
Detecting Global Optimality and Extracting Solutions in GloptiPoly

Didier Henrion^{1,2} and Jean-Bernard Lasserre¹

¹ LAAS-CNRS, 7 Avenue du Colonel Roche, 31 077 Toulouse, France.
`henrion@laas.fr, lasserre@laas.fr`

² Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic.

GloptiPoly is a Matlab/SeDuMi add-on to build and solve convex linear matrix inequality (LMI) relaxations of non-convex optimization problems with multivariate polynomial objective function and constraints, based on the theory of moments. In contrast with the dual sum-of-squares decompositions of positive polynomials, the theory of moments allows to detect global optimality of an LMI relaxation and extract globally optimal solutions. In this report, we describe and illustrate the numerical linear algebra algorithm implemented in GloptiPoly for detecting global optimality and extracting solutions. We also mention some related heuristics that could be useful to reduce the number of variables in the LMI relaxations.

1 Introduction

Consider the global optimization problem

$$\begin{aligned} p^* &= \min_x g_0(x) \\ \text{s.t. } g_i(x) &\geq 0, \quad i = 1, 2, \dots, m \end{aligned} \tag{1}$$

where the mappings $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 0, 1, \dots, m$ are real-valued polynomials, that is, $g_i \in \mathbb{R}[x_1, \dots, x_n]$ for all $i = 1, \dots, m$. Depending on its parity, let $\deg g_i = 2d_i - 1$ or $2d_i$, and denote $d = \max_i d_i$. Define

$$v_k(x) = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_n & x_1^2 & x_1x_2 & \dots & x_1x_n & x_2^2 & x_2x_3 & \dots & x_n^2 & \dots & x_1^k & \dots & x_n^k \end{bmatrix}^T \tag{2}$$

as a basis for the space of polynomials of degree at most k .

A polynomial $g \in \mathbb{R}[x_1, \dots, x_n]$ can be written

$$x \mapsto g(x) = \sum_{\alpha \in \mathbb{N}^n} g_\alpha x^\alpha$$

where

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

is a monomial of degree $|\alpha| = \sum_{i=1}^n \alpha_i$.

Following the methodology described in [12], we define for (generally non-convex) problem (1) a hierarchy $\{\mathbb{Q}_k\}$ of (convex) LMI relaxations

$$\mathbb{Q}_k \quad \begin{cases} p_k^* = \min_y \sum_{\alpha} (g_0)_{\alpha} y_{\alpha} \\ \text{s.t. } M_k(y) \succeq 0 \\ M_{k-d_i}(g_i y) \succeq 0, \quad i = 1, 2, \dots, m \end{cases} \quad (3)$$

where

- each decision variable y_{α} of $y = \{y_{\alpha}\}$ corresponds to a monomial x^{α} ,
- $M_k(y)$ is the positive semidefinite *moment matrix* of order k , and
- $M_{k-d_i}(y)$ is the positive semidefinite *localizing matrix* of order $k - d_i$ associated with the polynomial g_i , for all $i = 1, \dots, m$.

Solving the sequence $\{\mathbb{Q}_k\}$ of LMI relaxations (3) of increasing orders $k = d, d + 1, \dots$, it is proved in [12] that under some mild assumptions on the polynomials $\{g_i\}$, we obtain a monotone sequence of optimal values p_k^* converging asymptotically to the global optimal value p^* of the original optimization problem in (1), i.e. $p_k^* \uparrow p^*$ as $k \rightarrow \infty$. Experimental results reveal that in practice p_k^* is very close to p^* for relatively small values of k . In addition, in many cases the exact optimal value p^* is obtained at some particular relaxation \mathbb{Q}_k , that is, $p^* = p_k^*$ for some relatively small k .

GloptiPoly is a user-friendly Matlab/SeDuMi add-on to build and solve these LMI relaxations, see [10] and

www.laas.fr/~henrion/software/gloptipoly.

In this report we describe the algorithm used in GloptiPoly to detect whether the global optimum p^* in (1) has been reached at some LMI relaxation \mathbb{Q}_k in (3), i.e. whether $p_k^* = p^*$ for some index k . We also describe how to extract (one or several) global minimizers $x^* \in \mathbb{R}^n$ to original problem (1), given a solution y^* of the LMI relaxation \mathbb{Q}_k in (3).

Note that there exist a dual approach to build hierarchy of LMI relaxations, based on real algebraic geometry and sum-of-squares (SOS) decompositions of positive polynomials [16]. In contrast with the theory of moments which works in the space of measures on the primal space of solutions $x \in \mathbb{R}^n$, the SOS approach rather works in a (dual) space of polynomials, to obtain certificates ensuring validity of bounds on the objective function. As a result, and so far, in the latter approach there is no sufficient condition to check whether the exact optimal value is obtained, and no solution extraction mechanism.

In section 2 we state an algebraic condition ensuring global optimality of an LMI relaxation, and we describe the numerical linear algebra algorithm used to extract globally optimal solutions. In section 3 we mention some heuristics based on this algorithm that can be used to reduce significantly the number

of variables in the LMI relaxations. Finally, in section 4 we comment on a numerical behavior of GloptiPoly on unconstrained minimization problems. Illustrative numerical examples are inserted throughout the text.

