ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta Elektrotechnická

CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Electrical Engineering

Ing. Didier Henrion, PhD. CSc.

Convex optimization for solving non-convex polynomial optimization problems

Czech(*Convex optimization for solving non-convex polynomial optimization problems*)

January 2008

Summary

Polynomial optimization problems, where a multivariate polynomial function is to be minimized subject to polynomial inequality constraints, are ubiquitous in engineering, and in particular in systems control. In general, these optimization are non-convex, and hence typically difficult to solve numerically.

In this lecture we explain how convex optimization, and in particular semidefinite programming, can be used to solve globally these non-convex optimization problems, with a numerical certificate of global optimality.

Souhrn

Czech(*Polynomial optimization problems, where a multivariate polynomial function is to be minimized subject to polynomial inequality constraints, are ubiquitous in engineering, and in particular in systems control. In general, these optimization are non-convex, and hence typically difficult to solve numerically.*

In this lecture we explain how convex optimization, and in particular semidefinite programming, can be used to solve globally these non-convex optimization problems, with a numerical certificate of global optimality.)

Keywords: optimization, polynomials, convexity

Klíčová slova: Czech(*optimization, polynomials, convexity*)

Contents

1 Introduction

Most analysis and design problems in robust and nonlinear control can be formulated as global optimization problems with polynomial objective functions and constraints. Typical examples include robust stability analysis for characteristic polynomials with parametric uncertainty, simultaneous stabilization of linear systems, pole assignment by static output feedback, and stability analysis for polynomial systems by Lyapunov's second approach. In some specific cases, there exist computationally efficient techniques for solving these problems. For example, vertex or extremal results such as Kharitonov's Theorem or the Edge Theorem can be used to perform robust stability analysis without optimization. Similarly, static state-feedback design, or design of a controller of the same order as the plant, can be formulated as a convex linear matrix inequality (LMI) optimization problem [1], for which efficient interior-point methods are available [2].

Polynomial optimization problems arising from control problems are often highly non-convex, with several local optima, and are difficult to solve. Although general purpose global optimization algorithms can be applied, the computational cost is often an exponential function of the number of decision variables. To overcome the use of computationally intensive algorithms, researchers have focused on the development of relaxation or simplification techniques relying on convex optimization. A convex relaxation of a non-convex problem is obtained by removing non-convex constraints or replacing them with necessary (but generally not equivalent) convex constraints, hence simplifying and enlarging the set over which the optimization is carried out. In the last decade, semidefinite programming (SDP), or LMI optimization, has established itself as a popular convex relaxation technique in the systems and control community.

Conservatism is the price one has to pay when simplifying a non-convex problem. For example, convex sufficient stability conditions are frequently used instead of non-convex necessary and sufficient stability conditions when performing robust design. Generally speaking, there is a trade-off between the amount of conservatism and the computational cost when solving a non-convex problem. Due to the amount of conservatism inherent in LMI techniques, which is difficult to measure accurately for practical control problems [2], there has recently been a surge of interest in approaches that gradually increase computational complexity. Most of these approaches are based on sufficient conditions for the positivity of multivariable polynomials. For example, positivity of polynomials is replaced with the stronger sum of squares (SOS) constraint, which has an LMI formulation [5]. An alternative approach based on the theory of moments has been developed independently in [6].

For non-convex problems, the relaxation technique described in [6] enables the user to systematically construct an increasing sequence of convex LMI relaxations, whose optima are guaranteed to converge monotonically to the global optimum of the original non-convex global optimization problem. A Matlab implementation of the relaxation technique, called GloptiPoly [3], has been developed as an open-source freeware based on the LMI solver SeDuMi. Numerical experiments suggest that for most small- and medium-size problems in the technical literature on global optimization, the global optimum is reached with LMI relaxations of medium size, at a relatively low computational cost. Moreover, global optimality can sometimes be proved by using sufficient rank conditions and numerical linear algebra techniques.

The objective of this lecture is to describe the hierarchy of convex LMI relaxations for non-convex polynomial optimization. We keep the technical level elementary, focusing more on the main ideas than on the mathematical details. Two numerical examples are presented to illustrate the LMI relaxations.

