

The Newsletter of the INFORMS Optimization Society

Volume 7 Number 1 May 2017

Contents

Chair's Column	
The IOS 2018 Optimization Conference	
INFORMS Journal on Optimization	
Nominations for OS Prizes	19
Nominations for OS Officers	
Featured Articles	
LP formulations for mixed-integer polynomial mization problems	-
Gonzalo Muñoz	
Relative Entropy Convex Programs	
Venkat Chandrasekaran and Parikshit Shah	10

Send your comments and feedback to the Editor:

Marina A Epelman Industrial and Operations Engineering University of Michigan, Ann Arbor, USA mepelman@umich.edu

Chair's Column

David Morton

 $\begin{array}{c} \hbox{Industrial Engineering and Management Sciences} \\ \hbox{Northwestern University} \\ \hbox{USA} \end{array}$

david.morton@northwestern.edu

Dear Fellow IOS Members:

It's an exciting time to work in optimization! There are many supporting indicators:

- We see explosive growth of optimization in machine learning and high-dimensional statistics.
- The 2017 Mixed Integer Programming Workshop at HEC Montréal is "sold out."
- The INFORMS Board approved the new *IN-FORMS Journal on Optimization*.



2016 IOS Prizes winners (left to right): Sven Leyffer, Gonzalo Muñoz, Aharon Ben-Tal, Parikshit Shah, Hamza Fawzi, and Venkat Chandrasekaran

- IOS membership has doubled to 1627 members in the last 5 years.
- *INFORMS OS Today* (this newsletter) is doubling from one to two issues per year.

The INFORMS Journal on Optimization — yes, that's its name after a spirited discussion in our community — has been approved by the INFORMS Board with Dimitris Bertsimas as the inaugural Editor-in-Chief. His vision for the journal is included in this newsletter, and look for additional information in forthcoming issues.

Optimizers across the planet owe a large-scale thank you to Suvrajeet Sen. During his tenure as the previous Chair of IOS, Suvrajeet worked with unwavering energy to take the *INFORMS Journal on Optimization* from being a notion discussed fleetingly at IOS Business Meetings, and elsewhere, to a reality. Thank you, Suvrajeet!

We all know that semidefinite programs (SDPs) play a prominent role, both in generalizing important classes of convex optimization problems and in providing relaxations of hard optimization problems. The analogous role that relative entropy programs (REPs) can play is much less known. Venkat Chandrasekaran and Parikshit Shah received the 2016 IOS Young Researchers Prize for "Relative Entropy Relaxations for Signomial Programming." Their article in this issue summarizes that paper's results, including how REP provides an attractive hierarchy of convex relaxations for signomial programs, and further contrasting their hierarchy with that provided by sum-of-squares techniques.

Gonzalo Muñoz garnered the 2016 IOS Student Paper Prize for his paper, with Dan Bienstock, "LP Formulations for Mixed-Integer Polynomial Optimization Problems." In this issue, they describe a family of LP approximations that exploits sparsity in polynomial optimization problems, where sparsity is measured by the width of a tree-decomposition of the constraints' intersection graph. Their LP approximations yield near-optimal and near-feasible solutions of the polynomial optimization problem for any specified tolerance. A dense constraint causes difficulties for their first approach, and so in the context of polynomial optimization problems defined on networks, Bienstock and Muñoz discuss an equivalent reformulation that allows for small tree width.

Look for articles from Aharon Ben-Tal (Khachiyan Prize) and Sven Leyffer (Farkas Prize) in the second issue of our 2017 newsletter. These four prizes are highly competitive, and truly earn their characterization as "esteemed recognition." The deadline for submitting nominations for all four 2017 prizes is nearly upon us: <u>June 15th</u>. Please see the call for prize nominations in this newsletter.

Thank you to our IOS Officers — Necdet Serhat Aybat, Amitabh Basu, Güzin Bayraksan, Dan Bienstock, Pietro Belotti, Hande Benson, Austin Buchanan, Marina Epelman, Burcu Keskin, Siqian Shen, and Sauleh Siddiqui — who have growing responsibilities with our growing membership, handling more IOS sessions at our conferences, more newsletters, and more Web site updates, all enabling greater IOS contributions and visibility. Please see the call for nominations for IOS officers in this newsletter.

Finally, we organize an INFORMS Optimization Society Conference every other year, and when I say "we," for our 2018 conference, I mean: Steve Billups (University of Colorado Denver), Manuel Laguna (University of Colorado Boulder), and Alexandra Newman (Colorado School of Mines). The conference will take place March 23–25, 2018 at the University of Colorado's downtown Denver location, and "we" have an exciting line-up of plenary speakers, as detailed in this newsletter. When registration becomes available on-line, register quickly, before it sells out!



Denver, CO — location of the 2018 IOS Conference (image used with permission of the University of Colorado)

LP formulations for mixed-integer polynomial optimization problems

Daniel Bienstock and Gonzalo Muñoz
Columbia University
USA

dano@columbia.edu and gonzalo@ieor.columbia.edu

This article presents a summarized version of [3], where we develop a class of linear programming approximations for polynomial optimization problems that take advantage of structured sparsity of the constraints. This structured sparsity will be given by *tree-width*, a graph theoretical parameter that roughly measures how *tree-like* a given graph is.



Gonzalo Muñoz and Jim Luedtke



Hamza Fawzi (recipient of Honorable Mention for Student Paper) and Jim Luedtke

In this summary we will show how the tree-width parameter can be included in a polynomial optimization context, and how it can be exploited so as to obtain tractable approximations. More specifically, we will show two ways of introducing structured sparsity in optimization problems: the first uses the *intersection graph*, a graph defined for any instance of a polynomial optimization problem, and the second relies on the concept of Network Polynomial Optimization problems (NPO), a family of optimization problems with an underlying network given on its definition. We argue how different these two setting can be, and give small LP approximations for both cases.

1 Polynomial Optimization under structured sparsity

We consider general polynomial optimization problems

(PO): min
$$c^T x$$

s.t.: $f_i(x) \ge 0$ $1 \le i \le m$ (1a)
 $x \in [0,1]^n$, (1b)

where the f_i 's are polynomials. We focus on integer programming techniques to obtain linear programming approximations to (1) that attain any desired tolerance (both feasibility and optimality). Moreover, when problem (1) is appropriately sparse, our linear programs are of polynomial size.

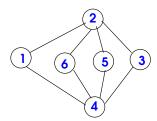
To model sparsity we rely on two standard concepts: the *intersection graph*, and *tree-width*.

Definition 1. The intersection graph [10] for the system of constraints (1a) is the undirected graph which has a vertex for each variable and an edge for each pair of variables that appear in any common constraint.

Example 2. Consider the following system of constraints on variables x_1, \ldots, x_6 :

$$\begin{array}{ll} 3x_1^2-x_2\geq 0, & -2x_2^2+x_3^3\geq \ 1, & x_2+x_6=1,\\ x_3^3-x_4^2\leq 2, & x_1+x_4\leq 0, & x_2+x_5\geq 0,\\ x_4-x_5^3\leq 0, & x_5^2-x_4^2=0. \end{array}$$

The intersection graph is as follows:



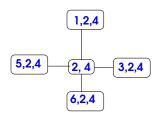
The intersection graph depicts the complexity of relationships among variables. If the intersection graph is dense then, potentially, problem (1) could prove difficult. In fact, we will need a specific parameterization of sparsity in order to obtain a tractable formulation. In what follows, given a graph H, we will use V(H) and E(H) to denote the vertex-set and edge-set of H, respectively; and $\delta_H(u)$ will be the set of edges incident with vertex u.

Definition 3. Let G be an undirected graph. A treedecomposition [19, 18] of G is a pair (T, Q) where T is a tree and $Q = \{Q_t : t \in V(T)\}$ is a family of subsets of V(G) (the vertices of G) such that

- (i) For all $v \in V(G)$, the set $\{t \in V(T) : v \in Q_t\}$ forms a subtree T_v of T, and
- (ii) For each $\{u,v\} \in E(G)$ there is a $t \in V(T)$ such that $\{u,v\} \subseteq Q_t$, i.e. $t \in T_u \cap T_v$.

The width of the decomposition is defined as $\max\{|Q_t|: t \in V(T)\} - 1$. The tree-width of G is the minimum width over all tree-decompositions of G.

In the case of Example 2 a tree-decomposition is:



The tree T is a star with four leaves, and the sets Q_t are indicated inside each vertex of T. The width of this tree-decomposition is 2.

Tree-width, roughly speaking, indicates how "tree-like" a graph is. Trees are the graphs with tree-width 1, and a clique on n vertices has tree-width n-1. It can be shown that a graph with tree-width ω and n vertices has $O(\omega^2 n)$ edges, and

so graphs of small tree-width are sparse, however not all sparse graphs have small tree-width. We refer the reader to [4, 6, 1, 9] for additional background.

The following notation will be used in our results. Given $\alpha \in \mathbb{Z}_+^n$ we write

$$x^{\alpha} \doteq \prod_{j=1}^{n} x_j^{\alpha_j}.$$

For the i^{th} constraint, $f_i(x) \geq 0$, we are given a (finite) set $I(i) \subseteq \mathbb{Z}_+^n$ and we represent $f_i(x)$ as a sum of weighted monomials:

$$f_i(x) = \sum_{\alpha \in I(i)} f_{i,\alpha} x^{\alpha},$$

where each $f_{i,\alpha}$ is rational. We write

$$||f_i||_1 = \sum_{\alpha \in I(i)} |f_{i,\alpha}|.$$

The degree of $f_i(x)$ is defined as $\max_{\alpha \in I(i)} \sum_j \alpha_j$. Finally, given a set of polynomial constraints $f_i(x) \geq 0$ i = 1, ..., m, and $\epsilon \geq 0$, we say a vector \hat{x} is ϵ -scaled feasible if $f_i(\hat{x}) \geq -\epsilon ||f_i||_1 \, \forall i = 1, ..., m$.