2 Extracting Globally Optimal Solutions

2.1 Global Optimality Condition

Let y^* be an optimal solution of the LMI relaxation \mathbb{Q}_k in (3) (of order k). A sufficient rank condition ensuring global optimality of the LMI relaxation is

$$\text{rank } M_k(y^*) = \text{rank } M_{k-d}(y^*). \quad (4)$$

This condition can be checked numerically with the help of the singular value decomposition [8]. Note however that the rank condition (4) is not necessary, i.e. the global optimum p^* may have been reached at some LMI relaxation of order k (i.e., $p^* = p_k$), and yet $\text{rank } M_k(y_k^*) > \text{rank } M_{k-d}(y_k^*)$.

That condition (4) is sufficient to ensure that $p^* = p_k$ is a consequence of a deep result of Curto and Fialkow [6]. In our present context, if condition (4) is true, then by Theorem 1.6 in [6], y^* is the vector of moments of a rank $M_k(y^*)$ -atomic measure supported on the feasible set $\mathbb{K} = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, m\}$.

In the important special case where the feasible set \mathbb{K} can be written

$$\mathbb{K} = \{x \in \mathbb{R}^n \mid g_i(x) = 0, i = 1, \dots, n; g_{n+j}(x) \geq 0, j = 1, \dots, m\},$$

and the polynomial ideal $I = \langle g_1, \dots, g_n \rangle \subset \mathbb{R}[x_1, \dots, x_n]$ is zero-dimensional and *radical*, then condition (4) is guaranteed to hold at some index k . For instance this is the case for boolean (or 0-1) optimization problems, and more generally, bounded discrete optimization problems. For more details the interested reader is referred to [13, 14, 15, 17].

2.2 Extraction Algorithm

Assume that the LMI relaxation \mathbb{Q}_k in (3) has been solved, producing a vector y^* . Assume further that the rank condition (4) is satisfied. Then the main steps of the extraction algorithm can be sketched as follows.

Cholesky Factorization

As condition (4) holds, y^* is the vector of a rank $M_k(y^*)$ -atomic measure supported on \mathbb{K} . Hence, by construction of the moment matrix $M_k(y^*)$, we have

$$M_k(y^*) = \sum_{j=1}^r v_k(x^*(j))(v_k(x^*(j)))^T = V^*(V^*)^T$$

where

$$r = \text{rank } M_k(y^*) \tag{5}$$

and

$$V^* = [v_k(x^*(1)) \ v_k(x^*(2)) \ \cdots \ v_k(x^*(r))]$$

where $v_k(x)$ is as in (2), and $\{x^*(j)\}_{j=1}^r$ are r global minimizers of (1).

Extract a Cholesky factor V of the positive semidefinite moment matrix $M_k(y^*)$, i.e. a matrix V with r columns satisfying

$$M_k(y^*) = VV^T. \tag{6}$$

Such a Cholesky factor can be obtained via singular value decomposition, or any cheaper alternative [8].

Matrices V and V^* span the same linear subspace, so the solution extraction algorithm consists in transforming V into V^* by suitable column operations. This is described in the sequel.

Column Echelon Form

Reduce matrix V to column echelon form

$$U = \begin{bmatrix} 1 & & & & & & \\ x & & & & & & \\ 0 & 1 & & & & & \\ 0 & 0 & 1 & & & & \\ x & x & x & & & & \\ \vdots & & \ddots & & & & \\ 0 & 0 & 0 & \cdots & 1 & & \\ x & x & x & \cdots & x & & \\ \vdots & & & & \vdots & & \\ x & x & x & \cdots & x & & \end{bmatrix}$$

by Gaussian elimination with column pivoting [8]. By construction of the moment matrix, each row in U corresponds to a monomial x_α in polynomial basis v . Pivot elements in U (i.e. the first non-zero elements in each column) correspond to monomials x_{β_j} , $j = 1, 2, \dots, r$ of the basis generating the r solutions. In other words, if

$$w = [x_{\beta_1} \ x_{\beta_2} \ \dots \ x_{\beta_r}]^T \tag{7}$$

denotes this generating basis, then it holds

$$v = Uw \tag{8}$$

for all solutions $x = x^*(j)$, $j = 1, 2, \dots, r$.

In summary, extracting the solutions amounts to solving polynomial system of equations (8).

Solving the Polynomial System of Equations

Once a generating monomial basis is available, it turns out that extracting solutions of polynomial system of equations (8) amounts to solving a linear algebra problem.

As pointed out to us by Monique Laurent, this fact has been rediscovered many times. It is called Stickelberger theorem in textbook [19], and it is credited to Stetter and Müller in [4], see also the recent work [9]. The method can be sketched as follows.

Multiplication Matrices

For each first degree monomial $x_i, i = 1, 2, \dots, n$ extract from U the r -by- r multiplication matrix N_i containing coefficients of monomials $x_i x_{\beta_j}, j = 1, 2, \dots, r$ in generating basis (7), i.e. such that

$$N_i w = x_i w, \quad i = 1, 2, \dots, n. \tag{9}$$

Common Eigenvalues

As shown in [19], the entries of solutions $x^*(j), j = 1, 2, \dots, r$ are common eigenvalues of multiplication matrices $N_i, i = 1, 2, \dots, n$.