2 A hierarchy of convex relaxations

Consider the multivariate polynomial optimization problem

$$\mathbb{P} : p^* = \begin{array}{ll} \min_x & g_0(x), \\ \text{s.t.} & g_k(x) \geq 0, \quad k = 1, \ldots, m, \end{array} \tag{1}$$

where $g_k \in \mathbb{R}[x_1, \ldots, x_n]$ are real-valued polynomials. Formulation (1) encompasses non-convex quadratic problems as well as discrete optimization problems, such as 0-1 nonlinear programming problems. Denote by $\mathbb{K}$ the feasible set of $\mathbb{P}$, that is,

$$\mathbb{K} = \{x \in \mathbb{R}^n : g_k(x) \geq 0, \quad k = 1, \ldots, m\}. \tag{2}$$

The idea behind the methodology in GloptiPoly is to construct a sequence of convex LMI relaxations of $\mathbb{P}$ of increasing size and whose sequence of optimal values converges to the global optimal value $p^* = \inf \mathbb{P}$. The proof of convergence of the LMI relaxations is based on recent results in real algebraic geometry concerning the representation of polynomials that are strictly positive on a semi-algebraic set. It turns out that the primal and dual LMI relaxations of GloptiPoly correspond to the dual theories of *moments* and *positive polynomials*.

Indeed, while the primal relaxations aim at finding the moments of a probability measure with mass concentrated on some global minimizers of $\mathbb{P}$, the dual relaxations aim at representing the polynomial $g_0(x) - p^*$, which is positive on the semi-algebraic feasible set $\mathbb{K}$ of $\mathbb{P}$, as a linear combination of the $g_k(x)$ with SOS polynomial weights.

2.1 Primal relaxations

In brief, the primal LMI relaxations $\{\mathbb{Q}_i\}$ of $\mathbb{P}$ are relaxations of the moment problem

$$p^* = \min_{\mu} \int_{\mathbb{K}} g_0 \, d\mu, \tag{3}$$

which is equivalent to $\mathbb{P}$, and where the unknown μ belongs to the Borel set of probability measures supported on semialgebraic set $\mathbb{K}$. For a multi-index $\alpha \in \mathbb{N}^n$, $g_0(x) = \sum_\alpha (g_0)_\alpha x^\alpha$ is a polynomial of the monomials $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and thus the objective function is a finite linear combination $\sum_\alpha (g_0)_\alpha y_\alpha$ of *moments*

$$y_\alpha = \int x^\alpha d\mu$$

of the probability measure μ. The relaxations of (3) are obtained by replacing the constraint that μ has its support in $\mathbb{K}$ with progressively stronger semidefinite programming conditions on its moments.

For instance, let $2v_k - 1$ or $2v_k$ be the degree of the polynomial g_k in the definition (2) of the set $\mathbb{K}$, and let $v = \max_k v_k$. Then, the relaxation of order i includes the constraints

$$\int f^2 g_k \, d\mu \geq 0, \quad k = 1, \ldots, m, \tag{4}$$

for all polynomials $f \in \mathbb{R}[x_1, \ldots, x_n]$ of degree at most $i - v$. Inequalities (4) translate into equivalent LMI constraints on the moments $\{y_\alpha\}$ of μ of order $|\alpha| \leq 2i$. Of course, the larger the order i, the larger the size of the associated LMI constraints.

2.2 Dual relaxations

On the other hand, the LMI relaxations $\{\mathbb{Q}_i^*\}$ that are dual to $\{\mathbb{Q}_i\}$ solve the optimization problems

$$\begin{array}{ll} \max_{p_i,\{q_k\}} & p_i \\ \text{s.t.} & g_0(x) - p_i = q_0 + \sum_{k=1}^m g_k(x) q_k(x) \end{array} \tag{5}$$

where the unknowns $\{q_k\}$ are polynomials in x, all sums of squares. Both the number of variables and the number of constraints in the relaxation $\mathbb{Q}_i^*$ depend on the maximum degree $2i$ allowed in the right-hand-side of (5). The increasing numbers of variables and constraints in the relaxations reflect that the degree $2i$ must be large enough in (5) for p_i to be as close as desired to p^* (and often to be exactly equal to p^*).

