We did so and called it coflow polyhedra.

So you appreciate the usefulness and beauty of extended formulations.

What I question is the value of small extended formulations of large NP systems to explain their efficiency. What polyhedra do we know with integer vertices and with a system of inequalities which is small enough to be efficiently listed explicitly? I think we only know totally unimodular systems. Paul Seymour showed that totally unimodular constraint matrices are essentially network flow matrices, dual flow matrices, one-, two-, three-sums of these and one small exception. So, from what is known to us, small combinatorial inequality system means, effectively, flow or dual flow. Wouldn't it be closer to heaven to prove that every class of small integer LPs is totally unimodular, or whatever? Didn't Hoffman and Kruskal almost do that? And then see what we get by projecting away variables? Of course 'class' needs to be defined somehow.

Let us jump to one more general question. It's decades ago that your papers on matching polytopes, and submodularity related polytopes, have been written. Are you happy with the development that has happened after your papers?

No. But it seems that God is to blame and not researchers. I expected that these, and perfect graph polytopes studied by Ray Fulkerson, were just the beginning of what we are calling JC. To some extent that has been true. However I expected many more. Instead everyone started finding that thousands of problems are NP hard. And 50 years later, JC classes we know of are still surprisingly rare.

Is this due to the fact that there simply are not so many relevant examples? Or do you think we are just too stupid or too lazy to find other significant ones?

Submodularity related classes of polyhedra are extensive and rich. So are perfect-graph related classes. Continuing research about various classes and new classes has been far from stupid, in fact very deep, while much of continuing combinatorial research has been in different directions like LP approximation methods and semi-definite programming.

For thousands of years the beautiful symmetries of a few polyhedra and tilings were at the center of refined math, and still are if you consider Klein's Erlangen Program, and the frightfully obscure Field-medal-winning Langlands Program, both based on group theory. Turing computers and operations research have displaced them – so far at least without group theory. Are JC classes the newer dodecahedra? I never expected that. Perhaps in our NP-hardness mode, we are missing some JC. Hopefully, there are undiscovered classes of well-describable polyhedra, i.e., JC classes, out there.

Do you have any hint into which directions there could be important ones to detect?

No, but perhaps if I did have I wouldn't say.

Thank you very much, Jack, for this interview, and even more for your beautiful mathematics!

* * *

References

[1] J. Edmonds, *A glimpse of heaven*. In J. K. Lenstra, A. H. G. Rinnooy Kan, A. Schrijver (eds.), History of Mathematical Programming – A Collection of Personal Reminiscences, CWI, Amsterdam and North-Holland, Amsterdam, 1991, pp. 32–54.

[2] M. Grötschel (ed.), *Optimization Stories*. DOCUMENTA MATHEMATICA, Journal der Deutschen Mathematiker-Vereinigung, Extra Volume ISMP (2012), www.zib.de/groetschel/publications/OptimizationStories.pdf.

[3] M. Jünger et al. (eds.), *50 Years of Integer Programming 1958–2008*. Springer, 2010.

[4] M. Jünger, G. Reinelt, and G. Rinaldi (eds.), *Combinatorial Optimization – Eureka, You Shrink!*. Springer, 2001.

* * *

The questions were posed by Volker Kaibel, Jon Lee, and Jeff Linderoth on January 8, 2015, in Aussois, France. For further background reading next to [1] we in particular refer to the chapters of [2] contributed by Bill Cunningham and Bill Pulleyblank, to Bill Cook's article in [3], and to [4]. Possibilities to attend lecture series by Jack Edmonds in the near future include:

- ECCO XXVIII, Catania, Italy, May 2015
  www.ecco2015.unict.it/index.html
- Belgrade, Serbia, June 2015
  www.mi.sanu.ac.rs/weekly/mini.htm
- London, GB, June 2015
  www.maths.qmul.ac.uk/~fink/Edmonds2015.html
- PoCo 2015, Pittsburgh, U.S.A., July 2015
  http://poco2015.org

Daniel Bienstock

Mathematical Programming Computation

The journal *Mathematical Programming Computation* was launched in 2009, with the goal of addressing computational issues in mathematical optimization. As the reader may already know, the journal has a two-track reviewing process, whereby a standard paper is refereed in the normal way, while simultaneously the software component of the submission is reviewed by special referees with the goal of running the experiments in the paper and verifying, as much as possible, that the software does what it purports to. The goal is to have high-quality publications in mathematical optimization which additionally adhere to standard scientific criteria for reproducibility of experiments. The initial editor was Bill Cook; please see www.mathopt.org/Optima-Issues/optima78.pdf and mpc.zib.de for a history of the journal.

