Global Optimization with Polynomials

Deren Han
Singapore-MIT Alliance
National University of Singapore
Singapore 119260
Email: smahdr@nus.edu.sg

Abstract— **The class of POP (Polynomial Optimization Problems) covers a wide rang of optimization problems such as $0-1$ integer linear and quadratic programs, nonconvex quadratic programs and bilinear matrix inequalities. In this paper, we review some methods on solving the unconstraint case: minimize a real-valued polynomial $p(x): \quad R^n \to R$, as well the constraint case: minimize $p(x)$ on a semialgebraic set K, i.e., a set defined by polynomial equalities and inequalities. We also summarize some questions that we are currently considering.**

I. Introduction

A polynomial p in $x_1, \cdots, x_n$ is a finite combination of monomials:

$$p(x) = \sum_{\alpha} c_\alpha x^\alpha = \sum_{\alpha} c_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad c_\alpha \in R,$$

where the sum is over a finite number of n-tuples $\alpha = (\alpha_1, \cdots, \alpha_n)$, α_i is a nonnegative integer. In this paper, we will consider the problem **P**:

$$p^* = \min_{x \in R^n} p(x),$$

where $p(x): \ R^n \to R$ is a real-valued polynomial. That is, finding the global minimum p^* of $p(x)$ and a minimizer x^*. We will also consider the constraint case P_K:

$$p_K^* = \min_{x \in K} p(x),$$

where K is a semialgebraic set defined by polynomial equalities and inequalities $g_i(x) \geq 0, \ i = 1, \cdots, m$, which includes many interesting applications and standard problems such as $0-1$ integer linear and quadratic programs as particular cases.

For the problem **P**, exact algebraic algorithms find all the critical points and then comparing the values of p at these points. We will discuss these methods in Section 2, which include Gröbner bases, resultants, eigenvalues of companion matrices [4], and numerical homotopy methods [16], [31].

A classic approach for P_K (also can be used to **P**) is convex relaxation methods. In recent years, there are various relaxation methods that have been studied intensively and extensively. For the $0-1$ integer program, a *lift-and-project* linear programming procedure by Bala, Ceria and Cornuéjols [1], *The reformulation-linearization technique* (RLT) by Sherali and Adams [24] and an SDP (*Semidefinite Programming*) relaxation method by Lovász-Schrijver [15] were regarded as their pioneering works. They have been modified, generalized and extended to various problems and methods. Most recently, some new SDP relaxation methods were proposed by Lasserre [12] and Parrilo [19], [20], and Kim and Kojima [9] showed that their *Second-Order-Cone-Programming* (SOCP) relaxation is a reasonable compromise between the effectiveness of the SDP relaxation and the low computational burden of the lift-and-project LP relaxation or RLT. We will discuss these relaxation methods in Section 3 and will complete the paper in Section 4 by giving some conclusion.

II. Solving Polynomial Equations

In this section, we will discuss computational algebraic methods for the problem **P**. These results are based on [19], [2] and [7]. For solving this problem, one often look at the first order conditions, which form a system of (nonlinear) equations.

A. Preliminary Notions and Notation

Throughout the paper, we suppose that $1 \leq n$ is an integer, C^n and R^n respectively denote the complex and real n-space, and x is the abbreviation of $(x_1, \cdots, x_n)$. We let $R[x]$ and $C[x]$ denote the ring of polynomials in n indeterminates with real and complex coefficients, respectively. We first recall some definitions and results regarding the solution set of system of polynomial equations.

Definition 1: The set $I \subseteq C[x]$ is an ideal if it satisfies:

1) $0 \in I$;
2) If $a, b \in I$, then $a + b \in I$;
3) If $a \in I$ and $b \in C[x]$, then $a \cdot b \in I$.

Definition 2: Given a set of polynomials $p_1, \cdots, p_s \in R[x]$, define the set

$$\langle p_1, \ \cdots, \ p_s \rangle = \{ f_1 p_1 + \cdots + f_s p_s : \ f_i \in R[x], \ i = 1, \cdots, s \}.$$

It can be easily shown that the set $\langle p_1, \ \cdots, \ p_s \rangle$ is an ideal, known as the ideal generated by $p_1, \ \cdots, \ p_s$.

