INFORMS Optimization Society
2019 Khachiyan Prize

Expository Article:
My Road in the Field of Optimization

Masakazu Kojima
Research Scientist
Department of Industrial and Systems Engineering
Chuo University

Professor Emeritus
Tokyo Institute of Technology
kojima@is.titech.ac.jp

informs. OPTIMIZATION

Published online by the INFORMS Optimization Society on November 3, 2020.
http://connect.informs.org/optimizationsociety/home

I feel very honored to have received the 2019 Khachiyan Prize from the INFORMS Optimization Society. My deep gratitude goes to the members of the selection committee, Robert Vanderbei, Don Goldfarb, Jean Lasserre, and James Renegar.

It is not just my sole fortune and effort that have brought me here. Winning this award would not have been possible without the inspiration and help I have received from my colleagues and co-workers, for whom I have the deepest respect, and from whom I have derived the strength to challenge myself at each stage of my academic career of almost 50 years.

I started to work for the Tokyo Institute of Technology in 1975, having been invited by Professor Hidenori Morimura two years after I received the Doctor of Engineering degree from Keio University. I am profoundly indebted to him for giving me a chance to have worked at Tokyo Institute of Technology, where I spent more than a half of my life.

Many of my research works were done jointly with others. In particular, I would like to thank my collaborators on the study of primal-dual interior-point method for linear programs, Shinji Mizuno, Akiko Yoshise, Toshihito Noma and Nimrod Megiddo. Nimrod invited us several times to his place, IBM Almaden Research Center. I still have wonderful memories of working jointly with Nimrod there. I have been fortunate to work with many excellent collaborators. I really want to share this honor with them.

1 At the starting line

My undergraduate major was industrial engineering and management science at Keio University. During my senior year, I joined Professor Tomoharu Sekine's group for my research. Mathematics and computer science both were my favorite subjects, so I decided to study operations research, in particular, optimization, for my research subject. I also liked my advisor's personality. My advisor had a very unique style when supervising his students. He bombarded the students with many questions when they gave talks on their research subjects (*e.g.*, the duality theorem of LPs), but he never offered any answers. He just waited for us to find the answers by ourselves, while reading comic books or weekly magazines on some occasions. He often told us to explain by drawing a figure of the subject, which I think implicitly meant that if we really understand the essential, we can illustrate it with a figure. This is one practice that I find useful until today. I learned a lot from him, even though he did not directly teach me. An important lesson I learned from him was "*let students find solutions by themselves*" when supervising students. But I soon realized that waiting for them to find answers on their own is often difficult, as it takes so long and telling them what we know may be natural for us. While working in Tomoharu Sekine's group, I found joy in conducting research, a feeling that I never had during my previous years at Keio.

I went on to graduate school at Keio University. At the time, complementarity theory and computing fixed points was one of the most popular research subjects in optimization. In particular, piecewise linear homotopy methods had been extensively studied. I clearly remember the fascination I felt when I found the paper [21] by Lemke and Howson. It was one of the cornerstone papers on the subject, which proposed the so-called Lemke's method for solving the linear complementarity problem induced from bimatrix games. The paper motivated me to work on the subject, and then led me to choose "a computational method for solving the nonlinear complementarity problem" for my doctoral thesis.

Later in my career, I worked on many subjects, among them, I would like to discuss the following three subjects in this article.

Stability of stationary solutions of nonlinear programs — Section 2

Primal-dual interior-point algorithms — Section 3
Exploiting sparsity in semidefinite programs — Section 4

2 Stability of stationary solutions of nonlinear programs

When I was at Northwestern University and University of Wisconsin in 1978 as a postdoc, I worked on this subject. I am greatly indebted to Romesh Saigal and Steve Robinson for providing me with the postdoc positions.