I became editor-in-chief in January 2015. As Area Editors we have a strong group: Alper Atamturk (Linear and Integer Programming), Robert Fourer (Modeling Languages and Systems), Andrew Goldberg (Graph Algorithms and Data Structures), Nicolas I.M. Gould (Nonlinear Optimization), Jeffrey T. Linderoth (Stochastic Optimization, Robust Optimization, and Global Optimization), F. Bruce Shepherd (Combinatorial Optimization) and Kim-Chuan Toh (Convex Optimization). Additionally, the journal relies on an extremely strong group of Associate Editors and Technical Editors (responsible for reviewing software). Please refer to the above websites for a complete listing of the editorial board.

In this column I will address some issues of interest to the MPC community and more broadly to the mathematical optimization community.

1. We are now rolling out an online publication tracking system. If you submit a paper to MPC you may (and in the future, will) interact with this system. As we are all aware, such systems are never perfect; however we hope that overall this is a good change.
2. We have started an online discussion forum. Any topic of interest (see below) can be raised in this forum. If you are interested in joining in, please send me an email at dano@columbia.edu.
3. We plan to occasionally have focused issues of MPC, used to highlight specific topics of interest. It is important to note that these will not be traditional "special issues" that rely on guest editors. Instead the normal editorial and reviewing process will be followed. The only difference will be that I will attempt to schedule the papers to appear in a common volume. The goal of these measures is to maintain a common quality level across all publications in the journal. If you think of a topic that could deserve a focused issue in MPC please let me know.
4. We are encouraging authors whose papers are accepted to archive the version of the software that was submitted, with MPC, and thus make it available to the public. This point has not proved universally popular and it is worth discussing here. There are a number of reasonable points that could be raised against making software in accepted papers publicly available. For instance, we are all familiar (perhaps painfully familiar) with the potential for misuse of software that we post. In the case of MPC-archived software, it should be understood that there is no presumption of support. If you download software, you are on your own, unless the authors kindly agree to help you. Moreover the archived software will remain unchanged, whereas it is possible that the authors will continue to develop new versions, which will not be posted at MPC. Another reason is that some of our authors have commercial associations, and their employers might object to posting software – even software that has been reviewed by the journal (I know that this can happen even if it may sound irrational). This would certainly be a valid excuse; as I noted above, we are "suggesting" rather than requiring archival.
   Now let me state two major reasons *for* software archival. One is that we want the community to be able to validate experiments. You learn by re-running experiments that other people have performed. A researcher studying a problem might benefit, years later, from running an old experiment verbatim; that is to say using the original software rather than a new version (which could be better, but who knows?). Another reason is that just like we learn by reading other people's proofs, we also learn by reading their software – and why is that any different? And just like published articles sometimes include errors, it could be the case that archived software is found to be buggy – why should we try to hide such errors? It would be to the benefit of the author that

such errors are found and corrected (else the manuscript that was published could become problematic). In any case, I would like to stress that the practice of archiving software, for public distribution, has a positive effect and it enhances our community's scientific standing.

5. There is another potential initiative for MPC that we have considered; and here I would like to stress that the MPC Board is not in complete agreement and thus this initiative may not be implemented. The board is considering the possibility of allowing "short communications" in MPC. Before presenting my rationale for such an initiative, let me first discuss what it would not be. First of all, there would not be an expectation of a special review process. To be sure, once a reviewer gets started the processing time might be faster with a shorter manuscript (and less software). Likewise, the time and effort needed for an author to write a short communication would also be decreased. However we would not promise a special review track for short communications. Also, we would not want authors to split up a normal paper into several "short" papers. Should an author engage in this practice, it will be quickly identified, and stopped.