Our first result is as follows:

Theorem 4. Consider an instance of problem (1) and any $\epsilon \in (0,1)$. Let ρ the maximum degree of any of the polynomials f_i . Given a treedecomposition of the intersection graph of the constraints of width ω , there is an LP formulation with $O\left((2\rho/\epsilon)^{\omega+1} n \log(\rho/\epsilon)\right)$ variables and constraints, such that any optimal solution \hat{x} for the LP satisfies:

- 1. \hat{x} is ϵ -scaled feasible for (1a),
- 2. $c^T \hat{x} < P^* + ||c||_{1}\epsilon$, where P^* is the value of (1).

The statement is indicative of the fact that as $\epsilon \to 0$, we converge to an optimal solution. We sketch the proof in Section 3, for a full proof see [3].

2 Network Polynomial Optimization

Our second point of focus concerns polynomial optimization problems (1) where there is network struc-

ture as an explicit element of the problem description. In these problems we are given a network, $^{1}\mathcal{G}$, where for each node u in \mathcal{G} there is a set \mathcal{X}_{u} of variables associated with u and a set \mathcal{K}_{u} of constraints associated with u. The optimization problem is of the following form:

(NPO): min
$$c^T x$$

s.t. :
$$\sum_{\substack{\{u,v\}\\ \in \delta_{\mathcal{G}}(u)}} p_{(u,v)}^{(k)}(x) \ge 0, k \in \mathcal{K}_u, u \in V(\mathcal{G}) \quad (2a)$$

$$x \in [0,1]^n. \quad (2b)$$

Here each $p_{(u,v)}^{(k)}$ is of the form

$$p_{(u,v)}^{(k)}(x) = p_{(u,v)}^{(k)}(\mathcal{X}_u \cup \mathcal{X}_v).$$

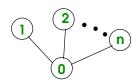
The notation stresses the *ordered* pair (u, v) and indicates that the polynomial only involves variables in $\mathcal{X}_u \cup \mathcal{X}_v$. The $p^{(k)}$ will be termed the *arc polynomials* of the problem. We do not assume that the sets \mathcal{X}_u are pairwise disjoint. Rather (for a technical reason) we allow intersections but we assume that for any variable x_j , the set of $u \in V(\mathcal{G})$ with $x_j \in \mathcal{X}_u$ induces a connected subgraph of \mathcal{G} .

There are many practical examples of problem (2), where the sets \mathcal{X}_u are small and the \mathcal{K}_u are small: the AC-OPF (Optimal Power Flow) problem in electrical transmission networks [17, 7], capacitated fixed-charge network flow models [14], and some optimization problems on gas networks [15].

In devising an algorithm for problem (2), a direct reliance on Theorem 4 runs into a difficulty as detailed next.

Example 5. Consider an NPO where

(i) The graph G is a star on n+1 vertices $0, 1, \ldots, n$ and edge set $\{0, j\}$ for $1 \le j \le n$:



¹We use the term 'network' to contrast with 'graph' which we reserve for intersection graphs. Likewise, we will typically use 'node' to refer to a vertex of a network.

- (ii) $\mathcal{X}_0 = \emptyset$ and $\mathcal{X}_j = \{x_j\}$ for $1 \le j \le n$.
- (iii) There is a single constraint (2a) and it is associated with node 0. In this constraint we have $p_{(0,j)}^{(1)}(\mathcal{X}_0 \cup \mathcal{X}_j) = jx_j^3 1/n \text{ for } 1 \leq j \leq n.$

Thus the (single) constraint (2a) for this example is: $\sum_{j=1}^{n} jx_{j}^{3} \geq 1$. This is a dense constraint, even though the input network \mathcal{G} is a tree. In fact, the intersection graph of this constraint is a clique on n vertices (with tree-width n-1), and, as a consequence, if we directly apply Theorem 4 we will obtain a formulation of size exponential in n.

In summary: even if an NPO problem arises from a network with small tree-width, the problem, if viewed directly as an instance of PO, may yield a very dense formulation. However, we argue that one can always reformulate any instance of NPO over a small tree-width graph \mathcal{G} so as to obtain an equivalent instance which, when viewed as a general problem (1) gives rise to an intersection graph with (still) small tree-width.

In the following result, we abbreviate constraint (2a) as $p^{(k)}(x) \geq 0$.

Theorem 6. Consider an instance of NPO (2) and any $0 < \epsilon < 1$. Let D, Δ and ρ be such that:

- The network \mathcal{G} has at most D edges incident with any node,
- The number of variables plus the number of constraints associated with any node of G is at most Δ.
- Every polynomial $p^{(k)}$ has maximum degree $\leq \rho$.

Then there is a linear program of size $O((D\rho/\epsilon)^{O(\Delta\omega)} n \log(\rho/\epsilon))$, such that any optimal solution \hat{x} for the LP satisfies:

- 1. \hat{x} is ϵ -scaled feasible for (2a),
- 2. $c^T \hat{x} \leq P^* + ||c||_1 \epsilon$, where P^* is the value of (2).

In Example 5, $\omega = 1$, D = n, $\rho = 3$ and $\Delta = 1$. Thus Thorem 6 yields an LP formulation of size $O((n/\epsilon)^{O(1)}\log(1/\epsilon))$.

²Using this notation we allow cases where only variables in \mathcal{X}_u or in \mathcal{X}_v are actually involved.

3 Proof sketch for Theorem 4

In this section we show a proof sketch for Theorem 4. Toward this goal, we will provide an approximation technique that replaces a polynomial optimization problem with one over binary variables. See [11, 2, 8, 13] for related material. In Section 3.1 we will reformulate this problem as a linear program, and in Section 3.2 we complete the proof.

Let r be a real with $0 \le r \le 1$. Then, given $0 < \gamma < 1$, we can approximate r as a sum of inverse powers of 2, within additive error γ . Let

$$L_{\gamma} \doteq \lceil \log_2 \gamma^{-1} \rceil.$$

Then there exist 0/1-values z_h , $1 \le h \le L_{\gamma}$, so that

$$\begin{split} \sum_{h=1}^{L_{\gamma}} 2^{-h} z_h & \leq r & \leq \sum_{h=1}^{L_{\gamma}} 2^{-h} z_h + 2^{-L_{\gamma}} \\ & \leq \sum_{h=1}^{L_{\gamma}} 2^{-h} z_h + \gamma & \leq 1. \end{split}$$

To apply this idea to problem (1), let $0 < \epsilon < 1$, and as before let ρ denote the maximum degree of any polynomial in (1). Choose γ so that

$$\epsilon = 1 - (1 - \gamma)^{\rho}.$$

Then for each j we approximately represent x_j as $\sum_{h=1}^{L_{\gamma}} 2^{-h} z_{j,h}$ where each $z_{j,h}$ is a (new) binary variable. We consider the following replacement for (1):

GB(
$$\gamma$$
): min $\sum_{j=1}^{n} c_{j} \left(\sum_{h=1}^{L_{\gamma}} 2^{-h} z_{j,h} \right)$
s.t. : $\sum_{\alpha \in I(i)} f_{i,\alpha} \prod_{j=1}^{n} \left(\sum_{h=1}^{L_{\gamma}} 2^{-h} z_{j,h} \right)^{\alpha_{j}} \ge -\epsilon \|f_{i}\|_{1},$
 $z_{j,h} \in \{0,1\}, \ 1 \le j \le n, \ 1 \le h \le L_{\gamma}.$

The following results describe the quality of the approximation to problem (1) provided by $GB(\gamma)$.

Lemma 7.

(a) Suppose \tilde{x} is feasible for (1). Then there is a feasible solution for $GB(\gamma)$ with objective value at most $c^T\tilde{x} + \epsilon \|c\|_1$.

(b) Suppose \hat{z} is feasible for $GB(\gamma)$. Define $\hat{x}_j = \sum_{h=1}^{L_{\gamma}} 2^{-h} \hat{z}_{j,h}$ for $1 \leq j \leq n$. Then \hat{x} is ϵ -scaled feasible for (1a).

Corollary 8. Let P^* be the optimal value of problem (1), and let \hat{z} be optimal for $GB(\gamma)$. Write $\hat{x}(\gamma)_j = \sum_{h=1}^{L_{\gamma}} 2^{-h} \hat{z}_{j,h}$ for $1 \leq j \leq n$. Then $c^T \hat{x}(\gamma) \leq P^* + \epsilon \|c\|_1$ and $\hat{x}(\gamma)$ is ϵ -scaled feasible for (1a).

Remark. As per the corollary, $\hat{x}(\gamma)$ achieves feasibility and optimality tolerance proportional to ϵ for problem (1). But as we mentioned above, $\hat{x}(\gamma)$ may actually be *super*-optimal for (1). Nevertheless for any sequence $\gamma_k \to 0^+$, the vectors $\hat{x}(\gamma_k)$ will have an accumulation point x^* , and this point necessarily must be feasible and thus optimal for (1).

3.1 From IPs to LPs

We have established that the pure binary problems $GB(\gamma)$ provide an approximation to (1) within guaranteed tolerance. This result holds for all problems (1) regardless of sparsity. In this section we will now consider sparse cases and analyze the complexity of the pure binary problems.