We consider mild technical assumptions on the feasible set $\mathbb{K}$, which are satisfied, for example, when $\mathbb{K}$ is a polytope, or when the level set $g_k(x) \geq 0$ is compact for some index k. Such assumptions can always be satisfied by enforcing a sufficiently large feasibility radius on the decision variables, that is, by introducing the additional Euclidean norm constraint $\|x\|^2 \leq R^2$ for sufficiently large R. Then it was proved in [6] that $\inf \mathbb{Q}_i$ converges to $\inf \mathbb{P}$ as i tends to infinity. In other words, letting p_i^* denote the optimum obtained by solving the LMI relaxation $\mathbb{Q}_i$ of order i, we obtain a monotone sequence of lower bounds p_i^* converging asymptotically to the globally optimal value p^* of the original optimization problem in (1). This monotonicity means that the sequence is designed to do better, or at least not worse, at each step. Moreover, our computational experiments on global optimization benchmark examples reveal that in practice p_i^* is very close to p^* for relatively small values of i. In addition, in many cases the exact optimal value p^* is obtained at some particular relaxation $\mathbb{Q}_i$, that is, $p^* = p_i^*$ for some relatively small i.

In our software GloptiPoly [3] we have implemented a numerical linear algebra algorithm that detects global optimality, for example, to determine whether the LMI relaxation $\mathbb{Q}_i$ provides the optimal value $p_i^* = p^*$, and another algorithm to extract global minimizers. Roughly speaking, detecting global optimality amounts to checking successive ranks of moment matrices, whereas global minimizer extraction amounts to computing a Cholesky factor of the moment matrix and solving an eigenvalue problem. All of these tasks can be carried out efficiently with standard algorithms of numerical linear algebra.

3 Examples

We now consider two examples of constructions of successive LMI relaxations. By emulating these examples, the reader should be able to build up LMI relaxations for more general polynomial optimization problems.

3.1 First example

Consider the non-convex optimization problem

$$\begin{array}{ll} \max & x_2 \\ \text{s.t.} & 3 + 2x_2 - x_1^2 - x_2^2 \geq 0 \\ & -x_1 - x_2 - x_1 x_2 \geq 0 \\ & 1 + x_1 x_2 \geq 0 \end{array}$$

where the linear objective function x_2 is maximized over a non-convex feasible set delimited by circular and hyperbolic arcs. The feasible region is shown in Figure 1.

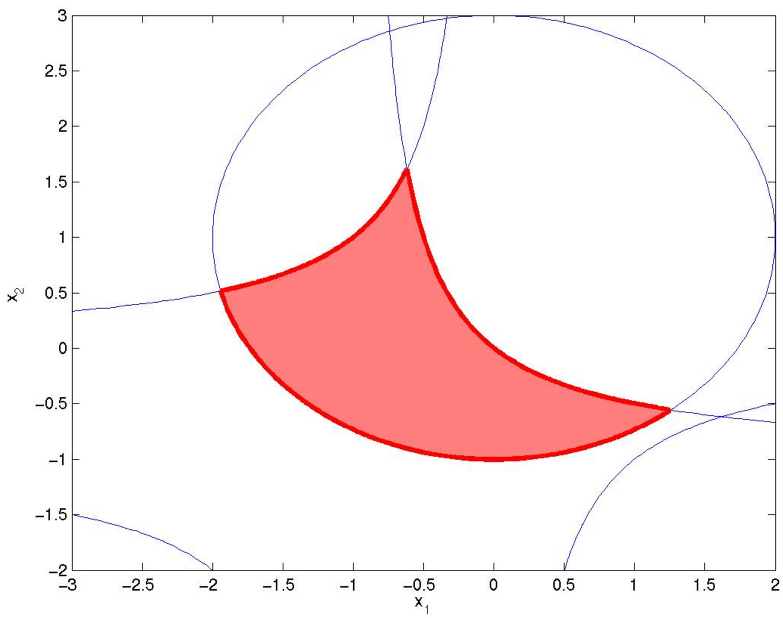

Figure 1: Feasible set for Example 1. The feasible set (shaded region) is non-convex and delimited by circular and hyperbolic arcs.