   Rather, the goal here is to facilitate the publication of compelling and fresh research that does not lend itself to a long manuscript. I am sure that readers have experienced instances, in other journals, where reviewers ask for "more material" in order to judge a manuscript acceptable. An extreme form of this might be that of a very theoretical article where reviewers demand computational experiments so as to validate the work. Such a request may or may not be valid; however there is no question that it delays publication of what otherwise might be a very compelling note. Another perspective concerns the significant experimental component of articles submitted to MPC. When we are engaged in computational research we often find ourselves performing very low-level studies (say, in a small set of problem instances) geared to understanding why a promising idea is instead misbehaving. Often, such studies are not deemed worthy of publication – instead, computational articles usually will focus on "macro" statistics such as running time, percentage of problem instances solved within a certain error margin in a certain amount of time, and so on. These are very valuable data; yet the detailed experiments I described above would also prove invaluable, in my opinion. Thus, for example, somebody working in integer programming would be interested in knowing why a promising branching rule fails to work well on some instances. Publishing a detailed account of this failure (or successes) could prove very useful. Such a manuscript might only amount to a short note, but if the experiment is suitably compelling, and the writing is informative and appropriate, the manuscript might prove a very valuable contribution. In my mind, this type of practice would place experimental work in optimization firmly within standard scientific norms used for decades or more in other disciplines, such as physics.

   In any case, I realize that the above paragraph may not naturally appeal to math optimization people. We have a discussion forum on this topic at MPC. If you are interested in joining, let me know. Ultimately the board will decide on the topic of short contributions, but of course it is in our interest to listen to what the community has to say.

I look forward to your input on this and any other topic of interest to the MPC community.

Daniel Bienstock, Department of Industrial Engineering and Operations Research and Department of Applied Physics and Applied Mathematics, Columbia University, 500 West 120th St., New York, NY 10027 USA
dano@columbia.edu

Nominations for 2015 MOS Elections

Nominations are solicited for all offices (Chair, Treasurer, and four At-Large Members of Council) of the Mathematical Optimization Society (MOS). The new Members-at-Large of the Council will take office at the time of the symposium, while the Chair-Elect and the Treasurer-Elect will take office one year later.

Candidates must be members of the Society (search at www.ams.org/cml/) and may be proposed either by Council or by any six members of the Society. No proper nomination may be refused, provided the candidate agrees to stand.

The following procedure will be observed.

1. Nomination to any office is to be submitted to William Cook (Chair of the MOS). Such nomination is to be supported by the nominator and at least five other members of the Society.
2. Nominations must be received by email (bico@uwaterloo.ca), on or before April 30, 2015.
3. In keeping with tradition, the next Chair should preferably not be a resident of North America. The membership is asked to consider only residents from other continents as candidates for the Chair.
4. Members of MOS on the rolls as of May 1, 2015 are eligible to vote. When the ballots are counted, the four At-Large candidates for Council having the highest number of votes will be elected; however, no more than two members having permanent residence in the same country may be elected.
5. The election will close on May 31, 2015.

Further information on the form of balloting will be forthcoming.

William Cook, MOS Chair

Call for Nominations
INFORMS John von Neumann Theory Prize

The John von Neumann Theory Prize is awarded annually to a scholar (or scholars in the case of joint work) who has made fundamental, sustained contributions to theory in operations research and the management sciences. The award is given each year at the INFORMS Annual Meeting if there is a suitable recipient. Although the Prize is normally given to a single individual, in the case of accumulated joint work, the recipients can be multiple individuals.

The Prize is awarded for a body of work, typically published over a period of several years. Although recent work should not be excluded, the Prize typically reflects contributions that have stood the test of time. The criteria for the Prize are broad, and include significance, innovation, depth, and scientific excellence. The award is $ 5000, a medallion and a citation.