To this effect, suppose that (T,Q) is a treedecomposition for the intersection graph of an instance of problem (1). We will now construct a treedecomposition for the intersection graph for the corresponding instance of problem $GB(\gamma)$. This treedecomposition will be of the form (T,Q') (note: same tree T) where for any vertex t of T we set

$$Q'_t = \{ z_{j,h} : x_j \in Q_t \text{ and } 1 \le h \le L_\gamma \}.$$

Lemma 9. (a) (T, Q') is a tree-decomposition of the intersection graph for problem $GB(\gamma)$. (b) Further, if (T, Q) has width w then (T, Q') has width at most $L_{\gamma}(w+1)-1$.

Remark. As a consequence of this result, not only does problem $GB(\gamma)$ provide a close approximation to problem (1), but when (1) is structurally sparse (small tree-width of the intersection graph) then so is $GB(\gamma)$, modulo the $O(L_{\gamma})$ multiplicative increase in tree-width. We now pause, and state an important result:

Theorem 10. Consider a polynomial optimization problem on binary variables

$$\min \ \bar{c}^T y \tag{4a}$$

s.t.:
$$\bar{f}_i(y) \ge 0 \quad 1 \le i \le \bar{m}$$
 (4b)

$$y \in \{0, 1\}^N,$$
 (4c)

where each \bar{f}_i is a polynomial. Suppose we have a tree-decomposition of the intersection graph of the constraints (4b) of width ϕ , say. Then there is an equivalent reformulation for (4) as a linear program with $O(2^{\phi}N)$ variables and constraints, with $\{0,1,-1\}$ -valued constraint coefficients.

This theorem is implied by the results in Section 8 of [16], although it is not explicitly stated there. An independent proof and generalizations can be found in [3]. Related results can be found in [5, 20].

3.2 Putting it all together

Here we use the above results to prove Theorem 4. Consider an instance of (1) on n variables. As before let ρ represent the maximum degree of any of the polynomials f_i . Suppose we have a tree-decomposition of width ω of the intersection graph of the constraints. Given $0 < \epsilon < 1$ we proceed as follows:

- 1. We choose $\gamma = 1 (1 \epsilon)^{1/\rho}$, so that $\epsilon = 1 (1 \gamma)^{\rho}$. Without loss of generality ϵ is small enough so that $\gamma > \frac{\epsilon}{2\rho}$.
- 2. We apply Theorem 10 to construct the linear programming reformulation of the all-binary problem $GB(\gamma)$. Let us call this linear program $LP(\gamma)$.
- 3. As per Corollary 8, GB(γ) and thus, LP(γ), yield a vector $\hat{x}(\gamma)$ that is ϵ -scaled feasible for (1a) and $c^T \hat{x}(\gamma) \leq P^* + \epsilon ||c||_1$.

Next we analyze the size of $LP(\gamma)$.

1. By Lemma 9, there is a tree-decomposition of the intersection graph for $GB(\gamma)$ of width at most $L_{\gamma}(\omega+1)-1$, where

$$L_{\gamma} = \lceil \log_2 \gamma^{-1} \rceil < \log_2(2\rho/\epsilon) + 1$$

for ϵ small enough.

- 2. Further, GB(γ) has $N = nL_{\gamma} = O(n \log_2(\rho/\epsilon))$ variables.
- 3. Thus, writing $\phi = L_{\gamma}(\omega + 1) 1$, by Theorem 10, the number of variables and constraints in $LP(\gamma)$ is

$$O(2^{\phi}N) = O((2\rho/\epsilon)^{\omega+1} n \log_2(\rho/\epsilon)).$$

The proof of Theorem 4 is now complete.

4 Proof sketch for Theorem 6

Now we return to the *network* polynomial optimization problems, or NPOs. In order to illustrate some of the ideas used in the proof of Theorem 6, we revisit Example 5. Here, the network \mathcal{G} was a "star" with node set $0, 1, \ldots, n$ and center node 0. The NPO had a single constraint (2a), associated with node 0, with arc polynomials $p_{(0,j)}^{(1)}(\mathcal{X}_0 \cup \mathcal{X}_j) = jx_j^3 - \frac{1}{n}$ for $1 \leq j \leq n$. The corresponding constraint (2a) reads: $\sum_{j=1}^n jx_j^3 \geq 1$. This is a dense constraint and a direct application of Theorem 4 will produce a formulation of size exponential in n.

This example illustrates the point that if, in an NPO, a node has high degree, the intersection graph of the NPO will likely have high tree-width. This observation suggests that we should try to reformulate an NPO into an equivalent NPO on a network where every node has small degree.

Consider an NPO, $\bar{\mathcal{P}}$, on a network $\bar{\mathcal{G}}$ all of whose nodes have degree at most three, and with associated sets of variables $\bar{\mathcal{X}}_u$. Let $(\bar{T}, \bar{\mathcal{Q}})$ be a tree decomposition of $\bar{\mathcal{G}}$. Form the pair $(\bar{T}, \bar{\mathcal{Q}}')$, where for each $t \in V(\bar{T})$ we define

$$\bar{\mathcal{Q}}'_t = \bigcup \{\bar{\mathcal{X}}_v : v \in \bar{\mathcal{Q}}_t \text{ or } \exists u \in \bar{\mathcal{Q}}_t \{u, v\} \in E(\bar{\mathcal{G}})\}.$$

We have that:

Theorem 11. (\bar{T}, \bar{Q}') is a tree-decomposition for the intersection graph of NPO $\bar{\mathcal{P}}$. If the width of $(\bar{T}, \bar{\mathcal{Q}})$ is \bar{w} , then the width of $(\bar{T}, \bar{\mathcal{Q}}')$ is at most $3(\bar{w}+1)\bar{k}-1$, where $\bar{k} = \max_{u \in V(\mathcal{G})} |\bar{\mathcal{X}}_u|$.

The result suggests a way to obtain Theorem 6. Namely, given \mathcal{P} an NPO on a general network of

"small" tree-width, we reformulate it as an equivalent NPO, $\bar{\mathcal{P}}$, on a network $\bar{\mathcal{G}}$ with nodes of degree at most three and such that $\bar{\mathcal{G}}$ also has "small" tree-width. If, in addition, the parameter \bar{k} in the statement of Theorem 11 is also "small" then our reformulation will be an NPO which, as a general polynomial optimization problem, has an intersection graph with small tree-width and can be handled using Theorem 4.

In the next section we follow this approach in order to find a good refomulation of the NPO in Example 5. For the general case we refer the reader to the full-length paper.

4.1 Reformulating the NPO in Example 5

In this section we will argue that the problem in Example 5 can be reformulated as an equivalent NPO on network of maximum degree three (and with small tree-width). And then, using the strategy in Theorem 11 we will argue that the intersection graph of such reformulation NPO has small tree-width.

Let us consider the case n=4. We will first produce the equivalent NPO by first constructing an extended formulation equivalent to $\sum_{j=1}^{4} jx_{j}^{3} \geq 1$. Later we will show that this extended formulation amounts to a new NPO.

The extended formulation has additional variables y_j (1 $\leq j \leq$ 7), and the following system of constraints whose sum yields $\sum_{j=1}^{4} j x_j^3 \geq 1$.

$$y_j \le jx_j^3 - \frac{1}{4}, \quad 1 \le j \le 4,$$
 (5a)

$$y_5 \le y_1 + y_2, \quad y_6 \le y_3 + y_4, \quad y_7 \le y_5 + y_6,$$
 (5b)

$$y_7 \ge 0. \tag{5c}$$

This system splits $\sum_{j=1}^{4} j x_j^3$ into partial sums with two terms each; y_5 and y_6 are stand-ins for these partial sums and y_7 represents the complete sum.³

System (5) is equivalent to $\sum_{j=1}^{4} jx_{j}^{3} \geq 1$ in the sense that the projection to x-space of the set of solutions to (5) equals $\{x \in \mathbb{R}^{4} : \sum_{j=1}^{4} jx_{j}^{3} \geq 1\}$. Let us put aside the issue that this equivalence might require some of the y_{j} to take values outside of the

range [0,1], which is not allowed in our formal definition (2) for an NPO. Modulo this point, we can argue that (5) is the system of constraints for an NPO. To construct this NPO we use a binary tree with vertices $1, \ldots, 7$ shown below. The variables

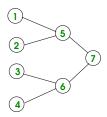


Figure 1: Binary tree replacement for star with leaves $1, \ldots, 4$.

for this NPO will be all the x_j and y_j . We associate with node 5 a family of arc polynomials that will yield (5a) for j=1,2 and the first inequality in (5b). Namely, we associate with node 5 the set of variables $\{x_1,\ldots,x_4\}\cup\{y_1,y_2,y_5\}$, define the arc polynomials

$$p_{(5,1)}^{(1)} = -y_1 + x_1^3 - \frac{1}{4}, \quad p_{(5,2)}^{(1)} = 0,$$

$$p_{(5,2)}^{(2)} = -y_2 + 2x_2^3 - \frac{1}{4}, \quad p_{(5,1)}^{(2)} = 0,$$

$$p_{(5,1)}^{(3)} = -\frac{1}{2}y_5 + y_1, \quad p_{(5,2)}^{(3)} = -\frac{1}{2}y_5 + y_2,$$

and impose the NPO constraints (associated with node 5):

$$p_{(5,1)}^{(k)} + p_{(5,2)}^{(k)} \ge 0, \quad k = 1, 2, 3.$$
 (7)

Likewise, we associate with node 6 the set $\{x_1, \ldots, x_4\} \cup \{y_3, y_4, y_6\}$, define

$$p_{(6,3)}^{(1)} = -y_3 + 3x_3^3 - \frac{1}{4}, \quad p_{(6,4)}^{(1)} = 0,$$

$$p_{(6,4)}^{(2)} = -y_4 + 4x_4^3 - \frac{1}{4}, \quad p_{(6,3)}^{(2)} = 0,$$

$$p_{(6,3)}^{(3)} = -\frac{1}{2}y_6 + y_3, \quad p_{(6,4)}^{(3)} = -\frac{1}{2}y_6 + y_4,$$

and impose

$$p_{(6,3)}^{(k)} + p_{(6,4)}^{(k)} \ge 0, \quad k = 1, 2, 3, \quad (9)$$

which yields (5a) for j = 3, 4 and the second inequality in (5b). Finally we associate with node 7 the variables $\{x_1, \ldots, x_4\} \cup \{y_7\}$, define

$$p_{(7,5)}^{(1)} = y_5 - \frac{1}{2}y_7, \quad p_{(7,6)}^{(1)} = y_6 - \frac{1}{2}y_7,$$

³This "splitting" technique is reminiscent of methods used in sparsification techniques for linear systems [12].

and set the constraint

$$p_{(7,5)}^{(1)} + p_{(7,6)}^{(1)} \ge 0.$$
 (11)

This yields the third inequality in (5b). The bound $y_7 \ge 0$ is implicit in the definition of an NPO. Thus, system (5) arises as the constraint set for an NPO.