The set of all simultaneous solutions in C^n of a system of equations

$$\{ x | \ p_1(x) = p_2(x) = \cdots = p_s(x) = 0 \}$$

is called the affine variety defined by $p_1, \ \cdots, \ p_s$, denoted by $V(p_1, \ \cdots, \ p_s)$. Given a polynomial ideal I we let

$$V(I) = \{ x \in C^n | f(x) = 0, \ \forall f \in I \}$$

as the affine variety associated with I.

B. Gröbner bases and Stetter-Möller Method

Obviously, any finite set of polynomials generated a polynomial ideal. Due to the Hilbert's Nullstellensatz, the converse is also true: any polynomial ideal I is generated by a finite set of polynomials, which is called a *basis* for I. Usually, the generated set is not unique. For a given term order $\prec$ on the polynomial ring $R[x]$, any nontrivial ideal has a unique monic *reduced Gröbner basis* [2], [4]. Let $\mathcal{G} = (g_1, g_2, \cdots, g_r)$ be a Gröbner basis for the critical ideal

$$I = \langle \frac{\partial p}{\partial x_1}, \frac{\partial p}{\partial x_2}, \cdots, \frac{\partial p}{\partial x_n} \rangle$$

with respect to $\prec$. Then, the elements of the quotient space $C[x]/I$ have the form $[f] = \hat{f} + I$ and $\hat{f} \in C[x]$ is unique:

$$\hat{f} = f - (f_1 g_1 + \cdots + f_r g_r), \quad f_i \in C[x], \quad i = 1, \cdots, r$$

and no term of $\hat{f}$ is divisible by any of the leading terms of the elements of $\mathcal{G}$. Obviously, the remainder $\hat{f} = 0$ if and only if $f \in I$ and polynomials in the same class have the same remainder.

Theorem 1: Let $I \subseteq C[x]$ be an ideal. The following conditions are equivalent:

a. The vector space $C[x]/I$ is finite dimensional.
b. The associate variety $V(I)$ is a finite set.
c. If $\mathcal{G}$ is a Gröbner basis for I, then for each i, $1 \le i \le r$, there is a $k_i \ge 0$ such that $x_i^{k_i}$ is the leading term of g for some $g \in \mathcal{G}$.

A monomial $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is *standard* if it is not divisible by the leading term of any element in the Gröbner basis $\mathcal{G}$. Let $\mathcal{B}$ be the set of standard monomials, then, it is a basis for the residue ring $C[x]/I$. For $f \in C[x]$, an arbitrary polynomial, define the endomorphism

$$A_f : C[x]/I \to C[x]/I, \qquad A_f([g]) = [fg].$$

The endomophism is represented in the basis $\mathcal{B}$ by a $\mu \times \mu$ matrix A_f, where μ is the number of elements of $\mathcal{B}$. The entry of A_f with row index $x^\alpha \in \mathcal{B}$ and column index $x^\beta \in \mathcal{B}$ is the coefficient of x^β in the normal form $x^\alpha f(x)$ with respect to $\mathcal{G}$.

The Stetter-Möller method [17] (also known as eigenvalue method) is to compute symbolically the matrix A_p and A_{x_i}, $i = 1, \cdots, n$, then compute numerically its eigenvalus and corresponding eigenvectors of A_p. Then, determine p^* and x^* according to the following result, which follows from Lemma 2.1 and Theorem 4.5 of [4].

Theorem 2: [19]. The optimal value p^* is the smallest real eigenvalue of the matrix A_p. Any eigenvector of A_p with eigenvalue p^* defines an optimal point $x^* = (x_1^*, \cdots, x_n^*)$ by the eigenvector identities $A_{x_i} \cdot v = p_i \cdot v$ for $i = 1, \cdots, n$.

C. Resultants

Let t be a new indeterminate and form the disciminant of the polynomial $p(x) - t$ with respect to $x_1, \cdots, x_n$:

$$\delta(t) := \Delta_x(p(x) - t)$$

and Δ_x is the $A - discriminant$, defined in [6], where A is the support of p together with the origin. From [6] we have that the discriminant δ equals the characteristic polynomial of the matrix A_f and

Theorem 3: The optimal value p^* is the smallest real root of $\delta(t)$.

The method of resultant is to compute $\delta(t)$, and minimal polynomials for the coordinates x_i^* of the optimal point, by elimination of variables using matrix formulas for resultants and discriminants [6].