The first LMI relaxation $\mathbb{Q}_1$ is

$$\begin{array}{ll} \max & y_{01} \\ \text{s.t.} & \left[\begin{array}{ccc} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{array}\right] \succeq 0 \\ & 3 + 2y_{01} - y_{20} - y_{02} \geq 0 \\ & -y_{10} - y_{01} - y_{11} \geq 0 \\ & 1 + y_{11} \geq 0 \end{array}$$

with optimal value $p_1 = 2$. The notation $\succeq 0$ stands for positive semidefinite. In this relaxation, the 3×3 matrix is a moment matrix of order up to 2. Problem constraints are linearized with the help of these moment variables.

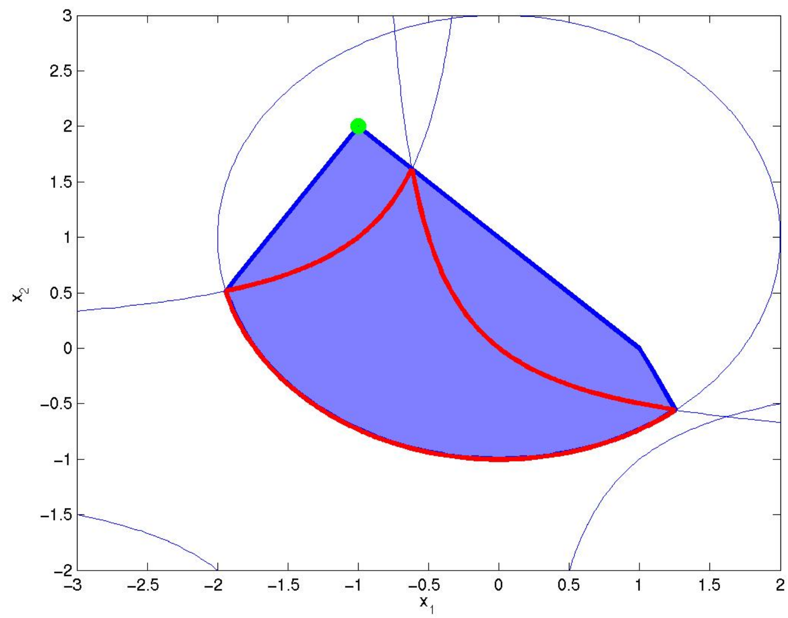

Figure 2: Feasible set of the first convex LMI relaxation for Example 1. The feasible set of the first LMI relaxation (shaded region) is obtained by projecting the first-order moments onto the plane. The optimum of the first LMI relaxation is attained at the upper vertex (dot) of the feasible set. The optimum is an upper bound on the global optimum of the original non-convex polynomial optimization problem.

In Figure 2 we show the projection of the feasibility set of LMI relaxation $\mathbb{Q}_1$ onto the plane y_{10}, y_{01} of first-order moments. This convex feasibility set inscribes the original non-convex feasible set. We can see that the optimum of the LMI relaxation is achieved at a point that is infeasible for the non-convex problem.

The second LMI relaxation $\mathbb{Q}_2$ is

$$\begin{array}{ll}
\max & y_{01} \\
\text{s.t.} & \left[\begin{array}{cccccc}
1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\
y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\
y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\
y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\
y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\
y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04}
\end{array}\right] \succeq 0, \\
& \left[\begin{array}{ccc}
3+2y_{01}-y_{20}-y_{02} & 3y_{10}+2y_{11}-y_{30}-y_{12} & 3y_{01}+2y_{02}-y_{21}-y_{03} \\
3y_{10}+2y_{11}-y_{30}-y_{12} & 3y_{20}+2y_{21}-y_{40}-y_{22} & 3y_{11}+2y_{12}-y_{31}-y_{13} \\
3y_{01}+2y_{02}-y_{21}-y_{03} & 3y_{11}+2y_{12}-y_{31}-y_{13} & 3y_{02}+2y_{03}-y_{22}-y_{04}
\end{array}\right] \succeq 0, \\
& \left[\begin{array}{ccc}
-y_{10}-y_{01}-y_{11} & -y_{20}-y_{11}-y_{21} & -y_{11}-y_{02}-y_{12} \\
-y_{20}-y_{11}-y_{21} & -y_{30}-y_{21}-y_{31} & -y_{21}-y_{12}-y_{22} \\
-y_{11}-y_{02}-y_{12} & -y_{21}-y_{12}-y_{22} & -y_{12}-y_{03}-y_{13}
\end{array}\right] \succeq 0, \\
& \left[\begin{array}{ccc}
1+y_{11} & y_{10}+y_{21} & y_{01}+y_{12} \\
y_{10}+y_{21} & y_{20}+y_{31} & y_{11}+y_{22} \\
y_{01}+y_{12} & y_{11}+y_{22} & y_{02}+y_{13}
\end{array}\right] \succeq 0
\end{array}$$