Application Process. The Prize Committee is currently seeking nominations, which should be in the form of a letter (preferably email) addressed to the prize committee chair (below), highlighting the nominee's accomplishments. Although the letter need not contain a detailed account of the nominee's research, it should document the overall nature of his or her contributions and their impact on the profession, with particular emphasis on the prize's criteria. The nominee's curriculum vitae, while not mandatory, would be helpful. Please compress electronic files if 10 MB.

Nominations should be submitted to the committee chair (see below) as soon as possible, but no later than June 1, 2015. Please see this page online for complete details: http://tinyurl.com/5vdb5lg

2015 Committee Chair: George Nemhauser, Professor, Georgia Institute of Technology, Atlanta, Georgia, 30332-0205, U.S.A.
george.nemhauser@isye.gatech.edu

Call for papers

Mathematical Programming Series B: DC Programming: Theory, Algorithms and Applications

(Guest editors: Le Thi Hoai An and Pham Dinh Tao)

Aims and Scope: DC (Difference of Convex) programming with local and global approaches, which constitute the backbone of nonconvex programming and global optimization, were extensively developed during the last two decades. These theoretical and algorithmic tools have been successfully applied for modeling and solving real-world nonconvex programs from different fields of Applied Sciences. This special issue, celebrating 30 years of developments of DC programming and DCA, aims at publishing contributions of high-quality which will be concerned with works on DC programming: Local and Global Approaches, from both a theoretical and an algorithmic point of view, and applications:

1. Refinement of local optimality conditions related to special classes of DC programs
2. Regularization techniques in DC programming
3. Improvement of solution algorithms with rate of convergence. New efficient approaches
4. Nonsmooth nonconvex equations systems
5. DC programming on Riemannian manifolds
6. Modeling and solving DC programs in combinatorial optimization, multiobjective/multilevel programming, VIP/MPEC, nonconvex programming with SOC/SDP constraints, sparse optimization, robust optimization, optimization under uncertainty, distributed/parallel nonconvex programming, for real-world applications to Transport-Logistic, Communication Systems, Network Optimization, Energy Optimization, Finance, Bioinformatics, Information Security, Cryptology, Mechanics, Image Processing, Robotics & Computer Vision, Automatic Control, Machine Learning.

Every paper must fit within the 'Aims and Scope' of this special issue.

All papers will be subjected to a standard refereeing procedure of Mathematical Programming before it can be accepted for publication. Accepted papers must meet the standards of the journal. Due to page limitations, we expect that each paper will have not more than 25 pages. All papers should be submitted through MP's web page: www.editorialmanager.com/mapr/ When the paper is submitted, the author is required to choose Jong-Shi Pang in Request Editor. The LaTeX style files for MPA are mandatory and can be downloaded from http://www.mathopt.org/?nav=journal#latex.

Deadline for submission of full paper to the special issues is 30 August, 2015. We plan to publish this special issue in Fall 2016. We look forward to receiving your contribution to this special MPB issue.

Call for papers

Mathematical Programming Series B: Variational Analysis and Optimization

(Guest editors: Samir Adly and Asen Dontchev)

Aims and Scope: Although there is already a rich literature in variational analysis and optimization, in the recent years there have been new developments not only in the theory but also various important applications in science, engineering and economics. This special issue aims to publish outstanding papers centered around the broad area of variational analysis and optimization and beyond, including in particular nonsmooth analysis, topics in functional analysis related to optimization, nonlinear programming, mathematical economics, risk theory, optimal control, numerical methods for optimization and optimal control, as well as applications related to all these areas. Some of the papers will be based but not limited to presentations at the forthcoming conference "Variational Analysis and Optimization", May 18–22, 2015 in Limoges, France, dedicated to R. Tyrrell Rockafellar on the occasion of his 80th birthday. Every paper must fit within the 'Aims and Scope' of this special issue.

All papers will be subjected to a standard refereeing procedure of Mathematical Programming before it can be accepted for publication. Accepted papers must meet the standards of the journal. Due to page limitations, we expect that each paper will have not more than 25 pages. All papers should be submitted through MP's web page: www.editorialmanager.com/mapr/ When the paper is submitted, the author is required to choose Jong-Shi Pang in Request Editor. The LaTeX style files for MPA are mandatory and can be downloaded from http://www.mathopt.org/?nav=journal#latex.

