

ON THE LASSERRE HIERARCHY OF SEMIDEFINITE PROGRAMMING RELAXATIONS OF CONVEX POLYNOMIAL OPTIMIZATION PROBLEMS

ETIENNE DE KLERK* AND MONIQUE LAURENT†

Abstract. The Lasserre hierarchy of semidefinite programming approximations to convex polynomial optimization problems is known to converge finitely under some assumptions. [J.B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* **19**, 1995–2014, 2009.] We give a new proof of the finite convergence property, that does not require the assumption that the Hessian of the objective be positive definite on the entire feasible set, but only at the optimal solution. In addition, we show that the number of steps needed for convergence depends on more than the input size of the problem. In particular, the size of the semidefinite program that gives the exact reformulation of the convex polynomial optimization problem may be exponential in the input size.

Key words. convex polynomial optimization, sum of squares of polynomials, positivstellensatz, semidefinite programming.

AMS subject classifications. 90C60, 90C56, 90C26

1. Polynomial optimization and the Lasserre hierarchy. We consider the polynomial optimization problem

$$p_{\min} = \min_{x \in \mathbb{R}^n} \{p_0(x) : p_i(x) \geq 0 \ (i = 1, \dots, m)\}, \quad (1.1)$$

where each $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 0, \dots, m$) is a polynomial. We denote the highest total degree of the polynomials p_0, \dots, p_m by d . We partition the index set $\{1, \dots, m\} =: \mathcal{I}_l \cup \mathcal{I}_n$ to differentiate between (affine) linear and nonlinear constraints, where \mathcal{I}_l consists of the indices i for which p_i is an affine or linear polynomial.

We denote the polynomials with real coefficients in the variables x by $\mathbb{R}[x]$. The subset of $\mathbb{R}[x]$ consisting of the sums of squares of polynomials is denoted by Σ^2 .

The feasible set of problem (1.1) is denoted by \mathcal{F} , i.e:

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid p_i(x) \geq 0 \ (i = 1, \dots, m)\}. \quad (1.2)$$

We assume that \mathcal{F} is compact so that problem (1.1) is guaranteed to have a minimizer.

The quadratic module generated by the polynomials p_i ($i = 1, \dots, m$) is defined as:

$$\mathcal{M}(p_1, \dots, p_m) := \left\{ \sigma_0 + \sum_{i=1}^m \sigma_i p_i \mid \sigma_i \in \Sigma^2 \ (i = 0, \dots, m) \right\}. \quad (1.3)$$

The truncated quadratic module of degree $2t$, denoted as $\mathcal{M}_t(p_1, \dots, p_m)$, is defined as the subset of $\mathcal{M}(p_1, \dots, p_m)$ where the sum of squares polynomials $\sigma_0, \dots, \sigma_m$ meet the additional conditions:

$$\deg(\sigma_0) \leq 2t, \deg(\sigma_i p_i) \leq 2t \quad (i = 1, \dots, m). \quad (1.4)$$

*Tilburg University, Department of Econometrics and Operations Research, 5000 LE Tilburg, Netherlands. Email: E.deKlerk@uvt.nl

†Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, and Tilburg University, Department of Econometrics and Operations Research, 5000 LE Tilburg, Netherlands. Email: M.Laurent@cwi.nl

Lasserre [10] introduced the following hierarchy of approximations to p_{\min} :

$$\rho_t := \max\{\lambda \mid p_0 - \lambda \in \mathcal{M}_t(p_1, \dots, p_m)\} \quad (1.5)$$

and showed that, under some assumptions, $\lim_{t \rightarrow \infty} \rho_t = p_{\min}$. Moreover, for each fixed t , ρ_t may be computed as the optimal value of a semidefinite program. In particular, this may be done in polynomial time to any fixed accuracy.

1.1. Convex polynomial optimization. Lasserre [11] recently showed that the hierarchy of approximations (1.5) exhibits finite convergence for certain classes of convex polynomial optimization problems (Theorem 3.4 in [11]).

We will prove finite convergence for convex polynomial optimization problems that meet the following conditions.