In order to compute these eigenvalues, we follow [4] and build a random combination of multiplication matrices

$$N = \sum_{i=1}^n \lambda_i N_i$$

where the $\lambda_i, i = 1, 2, \dots, n$ are non-negative real numbers summing up to one. Then, compute the ordered Schur decomposition [8]

$$N = QTQ^T \tag{10}$$

where

$$Q = [q_1 \ q_2 \ \dots \ q_r]$$

is an orthogonal matrix (i.e. $q_i^T q_i = 1$ and $q_i^T q_j = 0$ for $i \neq j$) and T is upper-triangular with eigenvalues of N sorted increasingly along the diagonal.

Finally, the i -th entry $x_i^*(j)$ of $x^*(j) \in \mathbb{R}^n$ is given by

$$x_i^*(j) = q_j^T N_i q_j, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, r. \tag{11}$$

2.3 Example

Consider the non-convex quadratic optimization problem [12, Ex. 5]

$$\begin{aligned}
 p^* &= \min_x -(x_1 - 1)^2 - (x_1 - x_2)^2 - (x_2 - 3)^2 \\
 \text{s.t. } &1 - (x_1 - 1)^2 \geq 0 \\
 &1 - (x_1 - x_2)^2 \geq 0 \\
 &1 - (x_2 - 3)^2 \geq 0.
 \end{aligned}$$

Applying the first ($k = 1$) LMI relaxation we obtain $p_1^* = -3$ and $\text{rank } M_1(y^*) = 3$.

With the second ($k = 2$) LMI relaxation we obtain $p_2^* = -2$ and $\text{rank } M_1(y^*) = \text{rank } M_2(y^*) = 3$. Rank condition (4) ensures global optimality, so $p^* = p_2^* = -2$.

The moment matrix of order $k = 2$ reads

$$M_2(y^*) = \begin{bmatrix} 1.0000 & 1.5868 & 2.2477 & 2.7603 & 3.6690 & 5.2387 \\ 1.5868 & 2.7603 & 3.6690 & 5.1073 & 6.5115 & 8.8245 \\ 2.2477 & 3.6690 & 5.2387 & 6.5115 & 8.8245 & 12.7072 \\ 2.7603 & 5.1073 & 6.5115 & 9.8013 & 12.1965 & 15.9960 \\ 3.6690 & 6.5115 & 8.8245 & 12.1965 & 15.9960 & 22.1084 \\ 5.2387 & 8.8245 & 12.7072 & 15.9960 & 22.1084 & 32.1036 \end{bmatrix}$$

and the monomial basis (2) is

$$v_2(x) = [1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1x_2 \quad x_2^2]^T.$$

The Cholesky factor (6) of the moment matrix is given by

$$V = \begin{bmatrix} -0.9384 & -0.0247 & 0.3447 \\ -1.6188 & 0.3036 & 0.2182 \\ -2.2486 & -0.1822 & 0.3864 \\ -2.9796 & 0.9603 & -0.0348 \\ -3.9813 & 0.3417 & -0.1697 \\ -5.6128 & -0.7627 & -0.1365 \end{bmatrix}$$

whose column echelon form reads (after rounding)

$$U = \begin{bmatrix} 1 \\ 0 \ 1 \\ 0 \ 0 \ 1 \\ -2 \ 3 \ 0 \\ -4 \ 2 \ 2 \\ -6 \ 0 \ 5 \end{bmatrix}.$$

Pivot entries correspond to the following generating basis (7)

$$w = [1 \quad x_1 \quad x_2]^T.$$

From the subsequent rows in matrix U we deduce from (8) that all the globally optimal solutions x satisfy the polynomial equations

$$\begin{aligned} x_1^2 &= -2 + 3x_1 \\ x_1x_2 &= -4 + 2x_1 + 2x_2 \\ x_2^2 &= -6 + 5x_2. \end{aligned}$$

Multiplication matrices (9) of monomials x_1 and x_2 in generating basis w are readily extracted from rows in U :

$$N_1 = \begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ -4 & 2 & 2 \end{bmatrix}, \quad N_2 = \begin{bmatrix} 0 & 0 & 1 \\ -4 & 2 & 2 \\ -6 & 0 & 5 \end{bmatrix}.$$

Then choose e.g.

$$N = 0.6909N_1 + 0.3091N_2 = \begin{bmatrix} 0 & 0.6909 & 0.3091 \\ -2.6183 & 2.6909 & 0.6183 \\ -4.6183 & 1.3817 & 2.9274 \end{bmatrix}$$

as a random combination of matrices N_1 and N_2 . The orthogonal matrix in Schur decomposition (10) is given by

$$Q = \begin{bmatrix} 0.4082 & 0.1826 & -0.8944 \\ 0.4082 & -0.9129 & -0.0000 \\ 0.8165 & 0.3651 & 0.4472 \end{bmatrix}.$$

From equations (11), we derive the 3 optimal solutions

$$x^*(1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad x^*(2) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad x^*(3) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

2.4 Numerical Stability

As shown in [8], all the operations of the solution extraction algorithm are numerically stable, except the Gaussian elimination step with column pivoting. Practical experiments with GloptiPoly however reveal that ill-conditioned problem instances leading to a failure of Gaussian elimination with column pivoting are very scarce. This experimental property of Gaussian elimination was already noticed in [8].