with optimal value $p_2 = 1.6180$, which is the global optimum p^* within numerical accuracy. In addition, first order moments $(y^*_{10}, y^*_{01}) = (-0.6180, 1.6180)$ provide an optimal solution of the original problem. This problem features a 6×6 moment matrix corresponding to moments of order up to 4. The three 3×3 LMI constraints are the LMI formulation of (4).

In Figure 3 we show the projection of the feasibility set of the LMI relaxation $\mathbb{Q}_2$ onto the plane y_{10}, y_{01} of first-order moments. By construction, the feasibility set of the LMI relaxation $\mathbb{Q}_2$ is included in the feasibility set of the LMI relaxation $\mathbb{Q}_1$. Compared to Figure 3, we can see that the feasibility set of the LMI relaxation $\mathbb{Q}_2$ is exactly the convex hull of the original non-convex feasible set, and the global optimum is now attained because the objective function x_2 is linear in the first-order moments.

3.2 Second Example

Consider the optimization problem

$$\begin{array}{ll}
\max & x_1^2 + x_2^2 \\
\text{s.t.} & 3 + 2x_2 - x_1^2 - x_2^2 \geq 0 \\
& -x_1 - x_2 - x_1 x_2 \geq 0 \\
& -1 - 4x_2 - 4x_1 x_2 \geq 0
\end{array}$$

where the objective function $\|x\|^2$, the squared Euclidean norm of x, is maximized over the non-connected set shown in Figure 4. This problem admits various local optima.

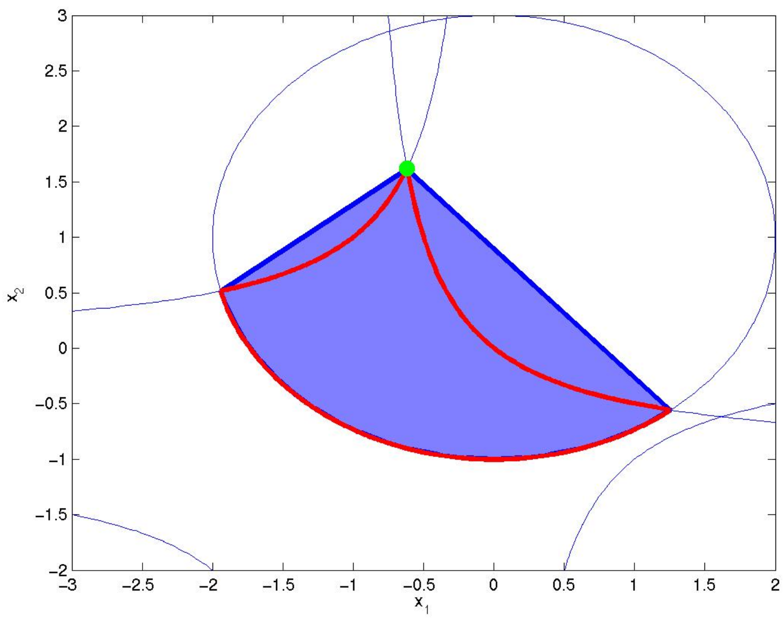

Figure 3: Feasible set of the second convex LMI relaxation for Example 1. The feasible set of the second LMI relaxation (shaded region) is obtained by projecting the first-order moments onto the plane. The optimum of the second LMI relaxation is equal to the global optimum (dot).