ASSUMPTION 1. *We make the following assumptions on problem (1.1):*

1. *The polynomials $p_0, -p_1, \dots, -p_m$ are convex;*
2. *The Slater condition holds:*

$$\exists x_0 \in \mathbb{R}^n : p_i(x_0) > 0 \text{ for } i \in \mathcal{I}_n \text{ and } p_i(x_0) \geq 0 \text{ for } i \in \mathcal{I}_l.$$

3. *The quadratic module $\mathcal{M}(p_1, \dots, p_m)$ is Archimedean:*

$$\exists R > 0 : R^2 - \|x\|^2 \in \mathcal{M}(p_1, \dots, p_m).$$

4. *$\nabla^2 p_0(x^*) \succ 0$ (i.e. the Hessian of p_0 at x^* is positive definite) if x^* is a minimizer of (1.1).*

Note that the third assumption implies that \mathcal{F} is compact (since it is contained in the ball $\{x : \|x\| \leq R\}$). Moreover, we may assume without loss of generality that $\|x\| < R$ for all $x \in \mathcal{F}$, since $\bar{R}^2 - \|x\|^2 \in \mathcal{M}(p_1, \dots, p_m)$ for all $\bar{R} \geq R$.

The fourth assumption implies that the minimizer of (1.1) is unique. It is a weaker assumption than the corresponding assumption in Theorem 3.4 of Lasserre [11] which requires that $\nabla^2 f(x) \succ 0 \forall x \in \mathcal{F}$. For example, consider the problem

$$\min_{x \in [-1, 1]} x^4 + 2x.$$

Here the Hessian is not positive definite at $x = 0$, but it is positive definite at the global minimizer $x^* = -2^{-1/3}$.

2. Convex optimization and the Farkas lemma. The following result is known as the extended (or convex) Farkas lemma; see [7] for a survey on the topic. A proof of the result (in this form) is e.g. given in [9, §2.2.3].

THEOREM 2.1 (Farkas). *Let f, g_1, \dots, g_m be given convex functions defined on a nonempty convex set \mathcal{C} , and assume that the Slater regularity condition is satisfied. The inequality system*

$$\begin{aligned} f(x) &< 0 \\ g_j(x) &\leq 0, \quad j = 1, \dots, m \\ x &\in \mathcal{C}, \end{aligned}$$

has no solution if and only if there exists a vector $\bar{y} \in \mathbb{R}_+^m$ such that

$$f(x) + \sum_{j=1}^m \bar{y}_j g_j(x) \geq 0 \quad \text{for all } x \in \mathcal{C}. \quad (2.1)$$

3. Finite convergence of the Lasserre hierarchy.

3.1. The general convex case. The aim in this section is to give a proof of the finite convergence result by Lasserre (Theorem 3.4 in [11]) under weaker assumptions. This result applies to a polynomial optimization problem of the form (1.1), where $p_0, -p_1, \dots, -p_m$ are convex polynomials (and satisfy some additional regularity condition).

A key lemma that we will need is the following *Positivstellensatz* by Scheiderer [16].

PROPOSITION 3.1 (Example 3.18 in [16]). *Let $p \in \mathbb{R}[x]$ be a polynomial for which the level set*

$$\mathcal{K} := \{x \in \mathbb{R}^n \mid p(x) \geq 0\}$$

is compact. Let $q \in \mathbb{R}[x]$ be nonnegative on \mathcal{K} . Assume that the following conditions hold:

1. *q has only finitely many zeros in \mathcal{K} , each lying in the interior of \mathcal{K} .*
2. *the Hessian $\nabla^2 q$ is positive definite at each of these zeroes.*

Then $q = \sigma_0 + p\sigma_1$ for some $\sigma_0, \sigma_1 \in \Sigma^2$.

We now prove the main result of this section, namely that the Lasserre SDP hierarchy has finite convergence for problem (1.1) under Assumption 1.

THEOREM 3.2 (cf. Theorem 3.4 in [11]). *Consider the polynomial optimization problem (1.1). Under Assumption 1, one has:*

$$p_0 - p_{\min} \in \mathcal{M}(p_1, \dots, p_m),$$

where the quadratic module $\mathcal{M}(p_1, \dots, p_m)$ was defined in (1.3).