2.5 Number of Extracted Solutions

In virtue of relation (5), the number of solutions extracted by the algorithm is equal to the rank of the moment matrix. Up to our knowledge, when solving an LMI relaxation there is no easy way to control the rank of the moment matrix, hence the number of extracted solutions.

If there is no objective function in problem (1), by default GloptiPoly minimizes the trace of the moment matrix. As a result, its rank is indirectly minimized as well. Note however that, in contrast with trace minimization, rank minimization under LMI constraints is a difficult non-convex problem. Practical experiments reveal that low rank moment matrices are preferable from the numerical point of view: they ensure faster convergence to the global optimum. See also example 3.1 for an illustration of the impact of the trace minimization heuristic on the number of extracted solutions.

3 Applications of the Extraction Algorithm

When rank condition (4) is not satisfied, then we still can attempt to apply the extraction algorithm described in section 2. If the algorithm is successful and returns feasible solutions reaching the relaxed optimal value p_k^* , then by definition of the relaxation \mathbb{Q}_k , these solutions are global minimizers. This is the topic of section 3.1. Unfortunately, this heuristic does not work systematically, and extracted solutions can be infeasible, as illustrated with a counterexample in section 3.2.

If the algorithm is not successful, the column echelon form of the Cholesky factor of the moment matrix may contain useful information that can sometimes be exploited to reduce significantly the number of variables, hence the computational burden, in subsequent LMI relaxations. This heuristic is described in section 3.3.

3.1 Rank Condition Non Satisfied but Global Optimum Reached

Even though rank condition (4) is not satisfied, the extraction algorithm can be applied successfully, as shown by the following example.

Trace Minimization Heuristic

With the help of this example we also return to the comments of section 2.5 on the number of extracted solutions and the trace minimization heuristic.

Consider the polynomial system of equations [4, Ex. 5.2]

$$\begin{aligned} x_1^2 + x_2^2 - 1 &= 0 \\ x_1^3 + (2 + x_3)x_1x_2 + x_2^3 - 1 &= 0 \\ x_3^2 - 2 &= 0. \end{aligned}$$

There is no objective function to be minimized, so as indicated above GloptiPoly solves the LMI relaxations by minimizing the trace of the moment matrix.

Applying the least order ($k = 2$) LMI relaxation we obtain $\text{rank}M_1(y^*) = 4$ and $\text{rank}M_2(y^*) = 7$, so global optimum cannot be ensured via rank condition (4).

With the third LMI relaxation ($k = 3$) we obtain $\text{rank } M_1(y^*) = \text{rank } M_2(y^*) = \text{rank } M_3(y^*) = 2$, so rank condition (4) ensures global optimality.

From the extraction algorithm we derive the two globally optimal solutions

$$x^*(1) = \begin{bmatrix} 0.5826 \\ -0.8128 \\ -1.4142 \end{bmatrix}, \quad x^*(2) = \begin{bmatrix} -0.8128 \\ 0.5826 \\ -1.4142 \end{bmatrix}.$$

Now replacing the minimum trace LMI objective function in GloptiPoly with a zero objective function, the third LMI relaxation returns $\text{rank } M_1(y^*) = 4$ and $\text{rank } M_2(y^*) = \text{rank } M_3(y^*) = 6$, so rank condition (4) cannot ensure global optimality.

However, by applying the extraction algorithm, we are able to extract 6 solutions

$$\begin{aligned} x^*(1) &= \begin{bmatrix} -0.8128 \\ 0.5826 \\ -1.4142 \end{bmatrix}, & x^*(2) &= \begin{bmatrix} 0.5826 \\ -0.8128 \\ -1.4142 \end{bmatrix}, & x^*(3) &= \begin{bmatrix} 0.0000 \\ 1.0000 \\ -1.4142 \end{bmatrix}, \\ x^*(4) &= \begin{bmatrix} 1.0000 \\ 0.0000 \\ -1.4142 \end{bmatrix}, & x^*(5) &= \begin{bmatrix} 0.0000 \\ 1.0000 \\ 1.4142 \end{bmatrix}, & x^*(6) &= \begin{bmatrix} 1.0000 \\ 0.0000 \\ 1.4142 \end{bmatrix} \end{aligned}$$

thus proving global optimality of the LMI relaxation.

3.2 Infeasible Extracted Solutions

When rank condition (4) is not satisfied, it may happen that solutions extracted by the algorithm are infeasible for the original optimization problem. Since solutions are extracted from a convex LMI relaxation, they may be feasible for a subset of the original constraints only.

Example

We consider the polynomial systems of equations arising from a test for numerical bifurcation, originally described in [11] and listed in problem collection [2]:

$$\begin{aligned} 5x_1^9 - 6x_1^5x_2 + x_1x_2^4 + 2x_1x_3 &= 0 \\ -2x_1^6x_2 + 2x_1^2x_2^3 + 2x_2x_3 &= 0 \\ x_1^2 + x_2^2 &= 0.265625. \end{aligned}$$

This system has 8 distinct real solutions.

The lowest order ($k = 5$) LMI relaxation yields $\text{rank } M_1(y^*) = 3$ and $\text{rank } M_2(y^*) = \text{rank } M_3(y^*) = \text{rank } M_4(y^*) = \text{rank } M_5(y^*) = 4$. Since $d = 5$, rank condition (4) cannot ensure global optimality.