Next we claim that as a *general* polynomial optimization problem, the problem with constraints (7)–(11) has an intersection graph of tree-width 2. The

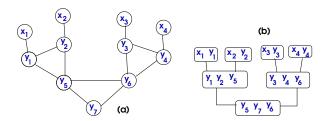


Figure 2: (a) Intersection graph for reformulation of Example 5. (b) A tree-decomposition.

intersection graph is shown in Figure 2(a) and a tree decomposition, of width 2, is given in Figure 2(b). Note that the tree in Figure 2(b) is isomorphic to that in Figure 1 — this is not an accidental event and, rather, reflects the structure of constraints (7)–(11). It is clear that the above process can be applied to the general case of $\sum_{j=1}^{n} jx_j^3 \geq 1$ so as to always yield a reformulation as an NPO on a binary tree with n leaves, a y_j variable for each internal node (and so less than n y_j variables), and such that the intersection graph has tree-width 2.

Regarding the aforementioned issue that the equivalence between system (5) and $\sum_{j=1}^{4} jx_{j}^{3} \geq 1$ might require that some y_{j} falls outside [0, 1], we handle this through a further reformulation, using the familiar trick of representing a real variable as the difference between two nonnegative variables and scaling. For these details we refer the reader to the full paper [3].

In summary, we have verified that the NPO in Example 5 can be reformulated as an equivalent NPO whose intersection graph has small tree-width. This makes it suitable for Theorem 4, which provides a polynomial-size linear programs that can approximate it.

In general, this reformulation requires a more delicate strategy for the vertex splitting procedure. As

pointed out in [3], splitting a vertex can potentially increase the tree-width of the resulting graph drastically. In the proof of Theorem 6 for the general case we show that the "right" vertex splitting always exists, and it is guided by the tree-decomposition of G.

REFERENCES

- [1] S. Arnborg, D. Corneil, and A. Proskurowski. Complexity of finding embeddings in a k-tree. SIAM Journal on Algebraic Discrete Methods, 8(2):277–284, 1987.
- [2] D. Bienstock. Histogram models for robust portfolio optimization. J. Comput. Finance, 11:1 – 64, 2007.
- [3] D. Bienstock and Muñoz G. LP approximations to mixed-integer polynomial optimization problems. Submitted. Available online in arXiv:1501.00288.
- [4] Daniel Bienstock and Michael A. Langston. Chapter 8 algorithmic implications of the graph minor theorem. In C.L. Monma M.O. Ball, T.L. Magnanti and G.L. Nemhauser, editors, Network Models, volume 7 of Handbooks in Operations Research and Management Science, pages 481 502. Elsevier, 1995.
- [5] Daniel Bienstock and Nuri Özbay. Tree-width and the Sherali-Adams operator. *Discrete Optimization*, 1(1):13–21, June 2004.
- [6] H. Bodlaender. Dynamic programming on graphs with bounded treewidth. In Timo Lepistö and Arto Salomaa, editors, Automata, Languages and Programming, volume 317 of Lecture Notes in Computer Science, pages 105–118. Springer Berlin Heidelberg, 1988.
- [7] Subhonmesh Bose, Dennice F. Gayme, S. H. Low, and K. Mani Chandy. Quadratically constrained quadratic programs on acyclic graphs with application to power flow. 2012.
- [8] Sanjeeb Dash, Oktay Günlük, and Andrea Lodi. On the mir closure of polyhedra. In *Integer Program*ming and Combinatorial Optimization, pages 337– 351. Springer Berlin Heidelberg, 2007.
- [9] Michael R. Fellows and Michael A. Langston. Nonconstructive advances in polynomial-time complexity. *Information Processing Letters*, 26(3):157?162, 1987.
- [10] D. R. Fulkerson and O. Gross. Incidence matrices and interval graphs. *Pacific Journal of Mathematics*, 15:835–855, 1965.
- [11] Fred Glover. Improved linear integer programming formulations of nonlinear integer problems. Management Science, 22(4):455–460, 1975.

- [12] J. Gondzio. Splitting dense columns of constraint matrix in interior point methods for large scale linear programming. Optimization, 24:285–297, 1992.
- [13] A. Gupte, S. Ahmed, M.S. Cheon, and S.S. Dey. Solving mixed integer bilinear problems using MIP formulations. SIAM Journal on Optimization, 23:721–744, 2013.
- [14] Mike Hewitt, George L. Nemhauser, and Martin W. P. Savelsbergh. Combining exact and heuristic approaches for the capacitated fixed-char ge network flow problem. *INFORMS Journal on Computing*, 22(2):314–325, 2010.
- [15] T. Koch, B. Hiller, M. Pfetsch, and L. Schewe, editors. Evaluating Gas Network Capacities. SIAM, Philadelphia, 2015.
- [16] M. Laurent. Sum of squares, moment matrices and optimization over polynomials. IMA, pages 1–147, 2010.
- [17] J. Lavaei and S. H. Low. Zero duality gap in optimal power flow problem. *IEEE Trans. Power Systems*, 27(1):92–107, 2012.
- [18] Neil Robertson and Paul D. Seymour. Graph minors II: Algorithmic aspects of tree-width. *Journal of Algorithms*, 7:309 – 322, 1986.
- [19] Neil Robertson and P.D Seymour. Graph minors. III. Planar tree-width. Journal of Combinatorial Theory, Series B, 36(1):49 – 64, 1984.
- [20] Martin J. Wainwright and Michael I. Jordan. Treewidth-Based conditions for exactness of the Sherali-Adams and Lasserre relaxations. Technical Report 671, University of California, September 2004.

Relative Entropy Convex Programs

Venkat Chandrasekaran

California Institute of Technology

and Parikshit Shah

Yahoo Research

USA

venkatc@caltech.edu and pshah@discovery.wisc.edu

We thank the INFORMS Optimization Society Young Researchers Prize Committee consisting of Sam Burer, Jon Lee, Andrew Schaefer, and Cole Smith for honoring our recent work "Relative Entropy Relaxations for Signomial Optimization" [7]. We begin by discussing the broader context for our work in Section 1, and then surveying the results of [7] in Sections 2 and 3.

1 Relative Entropy from a Computational Perspective

The relative entropy function

$$d(\boldsymbol{
u}, \boldsymbol{\lambda}) = \sum_{i=1}^n \boldsymbol{
u}_i \log(\boldsymbol{
u}_i/\boldsymbol{\lambda}_i), \ \boldsymbol{
u}, \boldsymbol{\lambda} \in \mathbb{R}^n_+$$

plays a significant role in information theory and in statistics — most prominently via error exponents — in the characterization of the performance of a



Parikshit Shah, Venkat Chandrasekaran, and Andrew Schaefer

variety of inferential procedures as well as in proofs of a number of fundamental inequalities [11]. This function also has a notable computational attribute in that it is convex jointly in both of its arguments. Further, convex programs with relative entropy inequalities can be solved efficiently to a desired accuracy via interior-point methods due to the existence of computationally tractable self-concordant barrier functions for the convex function $\nu \log(\nu/\lambda)$ for $\nu, \lambda \geq 0$ [22]. Building on these observations, we initiated a project a few years ago to develop a deeper understanding of the computational properties of relative entropy by investigating the expressive power of relative entropy programs (REPs), which are convex programs consisting of relative entropy inequalities. REPs offer a common generalization of a number of prominent families of convex optimization problems such as geometric programming (GP) in convex form [13, 6], second-order cone programming (SOCP) [4, 19] (and consequently, linear programming), and entropy maximization. The relation between semidefinite programs (SDPs) and REPs is less clear. It is still an open question whether REPs contain SDPs as a special case. In the other direction, SDPs do not contain REPs as a special case; this follows from the observation that the boundary of the constraint set of an REP is not algebraic in general, whereas constraint sets of SDPs have algebraic boundaries.

In recent papers [7, 8], we discuss the utility of REPs for solving a range of problems to which these other classes of convex programs are not directly applicable:

1. Permanent maximization The objective in this problem is to find a matrix with the largest permanent from a given collection. This problem is believed to be computationally intractable in general (indeed, even computing the permanent of a matrix is believed to intractable). We describe an approximation algorithm for the permanent maximization problem based on an REP relaxation [8]. We bound the quality of the approximation provided by our method via Van Der Waerden's inequality. As a parallel, recall that SDPs are useful for maximizing the determinant over an affine section of the cone of symmetric positive semidefinite

matrices [31].