The first LMI relaxation $\mathbb{Q}_1$ given by

$$\begin{array}{ll}\max & y_{20}+y_{02}\\ \text{s.t.} & \begin{bmatrix} 1 & y_{10} & y_{01}\\ y_{10} & y_{20} & y_{11}\\ y_{01} & y_{11} & y_{02}\end{bmatrix} \succeq 0\\ & 3+2y_{01}-y_{20}-y_{02}\geq 0\\ & -y_{10}-y_{01}-y_{11}\geq 0\\ & -1-4y_{02}-4y_{11}\geq 0\end{array}$$

yields the global optimum $p^\star = p_1 = 8.3492$ attained at $(y^*_{10}, y^*_{01}) = (-1.0935, 2.6746)$. The optimum is achieved on the boundary of the convex hull of the non-convex non-connected feasible set.

If we now wish to minimize (instead of maximize) $\|x\|^2$, the first LMI relaxation yields the global optimum $p^\star = p_1 = 0.059176$ attained at $(y^*_{10}, y^*_{01}) = (0.0535, -0.2372)$. From Figure 4 we can see that the global optimum is not achieved on the boundary of the convex hull of the feasible set.

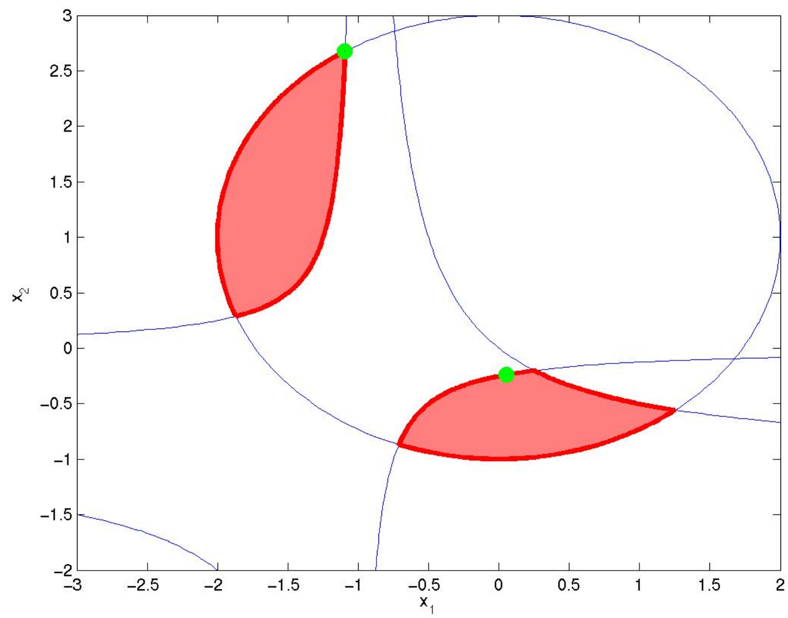

Figure 4: Feasible set for Example 2. The feasible set (shaded region) is non-convex and non-connected, delimited by circular and hyperbolic arcs. Also represented are global optima with minimum Euclidean norm (dot near the origin) and maximum Euclidean norm (dot at the top).

4 Conclusion

With the help of two simple numerical examples, we have briefly surveyed the general LMI relaxation methodology for polynomial optimization developed in [6] and implemented in the Matlab freeware GloptiPoly [3]. In the context of global optimization, the technique is original in the sense that it does not perform any problem splitting, and thus avoids the combinatorial explosion typical of branch and bound schemes.

In particular, this software can help solve various non-convex robust control problems as soon as they can be formulated as polynomial optimization problems. Conservatism can be reduced at the cost of a limited amount of additional computation.

Acknowledgment

Most of the material exposed here can be found in extended form in [4].

References

[1] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan. Linear Matrix Inequalities in System and Control Theory. SIAM, Philadelphia, PA, 1994.

[2] A. Ben-Tal, A. Nemirovski. Lectures on Modern Convex Optimization. SIAM, Philadelphia, PA, 2001.

[3] D. Henrion, J. B. Lasserre. GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi. ACM Transactions on Mathematical Software, Vol. 29, No. 2, pp. 165-194, 2003.

[4] D. Henrion, J. B. Lasserre. Solving nonconvex optimization problems - How GloptiPoly is applied to problems in robust and nonlinear control. IEEE Control Systems Magazine, Vol. 24, No. 3, pp. 72-83, 2004.