The extraction algorithm on moment matrix $M_2(y^*)$ returns 4 solutions

$$\begin{aligned}
 x^*(1) &= \begin{bmatrix} 0.3653 \\ -0.3636 \\ -0.0153 \end{bmatrix}, & x^*(2) &= \begin{bmatrix} 0.3653 \\ 0.3636 \\ -0.0153 \end{bmatrix}, \\
 x^*(3) &= \begin{bmatrix} -0.3653 \\ -0.3636 \\ -0.0153 \end{bmatrix}, & x^*(4) &= \begin{bmatrix} -0.3653 \\ 0.3636 \\ -0.0153 \end{bmatrix}.
 \end{aligned}$$

These solutions satisfy the second and third equations of the original problem, but not the first equation. Indeed, since the solutions are extracted from a convex relaxation of the original problem, they may be infeasible for a subset of the original constraints.

Proceeding with the 6th order LMI relaxation, we obtain $\text{rank } M_i(y^*) = 2$ for all $i = 1, 2, \dots, 6$, hence ensuring global optimality. The two extracted solutions are

$$x^*(1) = \begin{bmatrix} -0.2619 \\ 0.4439 \\ -0.0132 \end{bmatrix}, \quad x^*(2) = \begin{bmatrix} 0.2619 \\ 0.4439 \\ -0.0132 \end{bmatrix}.$$

3.3 Reducing the Number of LMI Variables

Suppose that at the LMI relaxation of order k , equation (8) holds for the solutions to be extracted, i.e. some monomials in standard basis (2) are expressed as linear combinations of monomials of generating basis (7).

If constraints of the original optimization problem become redundant when replacing linearly dependent monomials with combinations of generating monomials, then this results in a reduction of the monomial basis over which subsequent LMI relaxations are built. A similar idea is used in 0-1 quadratic problems to reduce the number of variables in successive LMI relaxations, see [14].

In summary, application of the reduction algorithm at earlier LMI relaxations – at which global optimality cannot be ensured with rank condition (4) – may result in a significant reduction of the problem dimensions. This can be seen as a (heuristic) alternative to the (systematic) algebraic reduction techniques of [7].

Example with Continuous Variables

Consider the following non-convex quadratic optimization problem suggested by Etienne de Klerk and Radina Dontcheva:

$$\begin{aligned}
 p^* &= \min_x -(x_1 - 1)^2 - (x_2 - 1)^2 - (x_3 - 1)^2 \\
 \text{s.t. } &1 - (x_1 - 1)^2 \geq 0 \\
 &1 - (x_2 - 1)^2 \geq 0 \\
 &1 - (x_3 - 1)^2 \geq 0
 \end{aligned} \tag{12}$$

whose global optimum is $p^* = -3$.

At the first ($k = 1$) LMI relaxation, the 4x4 moment matrix $M_1(y^*)$ has rank 4, so obviously no solution can be extracted. However, we obtain $p_1^* = -3$, so the global optimum is reached.

When $k = 2$, we have $\text{rank } M_1(y^*) = 4$ and $\text{rank } M_2(y^*) = 7$, and the column echelon form of the Cholesky factor of the 10x10 moment matrix $M_2(y^*)$ is given by

$$U = \begin{bmatrix} 1 & & & & & & & & & \\ & 0 & 1 & & & & & & & \\ & & 0 & 0 & 1 & & & & & \\ & & & 0 & 2 & 0 & & & & \\ & & & & 0 & 0 & 0 & 0 & 1 & \\ & & & & & 0 & 0 & 0 & 0 & 0 & 1 \\ & & & & & & 0 & 0 & 2 & 0 & 0 & 0 \\ & & & & & & & 0 & 0 & 0 & 0 & 0 & 1 \\ & & & & & & & & 0 & 0 & 0 & 2 & 0 & 0 & 0 \end{bmatrix}$$

in the monomial basis (2)

$$v_2(x) = [1 \quad x_1 \quad x_2 \quad x_3 \quad x_1^2 \quad x_1x_2 \quad x_1x_3 \quad x_2^2 \quad x_2x_3 \quad x_3^2]^T.$$

Pivot entries in matrix U correspond to the following generating basis (7)

$$w(x) = [1 \quad x_1 \quad x_2 \quad x_3 \quad x_1x_2 \quad x_1x_3 \quad x_2x_3]^T$$

which has 7 monomials.

From the rows in matrix U we deduce from (8) that solutions x to be extracted satisfy the polynomial equations

$$\begin{aligned} x_1^2 &= 2x_1 \\ x_2^2 &= 2x_2 \\ x_3^2 &= 2x_3. \end{aligned} \tag{13}$$

The extraction algorithm fails however, because third degree monomials are missing in U to build multiplication matrices (9).

Note however that when substituting monomials as in (13), constraints of the original problem (12) become redundant since $1 - (x_i - 1)^2 = -x_i^2 + 2x_i = 0 \geq 0$, for $i = 1, 2, 3$. We can therefore replace monomials x_i^2 with $2x_i$ and remove constraints in the next LMI relaxation.

So when $k = 3$, instead of having a basis (2) with 20 monomials of degree 3, we can use only 8 monomials to build the third LMI relaxation – with respect to the previous basis of the second LMI relaxation, the only new element is the third degree monomial $x_1x_2x_3$. Using 8 monomials instead of 20 reduces significantly the computational burden when solving the LMI relaxation. A further reduction is achieved since redundant constraints can be removed and the third LMI relaxation does not feature any localizing matrix.