D. Homotopy Methods

For the problem **P**, the critical equations form a squre system with n indeteminates and n equations. For solving such a square system, many *numerical homotopy continuation methods* were introduced, see for example [16], [31]. The basic idea of this class methods is to introduce a deformation parameter τ into the system such that the system at $\tau = 0$ breaks up into several systems and each of which consists of binomials. Thus, the system at $\tau = 0$ is easy to solve and the methods then trace the full solution set (with μ paths, μ is the Bézout's number) to $\tau = 1$.

If the system under consideration is sparse, then we usually use *polyhedral homotopies* which take the Newton polytops of the given equation into consideration. Under this case, the number μ is the *mixed volume* of the Newton polytopes [4], which is usually smaller than Bézout number.

III. Relaxation Methods

In the above section, we have reviewed some methods for the unconstrained global polynomial optimization problem **P**. The three classes of methods share the same feature that their running time is controlled by the number μ of complex critical points: In the Stetter-Möller method, we need to solve the eigenvalue-eigenvector problem on matrices with size $\mu \times \mu$; in the resultants methods, we must solve a univariate polynomial with degree μ; in the homotopy methods, we must trace μ paths from $\tau = 0$ to $\tau = 1$. These methods become infeasible if μ is large; this is the case even for small n or small total degree $2d$ of p, since $\mu = (2d - 1)^n$, which increases rapidly with n and $2d$. For example, when $n = 9$ and $2d = 4$, then $\mu = 1953125$ (see Table 1 in =[19]).

Various convex relaxation methods have been studied intensively and extensively in recent years, such as the *lift-and project* method for integer programs [1], [15], the *reformulation-linearization technique* of Sherali-Adams [24], [25], Sherali-Tuncbilek [26], [27], the semidefinite programming relaxation of Lasserre [12], [13] and Pariilo [19], [20], the second order cone programming relaxation of Kim and Kojima [8], [9]. In this section, we will review these methods detailly.

A. Linear Programming Relaxation

Let $\delta > 0$ be an integer. In the reformulation-linearization technique of Sherali and Tuncbilek [26], They first reformulate the constraints to the form

$$g_1(x)^{\alpha_1} g_2(x)^{\alpha_2} \cdots g_m(x)^{\alpha_m} \geq 0, \qquad |\alpha| := \sum_{i=1}^m \alpha_i \leq \delta, \tag{1}$$

which contain the bound factor product constraints ($0 \leq x_i \leq 1$) as well as the original constraints. Then, introducing a new variable y_α for each term in the objective function p and the new constraints, we obtain a linear programming, a relaxation of problem P_K:

$$P_\delta \to \min_y \{c_\delta^\top y \mid A_\delta y \geq b_\delta \text{ from (1) for every } |\alpha| \leq \delta\}. \tag{2}$$

The following results shows the reasonable of the LP relaxation. Here we assume with no loss of generality that the constant term of $p(x)$ is zero, i.e., $p(0) = 0$. For the proof, see [14].

Theorem 4: Consider the constraint polynomial optimization problem P_K and the LP relaxation P_δ in (2) defined from (1). Let ρ_δ be its optimal value:

(a) For every δ, $\rho_\delta \leq p^*$ and

$$p(x) - \rho_\delta = \sum_{|\alpha| \leq \delta} b_\alpha(\delta) g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m}, \tag{3}$$

for some nonnegative scalars $\{b_\alpha(\delta)\}$. Let x^* be a global minimizer of P_K and let $I(x^*)$ be the set of active constraints at x^*. If $I(x^*) = \emptyset$ (i.e., x^* is in the interior of the constraint set K) or if there is some feasible, nonoptimal solution $x \in K$ with $g_i(x) = 0,\ \forall i \in I(x^*)$, then $\rho_\delta < p^*$ for all δ, that is, no relaxation P_δ can be exact.

(b) If all the g_i are linear, that is, if K is a convex polytope, then (3) holds and $\rho_\delta \uparrow p^*$ as $\delta \to \infty$. If $I(x^*) = \emptyset$ for some global minimizer x^*, then in (3)

$$\sum_\alpha b_\alpha(\delta) \to 0 \quad \text{as } \delta \to \infty. \tag{4}$$

B. Semidefinite Programming Relaxation

The SDP relaxation of POP was inroduced by N.Z. Shor [29] and was recently further extended by Lasserre [12] and Parrilo [20]. Theoretically, it provides a lower bound of **P** or P_K while in practice it frequently agrees with the optimal value.