Deadline for submission of full paper to the special issues is 1 December, 2015. We plan to publish this special issue in 2016. We look forward to receiving your contribution to this special MPB issue.

ISMP 2015 in Pittsburgh

The 22nd International Symposium on Mathematical Programming (ISMP 2015) will take place in Pittsburgh, PA, USA, July 12–17, 2015. ISMP is a scientific meeting held every three years on behalf of the Mathematical Optimization Society (MOS). At ISMP 2015 there will be more than 1500 talks on all aspects of mathematical optimization and we expect more than 1700 participants!

Registration and Important Dates

April 15, 2015:	Early registration deadline
June 8, 2015:	Hotel reservation deadline
July 12, 2015:	Opening ceremony at Wyndham Grand Pittsburgh Downtown Hotel
July 17, 2015:	Scientific program ends at 6 pm

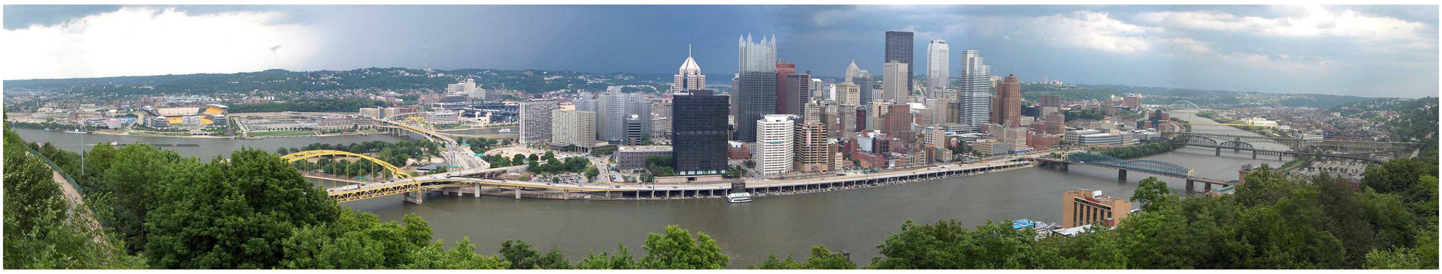
Conference registration is now open at www.ismp2015.org. Registration is available for participants even if you are not presenting.

Conference Venue. The symposium and opening ceremony will take place at the Wyndham Grand Pittsburgh Downtown Hotel located at the confluence of Pittsburgh's famed Three Rivers.

The opening ceremony will feature the presentation of awards by the Mathematical Optimization Society and will be followed by the welcome reception.

Pittsburgh is defined by its rivers, set on a "Golden Triangle" of land where the Allegheny and Monongahela Rivers meet to form the Ohio. The conference dinner will be on Wednesday July 15th featuring a river cruise on the Monongahela, Allegheny and Ohio Rivers, with panoramic views of beautiful landscapes and towering skyscrapers of Pittsburgh. The scientific talks will be scheduled on Monday–Friday 9 am–6 pm.

Plenary and Semi-plenary Speakers. ○ Laurent El Ghaoui (University of California Berkeley), *Optimization in the Age of Big Data: Sparsity and Robustness at Scale* ○ Jim Geelen (University of Waterloo, Canada), *Matroid Minors Project* ○ Daniel Kuhn (Ecole Polytechnique Federale de Lausanne, Switzerland), *A Distributionally Robust Perspective on Uncertainty Quantification and Chance Constrained Programming* ○ Daniel A. Spielman (Yale University), *Laplacian Matrices of Graphs: Algorithms and Applications* ○ Stephen J. Wright (University of Wisconsin-Madison), *Coordinate Descent Algorithms* ○ Samuel A. Burer (University of Iowa), *A Gentle, Geometric Introduction to*



Downtown Pittsburgh taken from the Mt. Washington Overlook at Grandview Ave & Bertha St. (Photo: Andrew "Dobie" Koch [CC BY 2.5])