When applying the reduction algorithm on the moment matrix $M_3(y^*)$ of rank 8, we obtain that monomial $x_1x_2x_3$ belongs to the generating basis. Multiplication matrices are readily obtained, and the 8 expected globally optimal solutions are extracted

$$\begin{aligned}
 x^*(1) &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, & x^*(2) &= \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, & x^*(3) &= \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, & x^*(4) &= \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \\
 x^*(5) &= \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, & x^*(6) &= \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, & x^*(7) &= \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}, & x^*(8) &= \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.
 \end{aligned}$$

Example with Discrete Variables

Consider the Max-Cut problem

$$\begin{aligned}
 \min & \quad -\frac{1}{2} \sum_{i < j} w_{ij}(1 - x_i x_j) \\
 \text{s.t.} & \quad x_i \in \{-1, +1\}
 \end{aligned}$$

in the case of a complete K5 graph with adjacency matrix

$$W = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

The first ($k = 1$) LMI relaxation yields $p_1^* = -6.25$ and $\text{rank } M_1(y^*) = 5$.

When $k = 2$ we obtain $p_2^* = -6.25$ and $\text{rank } M_1(y^*) = 5, \text{rank } M_2(y^*) = 10$.

When $k = 3$, we get $p_3^* = -6$ and $\text{rank } M_1(y^*) = 5, \text{rank } M_2(y^*) = 10, \text{rank } M_3(y^*) = 20$. The extraction algorithm returns

$$U = \begin{bmatrix} 1 \\ 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -2 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -2 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & -1 & 0 \\ 0 & -1 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

so that linearly dependent monomials in polynomial system of equations (9) are as follows

$$\begin{aligned} x_4x_5 &= -2 - x_1x_2 - x_1x_3 - x_1x_4 - x_1x_5 - x_2x_3 - x_2x_4 - x_2x_5 - x_3x_4 - x_3x_5 \\ x_1x_2x_3 &= x_4x_5 \\ x_1x_2x_4 &= x_1x_2x_3 \\ x_1x_2x_5 &= x_1x_2x_4 \\ x_1x_3x_4 &= x_1x_2x_5 \\ x_1x_3x_5 &= x_1x_3x_4 \\ x_1x_4x_5 &= -2x_1 - x_2 - x_3 - x_4 - x_5 - x_4x_5 - x_1x_2x_3 - x_1x_2x_4 - x_1x_2x_5 - x_1x_3x_4 \\ x_2x_3x_4 &= -x_1 - x_2 - x_3 - x_4 - x_4x_5 - x_1x_2x_3 - x_1x_2x_5 \\ x_2x_3x_5 &= -x_1 - x_2 - x_3 - x_5 - x_4x_5 - x_1x_2x_4 - x_1x_3x_4 \\ x_2x_4x_5 &= x_1 + x_3 + x_4x_5 + x_1x_2x_5 + x_1x_3x_4 \\ x_3x_4x_5 &= x_1 + x_2 + x_4x_5 + x_1x_2x_3 + x_1x_2x_4. \end{aligned}$$

From these relations, fourth degree monomials can be expressed in generating basis (7)

$$\begin{aligned}
 x_1x_2x_3x_4 &= (x_1x_2x_3)x_4 = (x_4x_5)x_4 = x_5 \\
 x_1x_2x_3x_5 &= (x_1x_2x_3)x_5 = (x_4x_5)x_5 = x_4 \\
 x_1x_2x_4x_5 &= (x_1x_2x_4)x_5 = (x_1x_2x_3)x_5 = (x_4x_5)x_4 = x_5 \\
 x_1x_3x_4x_5 &= (x_1x_3x_4)x_5 = (x_1x_2x_5)x_5 = x_1x_2 \\
 x_2x_3x_4x_5 &= (x_2x_3x_4)x_5 = (-x_1 - x_2 - x_3 - x_4 - x_4x_5 - x_1x_2x_3 - x_1x_2x_5)x_5 \\
 &= -x_1x_5 - x_2x_5 - x_3x_5 - x_4x_5 - 2x_4 - x_1x_2
 \end{aligned}$$

and the only fifth degree monomial readily follows

$$x_1x_2x_3x_4x_5 = (x_1x_2x_3x_4)x_5 = 1.$$

At this stage, it is useless to proceed with higher order LMI relaxations since no more linearly independent monomials of higher degree can be produced.

Consequently, the global optimum $p^* = p_3^* = -6$ has been reached and 20 globally optimal solutions can be extracted from the above matrix U .

4 A Remark on the Numerical Behavior of GloptiPoly

Finally, we want to comment on a nice and surprising behavior of GloptiPoly that we observed on some examples of unconstrained minimization.

In the case of unconstrained global minimization, that is when \mathbb{K} is R^n , only one LMI relaxation is useful, namely $M_k(y)$ if $\deg g_0 = 2k$ or $2k - 1$. Indeed,

- (a) either $g_0 - p^*$ is SOS and then $p_k^* = p^*$, or
- (b) $g_0 - p^*$ is *not* SOS and then $p_{k+j}^* = p_k^* < p^*$ for all $j = 1, 2, \dots$

Therefore there is no need to try relaxations with orders higher than k . However, in case (b) it may be worthy to still try higher order relaxations! Indeed, because of the numerical inaccuracies involved in the solving procedure, one may obtain convergence in a finite number of steps to a value and minimizers, very close to the exact value and the exact minimizers respectively! Let us try to explain why.