Let $\mathcal{F}$ denote the set of twice continuously differentiable maps from $\mathbb{R}^n$ into $\mathbb{R}^{1+\ell}$. For each $\boldsymbol{f} = (f_0, \ldots, f_\ell) \in \mathcal{F}$, consider a nonlinear program

$$\mathrm{P}(\boldsymbol{f}): \text{ minimize } f_0(\boldsymbol{x}) \text{ subject to } f_j(\boldsymbol{x}) \leq 0 \ (j = 1, \ldots, \ell).$$

Then the Karush-Kuhn-Tucker (KKT) condition for $\mathrm{P}(\boldsymbol{f})$ is given by

$$\begin{aligned} &\nabla f_0(\boldsymbol{x}) + \sum_{j=1}^{\ell} y_j \nabla f_j(\boldsymbol{x}) = \mathbf{0}, \\ &y_j \geq 0,\ z_j = -f_j(\boldsymbol{x}) \geq 0 \text{ and } y_j z_j = 0 \ (j = 1, \ldots, \ell). \end{aligned}$$

If $\boldsymbol{x}$ satisfies the condition above for some $(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{2\ell}$, $\boldsymbol{x}$ is called a stationary solution of $\mathrm{P}(\boldsymbol{f})$. It is well-known that if a local minimizer $\boldsymbol{x}$ of $\mathrm{P}(\boldsymbol{f})$ satisfies a certain assumption, which is often called a constraint qualification, then $\boldsymbol{x}$ is a stationary solution of $\mathrm{P}(\boldsymbol{f})$.

In my paper [16], I introduced a new notion of stability, *the strong stability* for stationary solutions, and discussed its fundamental characterization. In particular, a necessary and sufficient condition for the strong stability was given under the Mangasarian and Fromovitz constraint qualification. It can be summarized as stating that an isolated stationary solution $\boldsymbol{x}$ of $\mathrm{P}(\boldsymbol{f})$ is strongly stable if there exists a unique stationary solution of $\mathrm{P}(\boldsymbol{f}+\boldsymbol{g})$ which is continuous with respect to a sufficiently small (up to the second derivatives) perturbation $\boldsymbol{g} \in \mathcal{F}$ in a small neighborhood of $\boldsymbol{x}$. A distinguishing feature of the strong stability is that its definition is independent of any specific parametrization of the objective function and the constraint functions.

To characterize a strongly stable stationary solution of $\mathrm{P}(\boldsymbol{f})$, the paper [16] first reduced the KKT condition to a system of PC^1 (piecewise continuously differentiable) equations

$$\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) \equiv \begin{pmatrix} \nabla f_0(\boldsymbol{x}) + \sum_{j=1}^{\ell} y_j^+ \nabla f_j(\boldsymbol{x}) \\ y_1^- - f_1(\boldsymbol{x}) \\ \cdots \\ y_\ell^- - f_\ell(\boldsymbol{x}) \end{pmatrix} = \mathbf{0}.$$

Here $y_j^+ = \max\{0, y_j\}$ and $y_j^- = \min\{0, y_j\}$ $(j =, 1, \ldots, m)$. Then $\boldsymbol{x}$ is a stationary solution of $\mathrm{P}(\boldsymbol{f})$ if and only if $(\boldsymbol{x}, \boldsymbol{y})$ is a solution of the system of PC^1 equations $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ for some $\boldsymbol{y} \in \mathbb{R}^\ell$. This conversion enables us to apply some powerful mathematical tools and results from algebraic topology, which I had studied in connection with the piecewise-linear approximations of PC^1 mappings in [15], to the characterization of the strong stability; I utilized the degree theory of mappings in [16].

Around the same time, Robinson [27] proposed the strong regularity for generalized equations independently. The two concepts, the strong stability and the strong regularity are closely related, and they both played a fundamental and important role in the development of the stability analysis of various mathematical programming problems.

3 The primal-dual interior-point method for LPs and SDPs

In 1984, Karmarkar's polynomial-time interior-point method (abbreviated as IPM) [9] for LPs made a dramatic impact on the optimization society. I immediately started to work on this subject as many researchers did. A few years later, a variety of IPMs based on different ideas were proposed. When I wrote the paper [18] jointly with Mizuno and Yoshise on the primal-dual interior-point method (abbreviated as PDIPM), some important notions that would play essential roles in the development of IPMs, such as the barrier function, the Newton method, and the analytic center of a polytope, had been already introduced in some IPMs. Those IPMs generate a sequence of interior feasible solutions of either the primal LP or the dual LP, but not both simultaneously.