Let

$$\begin{array}{l} 1,\ x_1,\ x_2,\ \cdots, x_n,\ x_1^2,\ x_1x_2,\ \cdots,\ x_1x_n, \\ x_2^2,\ x_2x_3,\ \cdots,\ x_n^2,\ \cdots,\ x_1^r,\ \cdots,\ x_n^r \end{array} \tag{5}$$

be a basis of a real-valued polynomial of degree at most r and let $s(r)$ be its length.

The unconstraint POP is equivalent to

$$\max\ \lambda, \qquad s.t. \quad p(x) - \lambda \geq 0,\ \forall x \in R^n.$$

This is a very hard problem and we usually relax it to

$$\max\ \lambda, \qquad s.t. \quad p(x) - \lambda \ \text{ is sos} \tag{6}$$

where sos is the abbreviation of *sum of squares*. Now, we can assume that the degree of p is $2d$. Let X denote the column vector whose elements are as (5) with degree d. The length of X is

$$N = \begin{pmatrix} n+d \\ d \end{pmatrix}.$$

Let $\mathcal{L}_p$ denote the set of all real symmetric $N \times N$ matrix A such that $p(x) = X^\top AX$ and let E_{11} denote the matrix unit whose only nonzero entry is one on the upper left corner.

Theorem 5: [19] For any real number λ, the following two are equivalent:

1) The polynomial $p(x) - \lambda$ is a sum of squares in $R[x]$.
2) There is a matrix $A \in \mathcal{L}$ such that $A - \lambda E_{11}$ is positive semidefinite, that is, all eigenvalues of $A - \lambda E_{11}$ are nonnegative reals.

From this theorem, we can see that (6) is a semidefinite programming, which can be solved in polynomial time by interior point methods [18], [30]. For fixed n or for fixed d, The length N of X is polynomial of n, which, together with the above theorem, means that we can find the largest number λ of (6), denote by p^{sos}, in polynomial time. We always have that $p^{sos} \leq p^*$ and the inequality may be strict. An example is Motzkin's polynomial [23]

$$m(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2.$$

We can prove that $m(x, y) \geq -1$ but for any real number λ, $m(x, y) - \lambda$ is not sos, which means that $p^{sos} = -\infty$.

For the constraint case, we can find the largest number λ, such that $p(x) - \lambda \geq 0,\ \forall x \in K$. This condition is then relaxed to

$$p(x) - \lambda = u_0(x) + \sum_{j=1}^m u_j(x) g_j(x)$$

and

$$u_j(x) \text{ is sos},\ j = 0, \cdots, m.$$

This also leads to a semidefinite programming relaxation:

$$\begin{aligned} p^{sos} = \quad & \max \quad \lambda \\ & s.t. \quad p(x) - \lambda = u_0(x) + \sum_{j=1}^m u_j(x) g_j(x) \\ & \qquad\quad u_0,\ u_1,\ \cdots,\ u_m \text{ sos}. \end{aligned}$$

From a dual point of view, Lasserre [12] develop another SDP relaxation. Replace **P** and P_K with the equivalent problem

$$\mathcal{P} \to p^* := \max_{\mu \in \mathcal{P}(R^n)} \int p(x) \mu(dx)$$

and

$$\mathcal{P}_K \to p^* := \max_{\mu \in \mathcal{P}(K)} \int p(x) \mu(dx),$$

respectively, where $\mathcal{P}(R^n)$ and $\mathcal{P}(K)$ are the space of finite Borel signed measures on R^n and K, respectively. Then, the criterion to minimize is a linear criterion $a^\top y$ on the finite collection of moments $\{y_\alpha\}$, up to order m, the degree of p, of the probability measure μ. The problem is then how to describe the conditions on y to be a sequence of moments. For the history and recent development on the theory of moments, one is referred to [3], [5], [21] and references therein.

Lasserre [12] then relax $\mathcal{P}$ to the following SDP:

$$\mathcal{Q} \to \begin{cases} \inf_y & \sum_\alpha p_\alpha y_\alpha \\ s.t. & M_m(y) \succeq 0 \end{cases} \tag{7}$$

where $M_m(y)$ is the moment matrix of dimension $s(m)$ with rows and columns labelled by (5). Equivalently, (7) can be written as

$$\mathcal{Q} \to \begin{cases} \inf_y & \sum_\alpha p_\alpha y_\alpha \\ s.t. & \sum_{\alpha \neq 0} y_\alpha B_\alpha \succeq B_0 \end{cases} \tag{8}$$