If the space of polynomials $x \mapsto g(x) = \sum_{\alpha} g_{\alpha}x^{\alpha}$ is equipped with the norm $\|g\| = \sum_{\alpha} |g_{\alpha}|$, then the cone Σ_n of SOS polynomials is dense in the set of polynomials nonnegative over the multidimensional box $[-1, 1]^n$, see e.g. [1].

Therefore, consider a nonnegative polynomial g_0 that is *not* SOS, and assume that g_0 has a global minimizer $x^* \in [-1, 1]^n$ with $g_0(x^*) = p^*$. Then, one may hope that an SOS polynomial g_k , *close* to g_0 (i.e., with $\|g_k - g_0\| < \epsilon$) will provide a global minimizer close to x^* . Observe that for all $x \in [-1, 1]^n$,

$$|g_k(x) - g_0(x)| = \left| \sum_{\alpha} [(g_k)_{\alpha} - (g_0)_{\alpha}] x^{\alpha} \right| \leq \|g_k - g_0\| \leq \epsilon.$$

However, one does not know how to construct such a sequence of SOS polynomials $\{g_k\}$ with $\|g_k - g_0\| \rightarrow 0$.

But let us see how GloptiPoly behaves on the following well-known example of a non-negative polynomial which is not SOS, namely the polynomial obtained by dehomogenization of Motzkin’s form:

$$g_0(x) = \frac{1}{27} + x_1^2 x_2^2 (x_1^2 + x_2^2 - 1).$$

This polynomial is nonnegative ($p^* = \min_x g_0(x) = 0$ attained at $|x_1| = |x_2| = \sqrt{3}/3$) but is not SOS.

- The least order ($k = 3$) LMI relaxation is unbounded, returning no useful information. In principle one should stop here; we have detected that g_0 is not SOS.
- when $k = 4$ the LMI relaxation is unbounded too.
- when $k = 5$ the LMI relaxation returns $p_5^* = -0.4036$ and all the moment matrices have full rank (in GloptiPoly we use a relative threshold of 10^{-3} to evaluate the numerical rank of a matrix)
- when $k = 6$ the LMI relaxation returns $p_6^* = -0.08241$ and all the moment matrices have full rank
- when $k = 7$ the LMI relaxation returns $p_7^* = -0.01675$ and all the moment matrices have full rank
- when $k = 8$ the LMI relaxation returns an almost zero optimum $p_8^* = 3.022 \cdot 10^{-10}$, and $\text{rank } M_1(y^*) = 3$, $\text{rank } M_2(y^*) = \text{rank } M_3(y^*) = 4$, thus proving global optimality.

The moment matrix of second order reads

$$M_2(y^*) = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.3333 & 0.0000 & 0.3333 \\ 0.0000 & 0.3333 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.3333 & 0.0000 & 0.0000 & 0.0000 \\ 0.3333 & 0.0000 & 0.0000 & 0.1111 & 0.0000 & 0.1111 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.1111 & 0.0000 \\ 0.3333 & 0.0000 & 0.0000 & 0.1111 & 0.0000 & 0.1111 \end{bmatrix}$$

from which we readily extract the four globally optimal solutions

$$\begin{aligned} x^*(1) &= \begin{bmatrix} -0.5773 \\ -0.5773 \end{bmatrix}, & x^*(2) &= \begin{bmatrix} 0.5773 \\ -0.5773 \end{bmatrix}, \\ x^*(3) &= \begin{bmatrix} -0.5773 \\ 0.5773 \end{bmatrix}, & x^*(4) &= \begin{bmatrix} 0.5773 \\ 0.5773 \end{bmatrix}. \end{aligned}$$

From the dual LMI [12, 16], we can obtain the SOS decomposition

$$g_8(x) = \sum_{i=1}^{32} a_i^2 q_i^2(x) + \varepsilon r(x) \approx g_0(x) \quad \text{on } [-1, 1]^2,$$

where

- polynomials $q_i(x)$ and $r(x)$ are normalized such that their coefficient vectors have unit Euclidean norm,
- $\varepsilon \leq 10^{-8} < a_i^2$, i.e. positive scalar parameter ε is less than a given threshold, and positive scalar coefficients a_i^2 in the decomposition are greater than the threshold,
- $\deg q_i(x) \leq 8$, since GloptiPoly solved the eighth LMI relaxation,
- there are 32 (!) terms in the SOS decomposition.

The above SOS decomposition is approximate in the sense that parameter ε is small, but non-zero. It turns out that in GloptiPoly numerical inaccuracy (roundoff errors) helped to find a higher degree SOS polynomial g_8 close to Motzkin polynomial on $[-1, 1]^2$.

Thus, everything looks like if in the solving procedure of the dual relaxation \mathbb{Q}_k^* (see [12] for notations) the constraints

$$\langle X, B_\alpha \rangle = (g_0)_\alpha, \quad |\alpha| \leq 2k,$$

are replaced *automatically* by

$$\langle X, B_\alpha \rangle = (g_0)_\alpha + \epsilon_\alpha, \quad |\alpha| \leq 2k,$$

with *appropriate* small perturbations $\{\epsilon_\alpha\}$, chosen by the solver!