Unlike the simplex method, which was fully supported by the powerful duality theory for LPs, the IPM had not effectively utilized the duality theory at the time. My joint work with Mizuno and Yoshise aimed primarily to incorporate duality theory effectively into the IPM. We started from Megiddo's work in 1989 [22] that showed the existence of the smooth curve in the interior of the Cartesian product of the primal and dual feasible region, which converges to a primal-dual pair of optimal solutions. The smooth curve was later called the central trajectory.

The basic idea in the PDIPM [18] is to trace the central trajectory with

($\star$) a repeated application of the Newton iterations toward the central trajectory

in a neighborhood of the trajectory, which shrinks as the trajectory approaches the primal-dual pair of optimal solutions. We showed that our PDIPM attained an optimal solution in $O(nL)$ Newton iterations, where n denotes the number of variables in a primal LP to be solved and L is the number of bits required to represent the data of the LP. In our subsequent paper [17], we extended the PDIPM to a class of monotone linear complementarity problems, which includes convex quadratic programs, and improved the iteration complexity from $O(nL)$ to $O(\sqrt{n}L)$ by choosing a different neighborhood of the central trajectory. Around the time, Monteiro and Adler [24] independently proposed a PDIPM for convex quadratic programs with the same iteration complexity. Since then ($\star$) has been one of the most important principles not only in the theoretical development of the PDIPM but also in the development of many software packages incorporating the PDIPM for LPs.

In 1994, Nesterov and Nemirovskii [25] proposed a polynomial-time IPM for a fairly general class of convex programs as an extension of the IPM for LPs. Their class includes the semidefinite programs (SDP) and second-order cone programs (SOCP). Until then, they were not so popular subjects in the field of optimization, although SDPs were important and basic tools in control theory. But, as soon as their IPM algorithm showed potential efficiency in solving SDPs, a great deal of studies on interior-point algorithms for SDPs and their application to SDP relaxations of combinatorial optimization problems were initiated.

PDIPMs for SDPs were proposed independently by four groups, Alizadeh-Haeberly-Overton [1], Helmberg-Rendl-Vanderbei-Wolkowicz [8], Kojima-Shindoh-Hara [19], and Nesterov-Todd [26].

There exists a central trajectory converging to a primal-dual pair of optimal solutions of an SDP in the interior of the Cartesian product of the primal and dual feasible region which is assumed to be nonempty. However, the extension of the LP case to the SDP case is not straightforward, as the rigorous Newton direction toward the central trajectory cannot be defined. Among the Newton-like directions proposed for tracing the central trajectory, the

NT direction by Nesterov and Todd [26] and the HKM direction are well-known. The latter HKM direction was proposed independently by Helmberg-Rendl-Vanderbei-Wolkowicz [8] and Kojima-Shindoh-Hara [19], and later rediscovered by Monteiro [23] from a different formulation. This direction was implemented in popular software packages including SDPA and SDPT3 [28]. A preliminary version of SDPA was written by myself as I was learning C++ (Borland C++) more than 25 years ago, and then I gave it to Katsuki Fujisawa, who joined my group as a Ph.D. student, for its first version [4]. Since then, many people participated in improvements and extensions of SDPA, and it has now become a large SDPA family [2], including SDPARA (a parallel version), SDPA-M (with MATLAB interface), SDPA-P (with Python interface) and SDPA-GMP (with arbitrary precision arithmetic).

4 Exploiting sparsity in SDPs and polynomial optimization problems

While working on various subjects in optimization, I have always felt that practicality is essential in research and the goal of optimization is software. I am well aware that developing good software requires a great deal of patience and time. Based on my belief, I started to work on the sparsity exploitation for SDPs after developing SDPA.