where B_α and B_0 are easily understood from the definition of $M_m(y)$. The dual program of $\mathcal{Q}$ is

$$\mathcal{Q}^* \to \begin{cases} \sup_X & \langle X, -B_0\rangle (= -X(1,1)) \\ s.t. & \langle X, B_\alpha\rangle = p_\alpha \\ & X \succeq 0, \end{cases} \tag{9}$$

where X is a real-valued sysmmetric matrix and $\langle A, B\rangle$ is the Frobenius inner produce

$$\langle A, B\rangle = tr(AB) = \sum_{i,j=1}^{n} A_{ij} B_{ij}.$$

Lasserre proved that

Theorem 6: Assume that $\mathcal{Q}^*$ has a feasible solution. Then $\mathcal{Q}^*$ is solvable and there is no duality gap, that is

$$\inf \mathcal{Q} = \sup \mathcal{Q}^*.$$

Under some conditions, the relaxation is exact:

Theorem 7: Let $p(x)$: $R^n \to R$ be a $2m$-degree polynomial with global minimum p^*.

1) If the nonnegative polynomial $p(x) - p^*$ is a sum of squares of other polynomials, then **P** is equivalent to the semidefinite programming $\mathcal{Q}$ (7). More precisely, $\min \mathcal{Q} = p^*$ and if x^* is a global minimizer of **P**, then the vector

$$y^* := (x_1^*, \cdots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \cdots, (x_1^*)^{2m}, \cdots, (x_1^*)^{2m})$$

is a minimizer of $\mathcal{Q}$.
2) Conversely, if $\mathcal{Q}^*$ has a feasible solution, then $p^* = \min \mathcal{Q}$ only if $p(x) - p^*$ is a sum of squares.

As we have known from [23] and the above discussion that $p(x) - p^*$ may not be a sos. Then, suppose we know in advance that a global minimizer x^* of $p(x)$ has norm less than r for some $r > 0$, then, using the fact [3] that every polynomial $f(x) > 0$ on $K_r := \{x \mid r^2 - \|x\|^2 \geq 0\}$ can be written as

$$f(x) = \sum_{i=1}^{r_1} q_i(x)^2 + (r^2 - \|x\|^2) \sum_{j=1}^{r_2} t_j(x)^2$$

for some polynomials $q_i(x)$, $t_j(x)$, $i = 1, \cdots, r_1$, $j = 1, \cdots, r_2$. For every $N \geq m$, let

$$\mathcal{Q}_r^N \to \begin{cases} \inf_y & \sum_\alpha p_\alpha y_\alpha \\ s.t. & M_N(y) \succeq 0 \\ & M_{N-1}(\theta y) \succeq 0 \end{cases} \tag{10}$$

($\theta(x) = r^2 - \|x\|^2$) be the new relaxation. The dual of (10) is

$$\left(\mathcal{Q}_r^N\right)^* \to \begin{cases} \sup_{X,Z} & -X(1,1) - r^2 Z(1,1) \\ s.t. & \langle X, B_\alpha\rangle + \langle Z, C_\alpha\rangle = p_\alpha, \ \alpha \neq 0 \\ & X, \ Z \succeq 0, \end{cases} \tag{11}$$

where X, Z are real-valued sysmmetric matrices.

Lasserre [12] proved that

Theorem 8: Let $p(x)$: $R^n \to R$ be a $2m$-degree polynomial with global minimum p^* and $\|x^*\| \leq r$ for some $r > 0$ at some global minimizer x^*. Then

1) As $N \to \infty$, we have

$$\mathcal{Q}_r^N \uparrow p^*.$$

Moreover, for N sufficiently large, there is no duality gap between $\mathcal{Q}_r^N$ and its dual $(\mathcal{Q}_r^N)^*$, and the dual is solvable.
2) $\min \mathcal{Q}_r^N = p^*$ if and only if

$$p(x) - p^* = \sum_{i=1}^{r_1} q_i(x)^2 + (r^2 - \|x\|^2) \sum_{j=1}^{r_2} t_j(x)^2$$

for some polynomials $q_i(x)$ of degree at most N, and $t_j(x)$ of degree at most $N - 1$, $i = 1, \cdots, r_1$, $j = 1, \cdots, r_2$. In this case, the vector

$$y^* := (x_1^*, \cdots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \cdots, (x_1^*)^{2N}, \cdots, (x_1^*)^{2N})$$