Proof. We will apply the extended Farkas lemma (Theorem 2.1) with

$$f := p_0 - p_{\min}, \quad g_i := -p_i \quad (i = 1, \dots, m), \quad \text{and } \mathcal{C} := \{x \mid \|x\| \leq R\}.$$

By construction $f(x) \geq 0$ on the set $\{x \in \mathcal{C} \mid g_j(x) \leq 0 \ (j = 1, \dots, m)\}$, and the Slater assumption in Theorem 2.1 is met. Thus, by Theorem 2.1, there exists a $\bar{y} \in \mathbb{R}_+^m$ such that

$$p_0(x) - p_{\min} - \sum_{i=1}^m \bar{y}_i p_i(x) \geq 0 \quad \forall x \in \mathcal{C}.$$

We now show that the function

$$q(x) := p_0(x) - p_{\min} - \sum_{i=1}^m \bar{y}_i p_i(x) \tag{3.1}$$

has a unique root in \mathcal{C} and this root lies in the interior of \mathcal{C} . Indeed, $q(x^*) = 0$ at the minimizer x^* of problem (1.1), so that x^* is a minimizer of q over \mathcal{C} and a root of q in \mathcal{C} . As $\nabla^2 q(x^*) \succeq \nabla^2 p_0(x^*) \succ 0$, x^* is the unique minimizer of q in \mathcal{C} , which implies that x^* is the unique root of q in \mathcal{C} . Moreover, x^* lies in the interior of \mathcal{C} since we have assumed that $\|x\| < R$ for all $x \in \mathcal{F}$.

We may now apply Proposition 3.1 with $p(x) := R^2 - \|x\|^2$ and q as defined in (3.1), to conclude that

$$p_0(x) - p_{\min} = \sum_{i=1}^m \bar{y}_i p_i(x) + \sigma_0(x) + \sigma_1(x)(R^2 - \|x\|^2)$$

for some $\sigma_0, \sigma_1 \in \Sigma^2$. Since $R^2 - \|x\|^2 \in \mathcal{M}(p_1, \dots, p_m)$ by assumption, we obtain the required result. \square

REMARK 3.1. *Note that Theorem 3.2 remains valid under different constraint qualifications. For instance, instead of assuming the existence of a Slater point (as in Assumption 1 part 2), we may require the Mangasarian-Fromovitz constraint qualification:*

$$\exists w \in \mathbb{R}^n \quad w^T \nabla p_i(x^*) > 0 \quad \forall i \in J^*, \quad (3.2)$$

where x^* is a minimizer of (1.1) and $J^* = \{i \in \{1, \dots, m\} \mid p_i(x^*) = 0\}$ is the set of indices corresponding to the active constraints at x^* . Indeed, under (3.2), there exist multipliers $\bar{y}_i \geq 0$ for which $\nabla p_0(x^*) - \sum_i \bar{y}_i \nabla p_i(x^*) = 0$ and $\bar{y}_i p_i(x^*) = 0$ for all i (see e.g. [14, §12.6]).

Consider now, as before, the polynomial $q := p_0 - p_{\min} - \sum_i \bar{y}_i p_i$. As q is convex and $\nabla q(x^*) = 0$, x^* is a global minimizer of q over \mathbb{R}^n and thus $q \geq q(x^*) = 0$ on \mathbb{R}^n . We can now proceed as in the rest of the proof of Theorem 3.2.

REMARK 3.2. *The assumption that the Hessian of p_0 should be positive definite at the minimizer cannot be omitted in Theorem 3.2.*

To see this, consider the problem

$$p_{\min} = \min_{x \in \mathbb{R}^n} \{p_0(x) : 1 - \|x\|^2 \geq 0\}, \quad (3.3)$$

where p_0 is a convex form (i.e. homogeneous polynomial) of degree at least 4 that is not a sum of squares.

Then $p_{\min} = 0$. Indeed, convex n -variate forms are necessarily nonnegative on \mathbb{R}^n , since their gradients vanish at zero. On the other hand, they are not always sums of squares, as was shown by Blekherman [3].¹

By construction, problem (3.3) satisfies all the assumptions of Theorem 3.2, except for the positive definiteness of the Hessian at the minimizer.

Assume we have finite convergence of the Lasserre hierarchy for problem (3.3), i.e.

$$p_0 \in \Sigma^2 + (1 - \|x\|^2)\Sigma^2.$$

By Proposition 4 in De Klerk, Laurent and Parrilo [8], a form belongs to $\Sigma^2 + (1 - \|x\|^2)\Sigma^2$ if and only if it is a sum of squares. This contradicts our assumption that $p_0 \notin \Sigma^2$.

3.2. The convex quadratic case. Consider problem (1.1) in the special case when p_0, p_1, \dots, p_m are quadratic polynomials. Then, as shown in [10], the finite convergence result from Theorem 3.2 can be sharpened to show that the first relaxation in the hierarchy is exact. Moreover, we do not need to use Scheiderer's result (Proposition 3.1) since, as is well known, any nonnegative quadratic polynomial is a sum of squares.