- 2. Robust GPs In a range of applications, one desires a solution of a GP that is robust with respect to uncertainty in the underlying problem parameters; such a program is commonly referred to as a robust GP [3]. We describe the first REP-based method for robust GPs [8], which stands in contrast to previous GP-based reformulations. As a result, our technique enables exact and tractable solutions of robust GPs for a significantly broader class of uncertainty models than those considered in prior work [3, 15]. In recent work Han et al. [14] have employed our REP-based reformulations of robust GPs to optimally allocate resources to control the worst-case spread of epidemics in a network; here the exact network is unknown and it belongs to an uncertainty set.
- 3. Hitting-times in dynamical systems In this problem, we are given a linear dynamical system consisting of a region of feasible starting points and a target set, and the objective is to compute the smallest time required to hit the target set from an arbitrary feasible starting point. Hitting-times are of interest in system analysis and verification [26]. As an example, if a system has the property that the worst-case hitting-time to the target set from an arbitrary point in the initial feasible set is infinite, we have a certificate that the target region is not reachable from certain initial states. On the other hand, if the hitting-time is zero, then we have a proof that the initial set is contained inside the target set. We describe an REP-based approach for computing such hitting-times exactly for certain families of linear dynamical systems [8].
- 4. **Bounds on Signomial Programs** Signomial programs (SPs) are a class of non-convex (and NP-hard) problems that arise in a range of applications. In [7] we describe a family of tractable REP-based convex relaxations for obtaining bounds on SPs.

Software for solving REPs is available (and is constantly updated) at an online repository [1]. In this

article, we survey the results in [7] on REP relaxations for SPs.

2 Relaxations for Signomial Optimization

A signomial is a weighted sum of exponentials composed with linear functionals of a variable $\mathbf{x} \in \mathbb{R}^n$:

$$f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\left\{\alpha^{(j)'} \mathbf{x}\right\}$$
 (1)

Here $\mathbf{c}_{i} \in \mathbb{R}$ and $\boldsymbol{\alpha}^{(j)} \in \mathbb{R}^{n}$ are fixed parameters.¹ An SP is an optimization problem in which a signomial is minimized subject to constraints on signomials, all in a decision variable x [12, 13]. SPs are non-convex in general, and they include NPhard problems as special cases [9]. Consequently, we do not expect to obtain globally optimal solutions of general SPs in a computationally tractable manner. GPs constitute a prominent subclass of SPs in which a posynomial — a signomial with positive coefficients, i.e., the \mathbf{c}_i 's are all positive — is minimized subject to upper-bound constraints on posynomials [6, 13]. GPs are convex optimization problems that can be solved efficiently, and they have been successfully employed in a wide range of applications such as power control in communication systems [9], circuit design [5], approximations to the matrix permanent [18], and the computation of capacities of point-to-point communication channels [10]. However, the additional flexibility provided by SPs via constraints on arbitrary signomials rather than just posynomials is useful in a range of problems to which GPs are not directly applicable. Examples include resource allocation in networks [9], control problems involving chemical processes (see the extensive reference list in [20]), spatial frame design [32], and certain nonlinear flow problems in graphs [25].

A Stylized Application As a further illustration of the ubiquity of SPs in applications, we discuss the discrete choice model which is commonly employed in operations research [2]. The model of interest here is a mixture model of multinomial logits:

 $p_{\beta}(k; \mathbf{w}) \triangleq \text{Prob}[\text{user with attributes } \mathbf{w} \text{ chooses}]$

$$k\text{'th alternative}] = \sum_{i=1}^{\ell} \frac{1}{\ell} \cdot \frac{\exp\left\{\beta^{(i,k)}\mathbf{w}\right\}}{\sum_{k=1}^{m} \exp\left\{\beta^{(i,k)}\mathbf{w}\right\}}.$$

Learning the parameters β of this mixture model from data \mathcal{D} entails the solution of a maximum-likelihood optimization problem:

$$\sup_{\beta} \prod_{(k,\mathbf{w})\in\mathcal{D}} p_{\beta}(k;\mathbf{w}).$$

One can suitably transform this optimization problem to a constrained SP.

Several researchers have developed strategies for optimizing SPs based on variants of branch-and-bound methods [20] as well as heuristic techniques based on successive approximations that can be solved via LP or GP [6, 9, 27, 28]. The framework developed in [7] is qualitatively different as it is based on solving convex optimization problems involving linear and relative entropy functions to obtain guaranteed bounds on the optimal values of SPs. This approach is founded on the insight that the joint convexity of the relative entropy function leads to an effective convex parametrization of certain families of globally nonnegative signomials.

We focus this discussion on the problem of unconstrained minimization of a signomial; our approach to this problem can be adapted to constrained SPs [7]. Central to our development is a view of global minimization that is grounded in duality: globally minimizing a signomial is computationally equivalent to the problem of certifying global nonnegativity of a signomial. In other words, one may rephrase any unconstrained minimization problem as follows:

$$f^* = \inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

=
$$\sup_{\gamma \in \mathbb{R}} \gamma \text{ s.t. } f(\mathbf{x}) - \gamma \ge 0 \ \forall \mathbf{x} \in \mathbb{R}^n.$$

Note that $\gamma = \gamma \exp{\{\mathbf{0}'\mathbf{x}\}}$. Thus, for a given signomial $f(\mathbf{x})$ and for each fixed γ , the constraint in

¹In the literature [13], signomials are typically parametrized somewhat differently as weighted sums of generalized "monomials." A monomial consists of a product of variables, each raised to an arbitrary real power, and the variables only take on positive values. It is straightforward to transform such functions to sums of exponentials of the type (1) discussed in this paper; see [6, 13].

the maximization problem entails checking whether a signomial is globally nonnegative. Of course, as with global minimization, checking global nonnegativity of a signomial is also a computationally intractable problem in general. However, this reformulation suggests a natural strategy for obtaining lower bounds on f^* (note that upper bounds on f^* may be obtained by evaluating f at points in \mathbb{R}^n). Specifically, we devise a sufficient condition for certifying nonnegativity of a signomial, which can be checked by solving a convex REP feasibility problem. Employing this sufficient condition as a surrogate for global nonnegativity in the constraint in the maximization problem above leads to a lower bound on f^* .

To formalize matters, fix an indexed collection of exponents $\left\{\boldsymbol{\alpha}^{(i)}\right\}_{i=1}^{\ell} \subset \mathbb{R}^n$. Our objective is to certify nonnegativity of a signomial $f(\mathbf{x}) = \sum_{i=1}^{\ell} \mathbf{c}_i \exp\left\{\boldsymbol{\alpha}^{(i)'}\mathbf{x}\right\}$. Stated differently, we wish to certify membership of the coefficient vector $\mathbf{c} \in \mathbb{R}^{\ell}$ in the following set:

NNG
$$\left(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)}\right)$$

= $\left\{\mathbf{c} \in \mathbb{R}^{\ell} \mid \sum_{i=1}^{\ell} \mathbf{c}_{i} \exp \left\{\boldsymbol{\alpha}^{(i)}'\mathbf{x}\right\} \ge 0 \ \forall \mathbf{x} \in \mathbb{R}^{n}\right\}$ (2)

Although checking membership in NNG $(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)})$ for a general set of expo-

nents is a computationally intractable problem, certifying nonnegativity of a signomial with all but one coefficient being positive can be accomplished in a computationally tractable manner (for any exponent set). We fix some notation before demonstrating this point: for any $\mathbf{v} \in \mathbb{R}^k$, let $\mathbf{v}_{\setminus i} \in \mathbb{R}^{k-1}$ denote the vector obtained by removing the i'th coordinate from \mathbf{v} .

Consider a signomial $\sum_{j=1}^{\ell} \mathbf{c}_j \exp\left\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\right\}$ with $\mathbf{c}_{\setminus i} \in \mathbb{R}_+^{\ell-1}$. A natural approach to certify the nonnegativity of this signomial is to identify a set of weights $\boldsymbol{\delta} \in \mathbb{R}_+^{\ell}$ with $\mathbf{1}'\boldsymbol{\delta} = 1, \boldsymbol{\delta}_i = 0$ so that:

$$\prod_{j=1, j
eq i}^{\ell} \left(rac{\mathbf{c}_{j} \exp\left\{oldsymbol{lpha}^{(j)'} \mathbf{x}
ight\}}{oldsymbol{\delta}_{j}}
ight)^{oldsymbol{\delta}_{j}} \geq -\mathbf{c}_{i} \exp\left\{oldsymbol{lpha}^{(i)'} \mathbf{x}
ight\}.$$

The existence of such a δ then establishes nonnegativity as a consequence of the following weighted AM/GM inequality:

$$\sum_{j=1, j \neq i}^{\ell} \mathbf{c}_{j} \exp \left\{ \boldsymbol{\alpha}^{(j)'} \mathbf{x} \right\} \geq \prod_{j=1, j \neq i}^{\ell} \left(\frac{\mathbf{c}_{j} \exp \left\{ \boldsymbol{\alpha}^{(j)'} \mathbf{x} \right\}}{\boldsymbol{\delta}_{j}} \right)^{\boldsymbol{\delta}_{j}}.$$