is a minimizer of $(\mathcal{Q}_r^N)$. In addition, $\max \left(\mathcal{Q}_r^N\right)^* = \min \left(\mathcal{Q}_r^N\right)$ and for every optimal solution (X^*, Z^*) of $\left(\mathcal{Q}_r^N\right)^*$,

$$p(x) - p^* = \sum_{i=1}^{r_1} \lambda_i q_i(x)^2 + (r^2 - \|x\|^2) \sum_{j=1}^{r_2} \gamma_j t_j(x)^2,$$

where the vectors of coefficients of the polynomials $q_i(x)$, $t_j(x)$ are the eigenvectors of X^* and Z^* with respective to eigenvalues λ_i, γ_j.

In a similar way, Lasserre [12] deduced the following SDP relaxation for P_K:

$$\mathcal{Q}_K^N \to \begin{cases} \inf_y & \sum_\alpha p_\alpha y_\alpha \\ s.t. & M_N(y) \succeq 0 \\ & M_{N-\tilde{\omega}_i}(g_i y) \succeq 0, \ i = 1, \cdots, m, \end{cases} \tag{12}$$

where $\tilde{\omega}_i := \lceil \omega_i / 2 \rceil$ is the smallest integer larger than $\omega_i/2$, the degree of g_i and $N \geq \max\{\lceil m/2 \rceil, \max_i \tilde{\omega}_i\}$. Writing

$M_{N-\tilde{\omega}_i}(g_i y) = \sum_\alpha C_{i\alpha} y_\alpha$, the dual program is:

$$(\mathcal{Q}_K^N)^* \to \left\{ \begin{array}{ll} \sup_{X,Z_i} & -X(1,1) - \sum_{i=1}^m Z_i(1,1) \\ s.t. & \langle X, B_\alpha \rangle + \sum_{i=1}^m \langle Z_i, C_{i\alpha} = p_\alpha, \\ & \qquad \alpha \neq 0 \\ & X,\ Z_i \succeq 0,\ i = 1, \cdots, m. \end{array} \right. \tag{13}$$

Lasserre proved the following convergence result:

Theorem 9: Let $p(x)$: $R^n \to R$ be a m-degree polynomial with global minimum p_K^* and the compact set K is archimedean. Then

1) As $N \to \infty$, we have
$$\mathcal{Q}_K^N \uparrow\ p_K^*.$$
Moreover, for N sufficiently large, there is no duality gap between $\mathcal{Q}_K^N$ and its dual $(\mathcal{Q}_K^N)^*$ if K has a nonempty interior.
2) If $p(x) - p_K^*$ has the representation
$$p(x) - p_K^* = \sum_{i=1}^{r_1} q_i(x)^2 + (r^2 - \|x\|^2) \sum_{j=1}^{r_2} t_j(x)^2$$
for some polynomials $q_i(x)$ of degree at most N, and $t_j(x)$ of degree at most $N - \tilde{\omega}_i$, $i = 1, \cdots, r_1$, $j = 1, \cdots, r_2$, then $\min\ \mathcal{Q}_K^N = p_K^* = \max\ (\mathcal{Q}_K^N)^*$ and the vector
$$y^* := (x_1^*, \cdots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \cdots, (x_1^*)^{2N}, \cdots, (x_1^*)^{2N})$$
is a minimizer of $(\mathcal{Q}_K^N)$. In addition, for every optimal solution $(X^*, Z_1^*, \cdots, Z_m^*)$ of $(\mathcal{Q}_K^N)^*$,
$$p(x) - p_K^* = \sum_{i=1}^{r_1} \lambda_i q_i(x)^2 + \sum_{j=1}^{m} g_j(x) \sum_{i=1}^{r_i} \gamma_{ij} t_{ij}(x)^2,$$
where the vectors of coefficients of the polynomials $q_i(x)$, $t_{ij}(x)$ are the eigenvectors of X^* and Z_{ij}^* with respective to eigenvalues λ_i, γ_{ij}.

C. Second Order Cone Programming Relaxation

Lasserre [14] showed that the RLT of Sherali and Tuncbilek [26], [27] used implicity the Hausdorff moment conditions. Comparing with SDP relaxation, the LP relaxation has the following drawbacks:

1) The binomial coefficients involved in the reformulated constraints (see (3)), the Hausdorff moment condition numerically not stable.
2) In contrast the SDP relaxation, the asymptotic convergence of the LP relaxation is not guaranteed in general.
3) Even in the case of a convex polytope K, the LP relaxations cannot be exact in general.