THEOREM 3.3. [10] *Let $p_0, -p_1, \dots, -p_m$ be convex quadratic polynomials. Assume that the feasible set \mathcal{F} (as in (1.2)) is compact and let x^* be a minimizer of problem (1.1). Assume moreover that either there is a Slater point or x^* satisfies (3.2). Then, the Lasserre relaxation of order 1 is exact, i.e., ρ_1 is equal to the minimum of (1.1).*

¹It is interesting to note that Blekherman's proof is not constructive, and no actual examples are known of convex forms that are not sums of squares.

Proof. The first part of the proof is identical to that of Theorem 3.2 (or Remark 3.1), permitting to find multipliers $\bar{y}_i \geq 0$ for which $q = p_0 - p_{\min} - \sum_i \bar{y}_i p_i$ is nonnegative on $\mathcal{C} = \mathbb{R}^n$. Now, as q is quadratic polynomial, we can conclude directly that q is a sum of squares. \square

3.3. The bivariate quartic case. We now consider problem (1.1) in the special case when there is only one constraint ($m = 1$) and p_0, p_1 are quartic bivariate polynomials (i.e. $n = 2$). Again we do not need to use Scheiderer's result (Proposition 3.1), as we can instead use Hilbert's result claiming that any nonnegative bivariate quartic polynomial is a sum of squares.

THEOREM 3.4. *Let $p_0, -p_1$ be bivariate, quartic, convex polynomials and assume that $p_1(x_0) > 0$ for some $x_0 \in \mathbb{R}^2$. If p_0 is nonnegative on $\mathcal{F} = \{x \mid p_1(x) \geq 0\}$, then there exists a $\bar{y} \geq 0$ for which $p_0 - \bar{y}p_1$ is a sum of squares. Moreover, if p has a minimizer over \mathcal{F} , then the second relaxation in Lasserre's hierarchy is exact, i.e., ρ_2 is equal to the minimum of (1.1).*

Proof. Directly applying Farkas' lemma (Theorem 2.1) combined with the above mentioned result of Hilbert. \square

From this we can derive a result of Henrion [6] which gives an explicit semidefinite representation of $\text{cl}(\text{conv}(\mathcal{F}))$, the closure of the convex hull of the set \mathcal{F} . Clearly $\text{cl}(\text{conv}(\mathcal{F}))$ can be described as the intersection of all its supporting hyperplanes, which correspond to all linear polynomials nonnegative on \mathcal{F} . Let \mathcal{H} denote the subset consisting of all linear polynomials of the form $\sigma + \lambda p_1$, where $\lambda \geq 0$ and σ is a sum of squares of polynomials (thus of degree at most 4). Then

$$\text{cl}(\text{conv}(\mathcal{F})) \subseteq S(\mathcal{F}) := \{x \in \mathbb{R}^2 \mid f(x) \geq 0 \forall f \in \mathcal{H}\}.$$

The set $S(\mathcal{F})$ admits an explicit semidefinite programming formulation.

COROLLARY 3.5. [6] *Let p_1 be a bivariate concave quartic polynomial and assume that $p_1(x_0) > 0$ for some $x_0 \in \mathbb{R}^2$. Then $\text{cl}(\text{conv}(\mathcal{F})) = S(\mathcal{F})$.*

Proof. Directly from Theorem 3.4 as every linear polynomial nonnegative on \mathcal{F} belongs to \mathcal{H} . \square

4. Complexity results. A natural question is whether it is possible to give a bound on the (finite) number of steps required for convergence by the Lasserre hierarchy for problem (1.1) under Assumption 1.

Before addressing this question, we briefly discuss known complexity results for convex polynomial optimization, in order to place the discussion in the correct context.

4.1. Recognizing convex problems. A first point to make is that it is NP-hard in the Turing model of computation (described in e.g. [5]) to decide if a given instance of problem (1.1) is a convex optimization problem, due to the following result.