The results in [7] give an REP parametrization to identify such a set of weights, and more broadly, to characterize the set of all coefficients $\mathbf{c} \in \mathbb{R}^{\ell}$ with $\mathbf{c}_{\setminus i} \in \mathbb{R}^{\ell}_{+}$ so that the associated signomial is nonnegative:

$$AGE\left(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(i-1)}, \boldsymbol{\alpha}^{(i+1)}, \dots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(i)}\right) = NNG\left(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)}\right) \cap \left\{\mathbf{c} \in \mathbb{R}^{\ell} \mid \mathbf{c}_{\setminus_{i}} \in \mathbb{R}_{+}^{\ell-1}\right\} \\
= \left\{\mathbf{c} \in \mathbb{R}^{\ell} \mid \mathbf{c}_{\setminus_{i}} \in \mathbb{R}_{+}^{\ell-1}, \ \exists \boldsymbol{\nu} \in \mathbb{R}^{\ell} \text{ s.t. } \boldsymbol{\nu}_{\setminus_{i}} \in \mathbb{R}_{+}^{\ell-1}, d(\boldsymbol{\nu}_{\setminus_{i}}, e\mathbf{c}_{\setminus_{i}}) \leq \mathbf{c}_{i}, \sum_{j=1}^{\ell} \boldsymbol{\alpha}^{(j)} \boldsymbol{\nu}_{j} = 0, \boldsymbol{\nu}_{i} = -\mathbf{1}' \boldsymbol{\nu}_{\setminus_{i}}\right\}$$
(3)

Here e is Euler's constant. The reason for the terminology $AGE(\cdot;\cdot)$ is due to the connection with the AM/GM inequality. This characterization sug-

gests a natural sufficient condition for membership in NNG $(\alpha^{(1)}, \dots, \alpha^{(\ell)})$:

$$SAGE\left(\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(\ell)}\right)\triangleq \oplus_{i=1}^{\ell}AGE\left(\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(i-1)},\boldsymbol{\alpha}^{(i+1)},\ldots,\boldsymbol{\alpha}^{(\ell)};\boldsymbol{\alpha}^{(i)}\right)\subseteq NNG\left(\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(\ell)}\right). \tag{4}$$

SAGE here denotes "sums of AGE." Checking membership in SAGE $(\alpha^{(1)}, \ldots, \alpha^{(\ell)})$ is computationally efficient based on (3), and consequently, replacing the nonnegativity condition in (2) by the SAGE condition leads to a tractable convex lower bound on the infimum of a signomial. We refer to this relaxation as the SAGE relaxation.

Example 1. We consider signomials in three variables with seven terms of the form $f(\mathbf{x}) = \sum_{j=1}^{7} \mathbf{c}_{j} \exp{\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}}$ with $\mathbf{x} \in \mathbb{R}^{3}$ and the following parameters fixed: $\mathbf{c}_{1} = 0$, $\mathbf{c}_{2} = \mathbf{c}_{3} = \mathbf{c}_{4} = 10$, $\boldsymbol{\alpha}^{(1)} = (0,0,0)'$, $\boldsymbol{\alpha}^{(2)} = (10.2,0,0)'$, $\boldsymbol{\alpha}^{(3)} = (0,9.8,0)'$, and $\boldsymbol{\alpha}^{(4)} = (0,0,8.2)'$. The exponents $\boldsymbol{\alpha}^{(5)}, \boldsymbol{\alpha}^{(6)}, \boldsymbol{\alpha}^{(7)} \in \mathbb{R}^{3}$ are chosen to be random vectors with entries distributed uniformly in [0,3], and the coefficients $\mathbf{c}_{5}, \mathbf{c}_{6}, \mathbf{c}_{7}$ are chosen to be random Gaussians with mean 0 and standard deviation 10. We employ this construction — relatively large exponents $\boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(3)}, \boldsymbol{\alpha}^{(4)}$ in comparison to $\boldsymbol{\alpha}^{(5)}, \boldsymbol{\alpha}^{(6)}, \boldsymbol{\alpha}^{(7)}$, and the corresponding positive coefficients $\mathbf{c}_{2} = \mathbf{c}_{3} = \mathbf{c}_{4} = 10$ — to obtain signomials that are bounded below. An example of a signomial generated in this manner is:

$$f(\mathbf{x}) = 10 \exp\{10.2\mathbf{x}_1\} + 10 \exp\{9.8\mathbf{x}_2\} + 10 \exp\{8.2\mathbf{x}_3\}$$

$$- 14.6794 \exp\{1.5089\mathbf{x}_1 + 1.0981\mathbf{x}_2 + 1.3419\mathbf{x}_3\}$$

$$- 7.8601 \exp\{1.0857\mathbf{x}_1 + 1.9069\mathbf{x}_2 + 1.6192\mathbf{x}_3\}$$

$$+ 8.7838 \exp\{1.0459\mathbf{x}_1 + 0.0492\mathbf{x}_2 + 1.6245\mathbf{x}_3\}.$$
(5)

The SAGE relaxation applied to $f(\mathbf{x})$ gives the lower bound $f_{\mathrm{SAGE}} \approx -0.9747$. By applying a technique presented in [7], we obtain the point $\mathbf{x}^* = (-0.3020, -0.2586, -0.4010)'$ with $f(\mathbf{x}^*) \approx -0.9747$. Consequently, the lower bound f_{SAGE} is tight in this case. As another instance of the construction described here, consider the randomly generated signomial:

$$f(\mathbf{x}) = 10 \exp\{10.2\mathbf{x}_1\} + 10 \exp\{9.8\mathbf{x}_2\} + 10 \exp\{8.2\mathbf{x}_3\}$$

$$+ 7.5907 \exp\{1.9864\mathbf{x}_1 + 0.2010\mathbf{x}_2 + 1.0855\mathbf{x}_3\}$$

$$- 10.9888 \exp\{2.8242\mathbf{x}_1 + 1.9355\mathbf{x}_2 + 2.0503\mathbf{x}_3\}$$

$$- 13.9164 \exp\{0.1828\mathbf{x}_1 + 2.7772\mathbf{x}_2 + 1.9001\mathbf{x}_3\}.$$
(6)

In this case the SAGE relaxation gives a lower bound of -1.426, which turns out to not be tight. More generally, we generated 80 random signomials according

to the above description, and the SAGE relaxation was tight in 63% of the cases, while there was a gap in the remaining cases. (The results in [7] describe techniques for verifying when the SAGE relaxation is tight.)

We build on the preceding developments in [7] in three respects. First, we describe a tractable REP approach to obtain lower bounds on constrained SPs in an analogous manner to the SAGE relaxation. Second, we describe a principled framework to obtain a family of increasingly tighter lower bounds for general (constrained and unconstrained) SPs by solving hierarchies of successively larger convex programs based on relative entropy optimization; these hierarchies are derived by considering a sequence of tighter nonnegativity certificates for a signomial over a set defined by signomial constraints. Lastly, we develop a dual viewpoint that is intimately connected with representing the convex hull of the $\left\{ \left(\exp \left\{ \boldsymbol{\alpha}^{(1)'} \mathbf{x} \right\}, \dots, \exp \left\{ \boldsymbol{\alpha}^{(\ell)'} \mathbf{x} \right\} \right) \mid \mathbf{x} \in \mathbb{R}^n \right\},\,$ which in turn suggests a natural rounding procedure to obtain approximate minimizers of SPs from REP relaxations.

The hierarchy of convex relaxations that we describe in [7] has several notable features. First, GPs are solved exactly by the first level in this hierarchy; thus, our hierarchy has the desirable property that "easy problem instances remain easy." Second, the family of lower bounds is invariant under a natural transformation of the problem data. Specifically, the optimal value of an SP remains unchanged under the application of a non-singular linear transformation simultaneously to all the parameters $\alpha^{(j)}$ that appear in the exponents of the signomials in an SP (both in the objective and in the constraints). The hierarchy of relative entropy relaxations described in [7] leads to bounds that are invariant under such transformations. Third, it is desirable that any procedure for obtaining lower bounds of the optimal value of an SP be robust to small perturbations of the exponents $\alpha^{(j)}$ in an SP. As discussed in [7], this is also a feature of our proposed approach. Fourth, for broad families of SPs our approach leads to a convergent sequence of lower bounds, i.e., our hierarchy approximates the optimal value of the SP arbitrarily well from below. Such a property of a hierarchy is usually referred to as *completeness*.

3 Connections to Polynomial Optimization

A prominent example of hierarchies of tractable convex programs being employed for intractable problems is in the setting of polynomial optimization problems, for which SDP relaxations have been developed based on sum-of-squares techniques [17, 23, 24]. However, those methods are not directly relevant to SPs for several reasons, and we briefly highlight these distinctions here. As with SPs, polynomial optimization problems are also non-convex in general, and they include families of NP-hard Parrilo [23, 24] and Lasserre [17] deproblems. scribe computationally feasible methods to obtain lower bounds for polynomial optimization problems. These techniques rely on nonnegativity certificates for polynomials based on sum-of-squares decompositions [16, 21, 30], and the observation by Shor [29] that checking if a polynomial is a sum-of-squares can be recast as an SDP feasibility problem.