In a similar vein, it can be useful to add a redundant constraint of the type

$$g_1(x) = R^2 - \|x\|_2^2 \geq 0,$$

and consider the optimization problem $\min\{g_0(x) \mid g_1(x) \geq 0\}$, to obtain guaranteed convergence of the successive associated LMI relaxations.

Now consider problem (1) where $g_0(x)$ is the above Motzkin polynomial and $g_1(x)$ is the above radius constraint with $R = 1$ (to include the 4 global minima). With GloptiPoly we obtain already at the third LMI relaxation the SOS decomposition

$$g_0(x) = \sum_{i=1}^6 a_i^2 q_i^2(x) + g_1(x) \sum_{i=1}^2 b_i^2 r_i^2(x)$$

with only 6 and 2 terms such that $\deg q_i \leq 3$ and $\deg r_i \leq 2$, respectively.

5 Conclusion

Solution extraction is straightforward when the moment matrix has rank-one: in this case the solution vector is equal to the first order moment vector. When the moment matrix has rank greater than one, we have proposed in section 2 a systematic extraction procedure, implemented in version 2.2 of the GloptiPoly software.

The extraction algorithm is applied when moment matrices satisfy rank condition (4), in which case it is always successful and yields globally optimal solutions. However, as pointed out in section 3, when the rank condition is not satisfied, a heuristic consists in applying the extraction algorithm anyway. Either the algorithm is successful and we are done (see section 3.1) or the algorithm fails, but still some information can be exploited to reduce the number of variables in subsequent LMI relaxations (see section 3.3). Note however that these ideas are not currently implemented in GloptiPoly.

Note finally that an incomplete extraction procedure was sketched in [3] in the case of LMI relaxations for polynomial systems of equations, and partly motivated us to devise a more general algorithm. A specific extraction procedure was also described in [18, Section 5] in the case of quadratic optimization problems with one (possibly non-convex) quadratic constraint, or one linear constraint jointly with one concave quadratic constraint.

Acknowledgements.

This work benefited from numerous discussions with Etienne de Klerk, Monique Laurent, Arnold Neumaier and Jos Sturm, whose suggestions, references, and numerical examples are gratefully acknowledged.

References

1. Berg C (1987). The multidimensional moment problem and semi-groups. *Moments in Mathematics. Proc. Symp. Appl. Math.*, 37:110–124.
2. Bini D, Mourrain B (1998). Polynomial Test Suite. SAGA Project, INRIA Sophia Antipolis, France.
3. Chesi G, Garulli A, Tesi A, Vicino A (2000). An LMI-based Approach for Characterizing the Solution Set of Polynomial Systems. *Proc. IEEE Conf. Decision and Control*, 1501–1506, Sydney, Australia.
4. Corless R M, Gianni P M, Trager B M (1997). A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots. *Proc. ACM Int. Symp. Symbolic and Algebraic Computation*, 133–140, Maui, Hawaii.
5. Curto R E, Fialkow L A (1991). Recursiveness, positivity, and truncated moment problems. *Houston J. Math.* 17:603–635.
6. Curto R E, Fialkow L A (2000). The truncated complex K -moment problem. *Trans. Amer. Math. Soc.* 352:2825–2855.
7. Gatermann K, Parrilo P A (2004). Symmetry Groups, Semidefinite Programs, and Sums of Squares. *J. Pure and Appl. Algebra*, 192(1-3):95–128.
8. Golub G H, Van Loan C F (1996). *Matrix computations*. 3rd edition. The Johns Hopkins University Press, NY.
9. Jibeteau D (2003). *Algebraic Optimization with Applications to System Theory*. PhD Thesis, Centrum voor Wiskunde en Informatica (CWI), Amsterdam.
10. Henrion D, Lasserre J B (2003). GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi. *ACM Trans. Math. Software*, 29(2):165–194.

11. Kearfott R B (1987). Some Tests of Generalized Bisection. *ACM Trans. Math. Software*, 13(3):197–220.
12. Lasserre J B (2001). Global Optimization with Polynomials and the Problem of Moments. *SIAM J. Opt.* 11(3):796–817.
13. Lasserre J B (2001). Polynomials nonnegative on a grid and discrete optimization. *Trans. Amer. Math. Soc.* 354:631–649.
14. Lasserre J B (2002). An Explicit Equivalent Positive Semidefinite Program for Nonlinear 0-1 Programs. *SIAM J. Opt.* 12(3):756–769.
15. Laurent M (2002). Semidefinite representations for finite varieties. Preprint. Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 2002.
16. Parrilo P A (2003). Semidefinite Programming Relaxations for Semialgebraic Problems. *Math. Prog. Ser. B.* 96:293–320.
17. Parrilo P A (2002). An explicit construction of distinguished representations of polynomials nonnegative over finite sets. Preprint. ETH Zurich, Switzerland.
18. Sturm J F, Zhang S (2003). On Cones of Nonnegative Quadratic Functions. *Math. Op. Res.* 28:246–267.
19. Sturmfels B (2002). *Solving Systems of Polynomial Equations*. Amer. Math. Soc., Providence, RI.