THEOREM 4.1 (Ahmadi et al. [1]). *It is strongly NP-hard in the Turing model of computation to decide if a given form of degree $d \geq 4$ is quasi-convex.*

4.2. Complexity results via the ellipsoid method. The best known complexity result for solving problem (1.1) under Assumption 1 is by using the ellipsoid method of Yudin-Nemirovski. For given $\epsilon > 0$, the ellipsoid algorithm can compute an ϵ -feasible² x such that $|p_0(x) - p_{\min}| \leq \epsilon$ in at most

$$O\left(n^2 \ln\left(\frac{R}{\epsilon}\right)\right)$$

²We call x ϵ -feasible for problem (1.1) if the ball of radius ϵ and centered at x intersects the feasible set \mathcal{F} .

iterations, where each iteration requires the evaluation of the polynomials p_0, \dots, p_m as well as the gradient of p_0 and of one polynomial that is negative at the current iterate (in order to obtain a separating hyperplane); see e.g. [2, §5.2].

It will be convenient to only consider the real number model (also known as BSS model) of computation [4]. In the real number model, the input is a finite set of real numbers, and an arithmetic operation between two real numbers requires one unit of time. Thus, the size of the input of problem (1.1) may be expressed by four numbers:

1. n , the number of variables;
2. m , the number of constraints;
3. d , the largest total degree of p_0, \dots, p_m ;
4. the total number of nonzero coefficients of the polynomials p_0, \dots, p_m in the standard monomial basis, say $L := \sum_{i=0}^m L_i$, where L_i is the number of nonzero coefficients of p_i .

Note that

$$m + 1 \leq L \leq (m + 1) \binom{n + d}{d},$$

and the exact value of L depends on the sparsity of the polynomials p_0, \dots, p_m .

The n -variate polynomial p_i of total degree at most d may be evaluated in at most $O(dL_i)$ arithmetic operations. Thus the total complexity of the ellipsoid method becomes

$$O\left(dLn^2 \ln\left(\frac{R}{\epsilon}\right)\right). \quad (4.1)$$

Note that the ellipsoid algorithm uses the parameter R (and not only the fact that it is finite).

Also, neither the Slater assumption, nor the assumption that the Hessian of the objective is positive definite at a minimizer, is required by the ellipsoid method (cf. Assumption 1).

Finally, note that the number of constraints m only enters the complexity bound (4.1) implicitly, via the value L .

4.3. The rank of the Lasserre hierarchy. We now return to the question of giving a bound on the (finite) number of steps required for convergence of the Lasserre hierarchy for problem (1.1) under Assumption 1.

Recall that the Lasserre hierarchy computes the values ρ_t in (1.5) as the optimal value of suitable semidefinite programs. The size of the semidefinite program that yields ρ_t is as follows: it has $m + 1$ positive semidefinite matrix variables of order $\binom{n+t}{t}$, and there are $\binom{n+2t}{2t}$ linear equality constraints; see [10] or the survey [12] for details on the semidefinite programming reformulations.

In particular, ρ_t may be computed to ϵ relative accuracy in at most

$$O\left(\left((m + 1) \binom{n + 2t}{2t}\right)^{\frac{9}{2}} \ln\left(\frac{1}{\epsilon}\right)\right)$$

arithmetic operations using interior point algorithms; see e.g. [2, §6.6.3].

Note that this bound is only polynomial in (n, m, d, L) if $t = O(1)$.

We will call the smallest value of t such that $\rho_t = p_{\min}$ (see (1.1)), the *rank* of the Lasserre hierarchy.

We now show that, in a well-defined sense, the rank of the Lasserre hierarchy must depend on more than just the input size (n, m, d, L) of problem (1.1).

THEOREM 4.2. *Consider problem (1.1) under Assumption 1. If $\deg(p_0) \geq 4$, there is no integer constant $C > 0$ depending only on (n, m, d, L) , such that the Lasserre hierarchy converges in C steps.*

Proof. The proof uses a similar construction as in Remark 3.2. As in Remark 3.2, let p be a convex, n -variate form of degree d that is not a sum of squares.

We consider the behavior of the Lasserre hierarchy for the sequence of problems:

$$\min_{x \in \mathbb{R}^n} \left\{ p(x) + \frac{1}{k} \|x\|^2 \mid p_1(x) := 1 - \|x\|^2 \geq 0 \right\} \quad \text{for } k = 1, 2, \dots \quad (4.2)$$

By construction, for each k , problem (4.2) meets Assumption 1. By Theorem 3.2, the Lasserre hierarchy therefore converges in finitely many steps for problem (4.2) for each $k = 1, 2, \dots$