Let $\mathbb{S}^n$ denote the space of $n \times n$ symmetric matrices with the inner product $\boldsymbol{A} \bullet \boldsymbol{B} = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$, and $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$) the set of positive semidefinite (positive definite) matrices in $\mathbb{S}^n$. $\mathbb{S}^n_+$ forms a closed convex cone with its interior $\mathbb{S}^n_{++}$. Then a standard equality form SDP and its dual are described as

$$\begin{aligned} &\text{(P):} \qquad \min\left\{\boldsymbol{A}_0 \bullet \boldsymbol{X} : \boldsymbol{A}_k \bullet \boldsymbol{X} = b_k \ (k = 1, \ldots, m),\ \boldsymbol{X} \in \mathbb{S}^n_+\right\}, \\ &\text{(D):} \qquad \max\left\{\sum_{k=1}^m b_k y_k : \boldsymbol{Z} = \boldsymbol{A}_0 - \sum_{p=1}^m \boldsymbol{A}_k y_p \in \mathbb{S}^n_+\right\}. \end{aligned}$$

Here $\boldsymbol{A}_0,\ \boldsymbol{A}_k \in \mathbb{S}^n$ and $b_k \in \mathbb{R}$ $(k = 1, \ldots, m)$. In applying a standard PDIPM to (P) and (D) with m and/or n large, there are three main difficulties described below.

The first one is from the computation of a search direction $(d\boldsymbol{X}, d\boldsymbol{y}, d\boldsymbol{Z})$ at each iterate $(\boldsymbol{X}^p, \boldsymbol{y}^p, \boldsymbol{Z}^p) \in \mathbb{S}^n_{++} \times \mathbb{R}^m \times \mathbb{S}^n_{++}$, where the so-called Schur complement equations $\boldsymbol{B} d\boldsymbol{y} = \boldsymbol{r}$ with some $\boldsymbol{B} \in \mathbb{S}^m$ is solved for $d\boldsymbol{y}$ first, and then the rest of components $d\boldsymbol{X}$ and $d\boldsymbol{Z}$ of the search direction can be computed easily from $d\boldsymbol{y}$ in $O(mn^2)$ arithmetic operations. In the case of the HKM direction, each element of $\boldsymbol{B}$ is given by $B_{k\ell} = \boldsymbol{X}\boldsymbol{A}_k\boldsymbol{Z}^{-1} \bullet \boldsymbol{A}_\ell$ $(1 \le k, \ell \le m)$. Hence the computation of $\boldsymbol{B}$ requires $O(mn^3 + m^2n^2)$ arithmetic operations in a straightforward dense computation. As m and n increase, it would become the most expensive part in one iteration of the PDIPM. In practice, $\boldsymbol{A}_k$ $(k = 1, \ldots, m)$ are often sparse, although $\boldsymbol{X}^p$, $\boldsymbol{Z}^p$ and $\boldsymbol{B}$ become dense in general. For example, each binary constraint $x_k(1 - x_k) = 0$ induces an $\boldsymbol{A}_k$ with only three nonzero elements for the SDP relaxations of combinatorial quadratic optimization problems. In my joint paper [5] with Fujisawa and Nakata, we focused on this difficulty and proposed a technique for computing $B_{k\ell} = \boldsymbol{X}\boldsymbol{A}_k\boldsymbol{Z}^{-1} \bullet \boldsymbol{A}_\ell$ by taking account of the sparsity in $\boldsymbol{A}_k$ and $\boldsymbol{A}_\ell$ $(1 \le k, \ell \le m)$. This technique was incorporated in SDPA and also many other software packages including SDPT3 [28] to improve the performance of the PDIPM.