SPs with rational exponents can be transformed to polynomial optimization problems over the nonnegative orthant by clearing denominators in the exponents. This transformation generally leads to polynomials of very large degrees, thus making sum-ofsquares techniques ill-suited for general SPs. Concretely, consider the signomial $f_{\alpha,\beta}(x) = \exp\{dx\} +$ $\alpha \exp\{x\} + \beta$ with $x \in \mathbb{R}$. For each fixed $\alpha, \beta \in \mathbb{R}$ and for each $d \in \mathbb{Z}_{++}$, one can efficiently check whether $f_{\alpha,\beta}$ is globally nonnegative using the REP feasibility problem (3) (as β must be nonnegative for $f_{\alpha,\beta}$ to be globally nonnegative, we are left with a signomial with at most one negative coefficient). In particular, $f_{\alpha,\beta}$ is globally nonnegative if and only if $(\alpha, \beta) \in \{(a, b) \in \mathbb{R} \times \mathbb{R}_+ \mid \exists \boldsymbol{\nu} \in \mathbb{$ \mathbb{R}^2_+ s.t. $D(\nu, e(1, b)') \le a, (d-1)\nu_1 = \nu_2$. Note that the size of this description does not grow with d. One can also apply the transformation $y^2 \leftarrow \exp x$ to transform the question of nonnegativity of the signomial $f_{\alpha,\beta}$ to one of nonnegativity of the the polynomial $\tilde{f}_{\alpha,\beta}(y) = y^{2d} + \alpha y^2 + \beta$ for $y \in \mathbb{R}$. However, the corresponding nonnegativity certificates for $\tilde{f}_{\alpha,\beta}$ based on sum-of-squares methods grow in size with increasing d (see [17, 23, 24] for more details about constructing these certificates), which in turn entails the solution of large SDPs. Underlying this discussion is the insight that REP-based relaxations (based on the AM/GM inequality) offer a different proof system than SDP relaxations (based on sum-of-squares methods) for certifying nonnegativity.

More broadly, relative entropy relaxations for SPs also have the virtue that the bounds they provide are generally robust to small perturbations of the exponents

Example 2. Consider the following signomial obtained by perturbing the exponents of the signomial (5) in Example 1 (but leaving the coefficients unchanged):

```
\begin{split} f(\mathbf{x}) =& 10 \exp\{10.2070\mathbf{x}_1 + 0.0082\mathbf{x}_2 - 0.0039\mathbf{x}_3\} \\ &+ 10 \exp\{-0.0081\mathbf{x}_1 + 9.8024\mathbf{x}_2 - 0.0097\mathbf{x}_3\} \\ &+ 10 \exp\{0.0070\mathbf{x}_1 - 0.0156\mathbf{x}_2 + 8.1923\mathbf{x}_3\} \\ &- 14.6794 \exp\{1.5296\mathbf{x}_1 + 1.0927\mathbf{x}_2 + 1.3441\mathbf{x}_3\} \\ &- 7.8601 \exp\{1.0750\mathbf{x}_1 + 1.9108\mathbf{x}_2 + 1.6339\mathbf{x}_3\} \\ &+ 8.7838 \exp\{1.0513\mathbf{x}_1 + 0.0571\mathbf{x}_2 + 1.6188\mathbf{x}_3\}. \end{split}
```

The SAGE lower bound is tight for the perturbed signomial specified here, with the optimal value being equal to -0.9458, and the optimal solution being $\mathbf{x}^* = (-0.3016, -0.2605, -0.4013)'$. Recall that the SAGE lower bound was also tight for the signomial (5) of Example 1, with the optimal value being equal to -0.9747 and the optimal solution being $\mathbf{x}^* = (-0.3020, -0.2586, -0.4010)'$. Hence, this example provides numerical evidence for the robustness of our relative entropy relaxation methods with respect to small perturbations of the exponents.

Approaching this SP as a polynomial optimization problem clearly illustrates the shortcomings of that viewpoint, as small changes in the exponents can lead to very different polynomials after clearing denominators in the exponents. In turn, the quality of the bounds and amount of computation required to obtain them via SDP relaxations based on sum-of-squares techniques can vary dramatically for small changes in the exponents. However, relative entropy

relaxations for SPs based on SAGE decompositions are well-behaved under such small perturbations of the exponents.

Acknowledgements

We would like to thank Pablo Parrilo, Bernd Sturmfels, and Bill Helton for helpful discussions at various stages throughout this project.

REFERENCES

- [1] https://github.com/paryxyt/REO.
- [2] M. Ben-Avika and S. Lerman. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, 1985.
- [3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. Mathematics of Operations Research, 23:769– 805, 1998.
- [4] A. Ben-Tal and A. Nemirovskii. Lectures on Modern Convex Optimization. Society for Industrial and Applied Mathematics, 2001.
- [5] S. Boyd, S. J. Kim, D. Patil, and M. Horowitz. Digital Circuit Optimization via Geometric Programming. Operations Research, 53:899–932, 2005.
- [6] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi. A Tutorial on Geometric Programming. *Optimization and Engineering*, 8:67–127, 2007.
- [7] V. Chandrasekaran and P. Shah. Relative entropy relaxations for signomial optimization. SIAM Journal on Optimization, 26:1147–1173, 2016.
- [8] V. Chandrasekaran and P. Shah. Relative entropy optimization and its applications. *Mathematical Pro*gramming, 161:1–32, 2017.
- [9] M. Chiang. Geometric Programming for Communication Systems. Foundations and Trends in Communications and Information Theory, 2:1–154, 2005.
- [10] M. Chiang and S. Boyd. Geometric Programming Duals of Channel Capacity and Rate Distortion. *IEEE Transactions on Information Theory*, 50:245–258, 2004.
- [11] T. Cover and J. Thomas. Elements of Information Theory. Wiley, 2006.
- [12] R. J. Duffin and E. L. Peterson. Geometric Programming with Signomials. *Journal of Optimization The*ory and Applications, 11:3–35, 1973.

- [13] R. J. Duffin, E. L. Peterson, and C. M. Zener. Geometric Programming: Theory and Application. John Wiley and Sons, 1967.
- [14] S. Han, V. M. Preciado, C. Nowzari, and G. J. Pappas. Data-driven allocation of vaccines for controlling epidemic outbreaks. *IEEE Transactions on Network* Science and Engineering, 2:127–138, 2015.
- [15] K. L. Hsiung, S. J. Kim, and S. Boyd. Tractable approximate robust geometric programming. *Optimization* and Engineering, 9:95–118, 2008.
- [16] J. L. Krivine. Anneaux Prordonns. Journal d'Analyse Mathmatique, 12:307–326, 1964.
- [17] J. B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. SIAM Journal on Optimization, 11:796–817, 2001.
- [18] N. Linial, A. Samorodnitsky, and A. Wigderson. A Deterministic Strongly Polynomial Algorithm for Matrix Scaling and Approximate Permanents. *Combi*natorica, 20, 2000.
- [19] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Its Applications*, 284:193–228, 1998.
- [20] C. D. Maranas and C. A. Floudas. Global Optimization in Generalized Geometric Programming. Computers and Chemical Engineering, 21:351–369, 1997.
- [21] M. Marshall. Positive Polynomials and Sums of Squares. American Mathematical Society, 2008.
- [22] Y. Nesterov and A. Nemirovski. Interior-Point Polynomial Algorithms in Convex Programming. Society of Industrial and Applied Mathematics, 1994.
- [23] P. A. Parrilo. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. PhD thesis, California Institute of Technology, 2000.
- [24] P. A. Parrilo. Semidefinite Programming Relaxations for Semialgebraic Problems. *Mathematical Program*ming, 96:293–320, 2003.
- [25] E. L. Peterson. Geometric Programming. SIAM Review, 18:1–51, 1976.
- [26] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. Hybrid Systems: Computation and Control, Springer Lecture Notes in Computer Science, 2004.
- [27] P. Shen. Linearization Method of Global Optimization for Generalized Geometric Programming. Applied Mathematics and Computation, 162:353–370, 2005.

- [28] H. D. Sherali. Global Optimization of Nonconvex Polynomial Programming Problems having Rational Exponents. *Journal of Global Optimization*, 12:267– 283, 1998.
- [29] N. Z. Shor. Class of Global Minimum Bounds of Polynomial Functions. Cybernetics, 23:731–734, 1987.
- [30] G. Stengle. A Nullstellensatz and a Positivstellensatz in Semialgebraic Geometry. *Annals of Mathematics*, 207:87–97, 1974.
- [31] L. Vandenberghe, S. Boyd, and S. Wu. Determinant maximization with linear matrix inequality constraints. SIAM Journal on Matrix Analysis and Applications, 19:499–533, 1998.
- [32] K. Yun and C. Xi. Second-order Method of Generalized Geometric Programming for Spatial Frame Optimization. Computer Methods in Applied Mechanics and Engineering, 141:117–123, 1997.

Update on the 2018 INFORMS Optimization Society Conference

Steve Billups (University of Colorado Denver), Manuel Laguna (University of Colorado Boulder), and Alexandra Newman (Colorado School of Mines) are co-organizers of the 2018 INFORMS Optimization Society Conference, slated for March 23–25 on the University of Colorado Denver's downtown campus. Our "Mountains of Optimization" theme emphasizes many important, contemporary aspects of optimization. Plenary speakers include Shabbir Ahmed (Georgia Tech), Marcos Goycoolea (Universidad Adolfo Ibanez), Moritz Hardt (Google, UC Berkeley), Illya Hicks (Rice University), Karla Hoffman (George Mason University), John Hooker (Carnegie Mellon University), and Sven Leyffer (Argonne National Labs). We are working hard to make this an exciting conference! More details are forthcoming on the website.



University of Colorado Denver (image used with permission of the University of Colorado)



Denver, CO (image used with permission of the University of Colorado)

The Vision for INFORMS Journal on Optimization

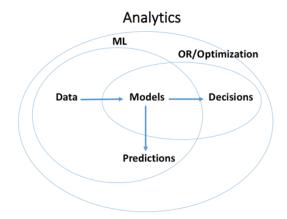
Dimitris Bertsimas

Editor in Chief of INFORMS Journal on Optimization dbertsim@mit.edu

Optimization has a long and distinguished history and has been at the core of INFORMS from the founding of ORSA in the early 1950s. I would argue that without the Simplex method, the key algorithm for Linear Optimization, ORSA might not have existed.