Assume now that there exists an integer $t > 0$ such that

$$p + \frac{1}{k} \|x\|^2 \in \mathcal{M}_t(p_1) \quad \forall k,$$

where $\mathcal{M}_t(p_1)$ is the truncated quadratic module of degree $2t$ generated by p_1 (see (1.4)). As the set $\{x : p_1(x) \geq 0\}$ has a nonempty interior, the set $\mathcal{M}_t(p_1)$ is closed (see [15]). As a consequence, the limit p of the sequence $p + \frac{1}{k} \|x\|^2$ (as k tends to ∞) must also belong to $\mathcal{M}_t(p_1)$. As explained in Remark 3.2, this contradicts the assumption that p is not a sum of squares. \square

In the construction used in the proof of Theorem 4.2, the smallest eigenvalue of the Hessian of the objective function in (4.2) at the minimizer $x^* = 0$ tends to zero as $k \rightarrow \infty$. This suggests that the rank of the Lasserre hierarchy may depend on the value of the smallest eigenvalue of the Hessian at the minimizer x^* . The smallest eigenvalue of the Hessian at x^* may in turn be viewed as a ‘condition number’ of the problem that is independent of (n, m, d, L) .

5. Conclusion and summary. We have given a new proof of the finite convergence of the Lasserre hierarchy for convex polynomial optimization problems, under weaker assumptions than were known before (Theorem 3.2). In Remark 3.2 we showed that our new assumption, namely that the Hessian of the objective is positive definite at the minimizer, is necessary for finite convergence.

We have also looked at the possibility of bounding the rank of the finite convergence, and gave a negative result about the dependence of such a bound on the problem data. In particular, we showed that the number of steps needed for convergence cannot be bounded by a quantity that depends only on the input size (in the real number model of computation). As a consequence, the worst-case complexity bound for solving problem (1.1) under Assumption 1 to fixed accuracy is not polynomial in the input size for the Lasserre hierarchy, in contrast to the ellipsoid method. Having said that, it is important to remember that the number of operations required by the ellipsoid method will typically equal the worst-case bound, whereas the Lasserre hierarchy can converge quickly for some convex problems (as we reviewed in Sections 3.2 and 3.3). Moreover, the worst-case complexity bound for the Lasserre hierarchy could possibly be improved by deriving error bounds on $p_{\min} - \rho_t$ (see (1.5)) in terms of t . For general polynomial optimization problems, deriving explicit error bounds for the

Lasserre hierarchy has proved difficult so far (see [13]), but the additional convexity assumption may simplify this analysis.

Acknowledgements. The authors would like to thank Amir Ali Ahmadi and Marianna Nagy for their valuable comments.

REFERENCES

- [1] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. Preprint, MIT, Cambridge, MA, 2010.
- [2] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.
- [3] G. Blekherman. Convex forms that are not sums of squares. arXiv:0910.0656v1, October 2009.
- [4] L. Blum, M. Shub, and S. Smale: On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. Amer. Math. Soc. (N.S.)* **21**(1) (1989), 1–46.
- [5] M.R. Garey and D.S. Johnson. *Computers and intractability: a guide to the theory of NP-completeness*. W.H. Freeman and Company, Publishers, San Francisco, USA, 1979.
- [6] D. Henrion. On linear and semidefinite programming relaxations of convex polynomial optimization problems. Preprint, 2008.
- [7] V. Jeyakumar. Farkas lemma: Generalizations. in *Encyclopedia of Optimization II*, C.A. Floudas and P. Pardalos Eds., Kluwer, Dordrecht, 87–91, 2001.
- [8] E. Klerk, M. Laurent, P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. In D. Henrion and A. Garulli (Eds.), *Positive Polynomials in Control*. Berlin: Springer Verlag. (LNCIS, 312), 2005.
- [9] E. de Klerk, C. Roos, and T. Terlaky. *Nonlinear Optimization*. Lecture Notes, 2005. Available at: <http://www.isa.ewi.tudelft.nl/~roos/courses/WI4207%20Optimization%20TU/wi387dic.pdf>
- [10] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, **11**:296–817, 2001.
- [11] J.B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization* **19**, 1995–2014, 2009.
- [12] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, Vol. 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivan (eds.), Springer, pages 157–270, 2009.
- [13] J. Nie and M. Schweighofer. On the complexity of Putinar’s Positivstellensatz. *Journal of Complexity*, **23**(1):135–150, 2007.
- [14] J. Nocedal and S. Wright. *Numerical Optimization*, Springer, 2000.
- [15] V. Powers and C. Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1:71–88, 2001.
- [16] C. Scheiderer. Sums of squares on real algebraic curves. *Mathematische Zeitschrift*, **245**, 725–760, 2003.