The second difficulty is that the primal variable matrix $\boldsymbol{X}$ in (P) is fully dense in general even when all data matrices $\boldsymbol{A}_k$ $(k = 0, \ldots, m)$ are sparse. (On the other hand, the dual variable matrix $\boldsymbol{Z}$ naturally inherits their sparsity). In each iteration p, we compute the Cholesky factorization $\boldsymbol{L}$ of $\boldsymbol{X}^p$ and the minimum eigenvalue of $\boldsymbol{L}^{-1}d\boldsymbol{X}\boldsymbol{L}^{-T}$, which require $O(n^3)$ arithmetic operations in the dense computation, to determine a step length. It becomes more expensive as the size n of $\boldsymbol{X}$ becomes larger even when the data matrices $\boldsymbol{A}_p$ $(p = 1, \ldots, m)$ are sparse. To overcome this difficulty, Fukuda, Murota, Nakata and myself [6] introduced the positive definite matrix completion technique [7] for the first time in SDPs. An undirected graph $G(N, E)$ with a node set $N = \{1, \ldots, n\}$ and an edge set $E \subset \{(i, j) : i, j \in N,\ i \neq j\}$ is said to be chordal if every cycle with at least 4 edges has a chord. We assumed that the aggregated sparsity of the data matrices $\boldsymbol{A}_k$ $(k = 0, \ldots, m)$ is characterized by a chordal graph $G(N, E)$ with maximal cliques $C_1, \ldots, C_\ell$ satisfying $\left\{(i, j) \in N \times N : [\boldsymbol{A}_k]_{ij} \neq 0 \text{ for some } k\right\} \subset \cup_{k=1}^{\ell} (C_k \times C_k)$. Then the positive semidefinite condition on $\boldsymbol{X}$ can be replaced by the family of positive semidefinite conditions on matrices $\boldsymbol{X}_{C_1}, \ldots, \boldsymbol{X}_{C_\ell}$, where $\boldsymbol{X}_{C_k}$ denotes the principal submatrix of $\boldsymbol{X}$ consisting of elements X_{ij} $(i, j) \in C_k \times C_k$. Based on this, we proposed two types of methods. The first method is a conversion method that converts (P) to an equivalent SDP with multiple smaller variable matrices $\boldsymbol{X}_{C_k}$ $(k = 1, \ldots, \ell)$. The other method is a completion method that incorporates the above fact and the positive semidefinite matrix completion technique in the PDIPM for solving (P) and (D) directly. The latter method was implemented in SDPAC and SDPARAC [2]. Both methods work very efficiently for large scale sparse SDPs.

The third difficulty is that the Schur complement matrix $\boldsymbol{B}$ is fully dense in general even if the aforementioned sparsity exploiting techniques are applied. Kobayashi, Kim and myself [14] showed that if the data matrices $\boldsymbol{A}_k$ $(k = 1, \ldots, m)$ satisfy *correlative sparsity*, a structured sparsity also characterized by a chordal graph, then the conversion method reduces (P) to an SDP in multiple smaller variable matrices whose Schur complement matrix has a sparse Cholesky factorization. The conversion method was extended to the domain and range space conversion methods by Kim, Mevissen, Yamashita and myself [10], and implemented in SparseCoLO [3].

Lasserre [20] proposed a hierarchy of SDP relaxations of polynomial optimization problems (abbreviated by POPs) in 2001. In theory, his method was shown to be very powerful to attain accurate lower bounds for POPs. In practice, however, the SDP relaxation problem often becomes too large to solve as the number of the variables and/or the degree of polynomials involved in a POP to be solved increase. To overcome this difficulty, Waki, Kim and myself [29] incorporated the basic idea in the conversion method described above into his method to develop SparsePOP [30], which can solve larger-scale sparse POPs.

5 Moving forward in the field of optimization

In my life of almost 50 years as a researcher, I have worked on various subjects. I chose the subjects that I liked and I wanted to make important contributions to the optimization field by writing papers on the subjects. I know, however, that many of those papers might not have made much contribution. Nevertheless, they provided me with learning experiences for the next step.

Since I retired from Tokyo Institute of Technology several years ago, I have been working mainly on two subjects. The first one is the completely positive (abbreviated as CPP) relaxation of QOPs and its extension to POPs [11, 12]. This subject is mainly from my theoretical interest. I would like to know more about the essentials of CPP relaxations. The other subject is efficient computation of the doubly nonnegative relaxation of large-scale QOPs and POPs [13]. It is my hope to make important contributions in these areas, as I walk forward in the field of optimization.

Acknowledgement

I would like to express my deep thanks to Sunyoung Kim for many helpful comments and suggestions on this article.

References

[1] F. Alizadeh, J. P. A. Haeberly, and M. L.Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM Journal of Optimization*, 8:746–768, 1998.

[2] K. Fujisawa, M. Fukuda, Y. Futakata, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata, and M. Yamashita. SDPA official page http://sdpa.sourceforge.net, March 2011.