Historically optimization and operations research have focused on the process of building models to derive optimal decisions. While models have been the protagonist, data has been a secondary actor. In recent years we have witnessed an unprecedented growth of data in electronic form and methods, typically associated with the area of machine learning, for prediction. I feel that one of the largest opportunities of the field of optimization is to embrace data in a protagonist role and combine it with machine learning. The vision of the future for the field of optimization I have is summarized in the graph below.

The definition of the field of Analytics I have is the science that starts with **data**, builds **models** to derive optimal **decisions** that add **value**.



Areas of interest for the journal: With these ideas in mind my vision of the new journal includes the following new areas:

• Data driven optimization;

- Optimization methods in machine learning;
- Exciting real world applications of optimization.

In addition, the journal will publish papers in more traditional areas of optimization.

A non-exclusive list of methodologies that the journal covers is: convex optimization (including linear optimization); general purpose nonlinear optimization; discrete optimization (combinatorial, integer, mixed integer optimization); optimization under uncertainty (dynamic programming, stochastic programming, robust optimization, simulation based optimization); infinite dimensional optimization; online optimization.

Especially welcomed are contributions studying new and significant applications. A non-exclusive list of application areas that the journal will cover includes: health care; inventory and supply chain management; logistics; revenue management and pricing; energy; the internet; interfaces with computer science; and finance.

There are other leading journals in optimization, including Mathematical Programming and SIAM Journal on Optimization, that are more focused in the methodology of optimization. The aspiration of the new journal is to add a new quality outlet that in addition to methodological papers will also publish papers that place particular emphasis on data driven optimization, interface with machine learning, and real world applications of optimization.

The journal will also invite reviews of new developments in optimization as well as new emerging areas of interest to a wider audience.

Criteria for publication: Accepted papers are expected to score highly in at least one of the following questions:

- (a) How new and significant is the application studied?
- (b) How original/creative is the optimization modeling?
- (c) How original/creative is the optimization methodology?
- (d) How significant is the impact in practice?

Addressing delays and negativity: There are two trends in our field that the journal aspires to reverse: (a) long delays, (b) negativity in the reviewing process. To address these trends, the journal will try

to solicit reviews from three reviewers of each paper and aspires to get a response to the authors within 3 months of submission. The journal will not reject a paper with one (out of three) negative report, unless there are serious questions on correctness. I very much believe in finding what is good in an idea and trying to improve it.

The review process: Authors submit a paper to the journal and recommend 3 AEs and 6 reviewers. If the paper is not sent to an AE or is not sent to reviewers, we will communicate this negative decision within 2 weeks. Assuming the paper enters the reviewing process, I will assign the paper to an AE (but not necessarily one of the 3 chosen) trying to address issues of balance of the workload among the AEs. The AE will solicit 3 reviews, making a recommendation within 3 months. A positive recommendation needs to justify which of the four criteria for publication the paper answers successfully to deserve publication. I will then make the final decision. As already mentioned, I aspire that the total time to get back to the authors will be 3 months.

Editorial board: The inaugural editorial board consists of

- Shabbir Ahmed
- John Birge
- Dick den Hertzog
- Brian Denton
- Laurent El Ghaoui
- Vivek Farias
- Dan Iancu
- Garud Iyengar
- Patrick Jaillet
- Andrea Lodi
- David Morton
- Georgia Perakis
- Jim Renegar
- Suvrajeet Sen
- David Shmoys
- Melvyn Sim
- Ben van Roy

It is anticipated that the journal will start accepting submissions by the end of July, 2017.

Nominations for Society Prizes Sought

The INFORMS Optimization Society awards four prizes annually at the INFORMS annual meeting. We seek nominations (including self-nominations) for each of them, due by June 15, 2017 (please note the earlier-than-usual deadline, due to the timing of the 2017 Annual Meeting). Each of the four awards includes a cash amount of US\$1,000 and a citation plaque. The award winners will be invited to give a presentation in a special session sponsored by the Optimization Society during the INFORMS annual meeting in Houston, TX, in October 2017 (the winners will be responsible for their own travel expenses to the meeting). Award winners are also asked to contribute an article about their awardwinning work to the Optimization Society newsletter.

The four awards are listed below. Additional information on the awards, and nomination instructions, can be found on the society website (http://connect.informs.org/optimizationsociety/prizes).

Please see the website for nomination instructions. Inquiries should be sent directly via email to the chair of the corresponding prize committee.

The Khachiyan Prize is awarded for outstanding lifetime contributions to the field of optimization by an individual or team. The topic of the contribution must belong to the field of optimization in its broadest sense. Recipients of the INFORMS John von Neumann Theory Prize or the MPS/SIAM Dantzig Prize in prior years are not eligible for the Khachiyan Prize. The prize may be awarded once in a lifetime to any individual. Nominations should be submitted to the Chair of the committee.

The prize committee for this year's Khachiyan Prize is as follows:

- Gerald Brown (Chair) GBrown@nps.edu
- Bill Cook
- Andrzej Ruszczyński
- Yinyu Ye

The Farkas Prize is awarded for outstanding contributions by a mid-career researcher to the field of optimization, over the course of their career. Such contributions could include papers (published or submitted and accepted), books, monographs, and software. The awardee will be within 25 years of their terminal degree as of January 1 of the year of the award. The prize serves as an esteemed recognition of colleagues in the middle of their career. The prize may be awarded at most once in their lifetime to any person. A nomination shall consist of: (i) a letter of nomination, not exceeding two pages, summarizing the nominee's contributions with explanations of their importance and impact; (ii) a curriculum vitae for the nominee, not exceeding four pages; and (iii) two support letters, each not exceeding two pages. These letters can be sent directly to the committee chair or to the nominator, to be included in the nomination package.

The prize committee for this year's Farkas Prize is as follows:

- Margaret Wright (Chair) mhw@cims.nyu.edu
- Patrick Jaillet
- Zhi-Quan (Tom) Luo
- Pascal Van Hentenryck

The **Prize for Young Researchers** is awarded to one or more young researcher(s) for an outstanding paper in optimization. The paper must be published in, or submitted to and accepted by, a refereed professional journal within the four calendar years preceding the year of the award. All authors must have earned their most recent degree within the eight calendar years preceding the year of the award or be enrolled in a degree-granting program. The prize serves as an esteemed recognition of promising colleagues who are at the beginning of their academic or industrial career. Nominations should be submitted to the Chair of the committee.

The prize committee for this year's Prize for Young Researchers is as follows:

- Michael Ferris (Chair) ferris@cs.wisc.edu
- Jonathan Eckstein
- Simge Küçükyavuz
- Suvrajeet Sen

The **Student Paper Prize** is awarded to one or more student(s) for an outstanding paper in optimization that is submitted to and received, or published in a refereed professional journal no more than three years before the closing date of nomination. Every nominee/applicant must be a student on the first of January of the year of the award. The prize serves as an esteemed recognition of promising students who are looking for an academic or industrial career. A complete entry consists of: (i) an electronic PDF copy of the paper; (ii) an electronic PDF file of a letter signed by all co-authors attesting that the majority of the work was done by the student(s); (iii) an electronic PDF file of a nomination letter attesting that the eligibility conditions have been satis field by the entrant(s) and the paper. Nominations should be submitted via https://easychair.org/ conferences/?conf=iosstudentprize2017.

The prize committee for this year's Student Paper Prize is as follows:

- Vineet Goyal (Chair) vgoyal@ieor.columbia.edu
- Kiavash Kianfar
- Javier Peña
- Ermin Wei

Nominations of Candidates for Society Officers Sought

Four Society Vice-Chairs will be completing their two-year terms in 2017: Hande Benson, Amitabh Basu, Sauleh Siddiqui, and Austin Buchanan. Burcu Keskin will be completing her term as the Secretary/Treasurer as well. We would like to thank these officers for their work!

We are currently seeking nominations of candidates for the following positions:

- Vice-Chair for Computational Optimization and Software
- Vice-Chair for Integer and Discrete Optimization
- Vice-Chair for Linear and Conic Optimization
- Vice-Chair for Network Optimization
- Secretary/Treasurer

Self-nominations for all of these positions are encouraged.

Vice-Chairs serve a two-year term. According to Society Bylaws, "The main responsibility of the Vice Chairs will be to help INFORMS Local Organizing committees identify cluster chairs and/or session chairs for the annual meetings. In general, the Vice Chairs shall serve as the point of contact with their sub-disciplines."

The Secretary/Treasurer serves a two-year term. According to Society Bylaws, "The Secretary/Treasurer shall conduct the correspondence

of the OS, keep the minutes and records of the Society, maintain contact with INFORMS, receive reports of activities from those Society Committees that may be established, conduct the election of officers and Members of Council for the OS. make arrangements for the regular meetings of the Council and the membership meetings of the OS. As treasurer, he/she shall also be responsible for disbursement of the Society funds as directed by the OS Council, prepare and distribute reports of the financial condition of the OS, help prepare the annual budget of the Society for submission to INFORMS. It will be the responsibility of the outgoing Secretary/Treasurer to make arrangements for the orderly transfer of all the Society's records to the person succeeding him/her." Secretary/Treasurer is allowed to serve at most two consecutive terms.

Additional details on officer responsibilities and elections can be found in**Bylaws** at http://connect.informs.org/ optimizationsociety/aboutios/bylaws

Please send your nominations or self-nominations to Burcu Keskin (bkeskin@cba.ua.edu), including contact information for the nominee, by June 30, 2017. Online elections will begin in mid-August, with new officers will assume their duties on January 1st, 2018.