Copositive Optimization ○ Roberto Cominetti (Universidad de Chile, Chile), *Equilibrium routing under uncertainty* ○ Michele Conforti (Universita' di Padova, Italy), *A geometric approach to cut-generating functions* ○ Tamara G. Kolda (Sandia National Laboratories), Optimization Challenges in Tensor Factorization ○ Andrea Lodi (University of Bologna, Italy), *On Mathematical Programming with Indicator Constraints* ○ Asu Özdaglar (Massachusetts Institute of Technology), *Fast Distributed Algorithms for Multi-agent Optimization* ○ Werner Römisch (Humboldt Universität zu Berlin, Germany), *Quasi-Monte Carlo methods for linear two-stage stochastic programming problems* ○ Frank Vallentin (University of Köln, Germany), *Mathematical Optimization for Packing Problems* ○ Pascal van Hentenryck (NICTA, Australia), *Complexity, Approximation, and Relaxation of the Power Flow Equations* ○ Yaxiang Yuan (Chinese Academy of Sciences, China), *Recent Advances in Trust-Region Algorithms*
For further details including the abstracts of these talks please visit www.ismp2015.org/?page_id=13.

Website. More details (including clusters, cluster chairs, all registration rates, discounted MOS membership for 2016–2018, hotel, sponsorship opportunities, exhibits etc.) will soon be available on the conference web site at www.ismp2015.org.

Summer School on Polyhedral Combinatorics (PoCo 2015)

The Summer School on Polyhedral Combinatorics (PoCo 2015) will be held on the campus of Carnegie Mellon University, Pittsburgh, from July 8–12, 2015.

The speakers will be Egon Balas, Amitabh Basu, Jesus De Loera, Xiaotie Deng, Jack Edmonds, Nick Harvey, Volker Kaibel, Gleb Koshevoy, Gianpaolo Oriolo, Thomas Rothvoss, Laura Sanita, Francisco Santos and Rico Zenklusen.

Polyhedral Combinatorics is at the heart of combinatorial optimization, a thriving area of mathematics with numerous applications. The Summer School on Polyhedral Combinatorics is aimed at students who want to learn more about this beautiful field. Topics covered will include: ○ Matroids, ○ submodularity, ○ matchings and stable sets in graphs, ○ construction of extended formulations and complexity, ○ diameter of polytopes, ○ disjunctive programming, ○ multi-row cuts, ○ PPAD completeness of bimatrix games, ○ algebraic techniques in discrete optimization, and ○ linear programming-based methods for approximation algorithms.
Please register at http://poco2015.org.

Application for Membership

I wish to enroll as a member of the Society. My subscription is for my personal use and not for the benefit of any library or institution.

☐ I will pay my membership dues on receipt of your invoice.
☐ I wish to pay by credit card (Master/Euro or Visa).

Credit card no. *Expiration date*

Family name

Mailing address

Telephone no. *Telefax no.*

E-mail

Signature

Mail to:
Mathematical Optimization Society
3600 Market St, 6th Floor
Philadelphia, PA 19104-2688
USA

Cheques or money orders should be made payable to The Mathematical Optimization Society, Inc. Dues for 2015, including subscription to the journal *Mathematical Programming*, are US $90. Retired are $45. Student applications: Dues are $22.50. Have a faculty member verify your student status and send application with dues to above address.

Faculty verifying status

Institution

IMPRINT
Editor: Volker Kaibel, Institut für Mathematische Optimierung, Otto-von-Guericke Universität Magdeburg, Universitätsplatz 2, 39108 Magdeburg, Germany. kaibel@ovgu.de ■ **Co-Editors:** Samuel Burer, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242-1994, USA. samuel-burer@uiowa.edu ■ Jeff Linderoth, Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI 53706-1572, USA. linderoth@wisc.edu ■ **Founding Editor:** Donald W. Hearn ■ Published by the Mathematical Optimization Society ■ Design and typesetting by Christoph Eyrich, Berlin, Germany. optima@0x45.de ■ Printed by Oktoberdruck AG, Berlin, Germany.