[3] K. Fujisawa, S. Kim, M. Kojima, Y. Okamoto, and M. Yamashita. SparseCoLo: Conversion methods for Sparse COnic form Linear optimization, 2008.

[4] K. Fujisawa, M. Kojima, and K. Nakata. SDPA (semidefinite programming algorithm) — user's manual —. Technical Report B-308, Dept. of Math. and Comp. Sci., Tokyo Institute of Technology, Tokyo 152, December 1995.

[5] K. Fujisawa, M. Kojima, and K. Nakata. Exploiting sparsity in primal-dual interior-point methods for semidefnite programming. *Mathematical Programming*, 79:235–253, 1997.

[6] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting sparsity in semidefinite programming via matrix completion. I: General framework. *SIAM J. Optim.*, 11:647–674, 2000.

[7] R. Gron, C. R. Johnson, E. M. Sá, and H. Wolkowicz. Positive definite completions of partial hermitian matrices. *Linear Algebra Appl.*, 58:109–124, 1984.

[8] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM J. Optimization*, 6:342–361, 1996.

[9] N. Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on theory of computing*, pages 302–311. ACM, 1984.

[10] S. Kim, M. Kojima, M. Mevissen, and Yamashita M. Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. *Math. Program.*, 129:33–68, 2011.

[11] S Kim, M. Kojima, and K. C. Toh. A geometrical analysis of a class of nonconvex conic programs for convex conic reformulations of quadratic and polynomial optimization problems. *SIAM J. on Optimization*, 30:1251–1273, 2020.

[12] S. Kim, M. Kojima, and K.C. Toh. Doubly nonnegative relaxations are equivalent to completely positive reformulations of quadratic optimization problems with block-clique graph structures. *Journal of Global Optimization*, 77:513–541, 2020.

[13] S. Kim, M. Kojima, and K.C. Toh. A Newton-bracketing method for a simple conic optimization problem. *Optimization Methods and Software*, To appear.

[14] M. Kobayashi, S. Kim, and M. Kojima. Correlative sparsity in primal-dual interior-point methods for LP, SDP, and SOCP. *Appl. Math. Optim.*, 58:69–88, 2008.

[15] M. Kojima. Studies on piecewise-linear aproximations of piecewise-c1 mappings in fixed pointes and complementarity theory. *Mathematics of Operations Research*, 3:17–36, 1978.

[16] M. Kojima. *Strongly Stable Stationary Solutions in Nonlinear Programs*, pages 93–138. Academic Press, New York, 1980.

[17] M. Kojima, S. Mizuno, and A. Yoshise. A polynomial-time algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44(1-26), 1989.

[18] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior poiint algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*, chapter 2, pages 29–47. Springer-Verlag, 1989.

[19] M. Kojima, S. Shindoh, and S. Hara. Interior-point methods for the monotone linear complementarity problem in symmetric matrices. *SIAM J. Optim.*, 7:86–125, 1997.

[20] J. B. Lasserre. Global optimization with polynomials and the problems of moments. *SIAM J. Optim.*, 11:796–817, 2001.

[21] C. E. Lemke and J. T. Jr. Howson. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12:413–423, 1964.

[22] N. Megiddo. Path ways to the optimal set in linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*, chapter 8, pages 131–158. Springer-Verlag, 1989.

[23] R. C. Monteiro. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. Optim.*, 7:663–678, 1997.

[24] R. C. Monteiro and I. Adlar. Interior path following primal-dual algorithms. part II: Convex quadratic programming. *Mathematical Programming*, 44:43–66, 1989.

[25] Y. E. Nesterov and A. Nemiroskii. *Interior-Point Polynomial Algorityms in Convex Programming*. SIAM, Philadelphia, USA, 1994.

[26] Y. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8:324–363, 1998.

[27] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.

[28] K.C. Toh, M.J. Todd, and R.H. Tutuncu. SDPT3 — a Matlab software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.

[29] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. *SIAM J. Optim.*, 17:218–242, 2006.

[30] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto. Algorithm 883: Sparsepop: A sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Trans. Math. Softw.*, 35(15), 2008.