[5] P. A. Parrilo. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. *PhD Thesis*, California Institute of Technology, Pasadena, CA, 2000.

[6] J. B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, Vol. 11, No. 3, pp. 796–817, 2001.

Curriculum Vitae of Ing. Didier Henrion, PhD. CSc.

Didier Henrion was born in Creutzwald, Northeastern France, in 1971.

He received the Diplôme d'Ingénieur (Engineer's Degree) and the Diplôme d'Etudes Approfondies (Masters' Degree) with specialization in control from the Institut National des Sciences Appliquées (INSA, National Institute for Applied Sciences), Toulouse, Southwestern France, in September 1994. From October 1994 to December 1995 he was a research assistant at the Centro de Automatización Industrial of the Universidad Simón Bolívar, Caracas, Venezuela. From February 1996 to December 1998 he was a Ph.D. student at the Ústav Teorie Informace a Automatizace (ÚTIA, Institute of Information Theory and Automation), Prague, Czech Republic. He received the C.Sc. (Ph.D.) degree from the Akademie Věd České Republiky (Academy of Sciences of the Czech Republic) in December 1998. From October 1996 to August 1999 he was a Ph.D. student at the Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS, Laboratory of Analysis and Architecture of Systems) of the Centre National de la Recherche Scientifique (CNRS, National Center for Scientific Research) in Toulouse, France. He received the Diplôme de Doctorat (Ph.D. degree) from INSA in October 1999, and the Diplôme d'Habilitation à Diriger des Recherches (habilitation degree) from the University of Toulouse in December 2007.

Since October 2000, D. Henrion has been a Chargé de Recherche au CNRS (CNRS Researcher) at LAAS and a researcher at the Czech Technical University in Prague. In 2004 he was awarded the Bronze Medal, the highest distinction from CNRS for young reseachers.

From 1998 to 2001, D. Henrion was the French project leader of the European Network of Excellence EUROPOLY funded by the European Commission. From 1997 to 2002, he was the French project leader of the French-Czech joint project Barrande. From 2001 to 2003, he coordinated a NATO Collaborative linkage grant with the Czech Technical University in Prague. From 2002 to 2004, he was the French project leader of a bilateral agreement between the French CNRS and the Czech Academy of Sciences. In 2001 he was awarded a two-year Young Researcher Grant by CNRS. In 2006 he was awarded a three-year research grant by the Grant Agency of the Czech Republic. From 2006 to 2007, he was coordinating an ECONET project funded by the French Ministry of Foreign Affairs, involving research groups in France, Czech Republic, Slovakia and Russia.

D. Henrion is an Associate Editor of the IFAC Automatica journal (since 2002), the European Journal of Control (since 2005) and the IEEE Transactions on Automatic Control (since 2005). From 1999 to 2004 he was an Associate Editor at the Conference Editorial Board of the IEEE Control Systems Society. He is chairing the IEEE Technical Committee on Computer Aided Control System Design, vice-chairing the IFAC Technical Committee on Control Design, and member of the IFAC Technical Committee on Robust Control. He is also a regular member of the panel of reviewers of the AMS Mathematical Reviews and Zentralblatt MATH.

D. Henrion's main research topics are numerical algorithms for polynomial matrices and convex optimization over linear matrix inequalities with applications to linear systems and robust control. He is also involved in the application of multivariable robust design techniques to the control of turbofan engines in collaboration with the French aerospace propulsion and equipment group SAFRAN (formerly SNECMA Moteurs).

D. Henrion has been involved in the development of various software tools, including the Polynomial Toolbox for Matlab (since 1996, jointly with M. Šebek and H. Kwakernaak), GloptiPoly (since 2002, jointly with J. B. Lasserre) and HIFOO (since 2005, jointly with J. V. Burke, A. S. Lewis and M. L. Overton).

D. Henrion has given courses on LMI optimization and polynomial methods in system control at Mérida, Venezuela (October 2001), Prague, Czech Republic (November 2003, February 2005, February 2006), Toulouse, France (May 2004, April 2005), Leuven, Belgium (June 2006), Dortmund, Germany (October 2006), Rende, Italy (March 2007) and Barcelona, Spain (June 2007).