On the other hand, LP software packages can handle very large-size problems, while the present status of SDP software packages excludes their uses in practice. Recently, Kim and Kojima [8], [9] showed that their SOCP relaxation is a reasonable compromise between the effectiveness of the SDP relaxation and the low computation cost of LP relaxation.

Their method is for solving nonconvex quadratic programs and the basic idea is simple: they just replaced the semedefinite condition $X \succeq 0$ by a necessary condition
$$(X_{kj})^2 \leq X_{kk} X_{jj}.$$
In some case, this relaxation is as powerful as the original condition $X \succeq 0$, while the computational cost is much less than SDP.

D. Tighter Relaxation by Redundant Constraints

By adding redundant constraints, tighter bound for the original problem can be found. Recently, Kojima, Kim and Waki [10] gave a general framework for convex relaxation of polynomial optimization over cones. They summarized that we can add two classes valid inequalities to the original problem to enhance the relaxation: Universally valid polynomial constraints and deduced valid inequalities. We say that a constraint is universally valid if it holds for any $x \in R^n$.

1) *Universally valid polynomial constraints*. Let u be a mapping from R^n into R^m whose jth component u_j is a polynomial in x. Then the $m \times m$ matrix $u(x)u(x)^\top$ is positive semidefinite for all $x \in R^n$. We can add the constraint $u(x)u(x)^\top \in S_+^m$ to the original problem.
Another universally valid constraint is the second order cone constraint. Let u_1 and u_2 be two mappings from R^n into R^m whose jth component is a polynomial in x. By the Cauthy-Schwarz inequality, we see that
$$(u_1(x)^\top u_2(x))^2 \leq (u_1(x)^\top u_1(x))(u_2(x)^\top u_2(x)),$$
which can be converted to
$$\begin{pmatrix} u_1(x)^\top u_1(x) + (u_2(x)^\top u_2(x) \\ u_1(x)^\top u_1(x) - (u_2(x)^\top u_2(x) \\ 2u_1(x)^\top u_2(x) \end{pmatrix} \in \mathcal{N}_2^3,$$
where $\mathcal{N}_2^3$ denotes the 3-dimensional second cone.
2) *Deduced valid inequalities*. We can also deduce valid inequalities from the original constraints. For example, in the RLT, they added the products of the original inequalities. Kojima, Kim and Waki [10] summarize some technique of this class, including Kronecker products of positive semidefinite matrix cones, Hadamard products of p-order cones ($p \geq 1$), linear transformation of cones, quadratic convexity and constraints from numerical computation.

IV. Conclusion

In this review paper, we have summarized the current development of the global polynomial optimization problems, constrained and unconstrained. There are many methods for this class of problems, algebraically and numerically. The algebraic methods usually provide good approximation of the optimal value as well as the global minimizer while the computation cost is huge. The LP, SDP and SOCP are well-developed and they can be used as convex approximation of the original nonconvex problem. Among the three convex relaxation methods, the SDP the most attractive but the status of its software packages exclude it from utilization, LP is mostly used in practice for large-size problems and SOCP is a compromise between the effectiveness of SDP and efficiency of LP.

Sherali and Tuncbilek [26], [28] have combined their LP relaxation with other global optimization methods such as *branch-and-bound*. Since that SDP relaxation will outperform than LP for small-size problem, it is also possible to choose SDP as the subproblem in branch and bound methods. But how to choose a suitable N to make use the effectiveness of SDP sufficiently and on the same time do not increase the computational task is not an easy problem.

Parrilo [19] and Qi and Teo [22] listed some interesting open questions on POP.

ACKNOWLEDGMENT

The author would like to thank Prof. Jie Sun for his constant help.

REFERENCES

[1] E. Balas, S. Ceria and G. Gornué jols, *A Lift-and-Project Cutting Plane Algorithm for Mixed* $0-1$ *Programs*, Mathematical Programming **58** (1993), 295-323.

[2] D.A. Cox, J.B. Little and D. O'Shea, *Ideals, Varieties and Algorithms*, Undergraduate Texts in Mahtematics, Springer-Verlag, 1997.

[3] C. Berg, *The Multidimensional Moment Problem and Semigroups*, in Moments in Mathematics, H.J. Landau, ed., AMS, Providence, RI, 1980, 110-124.

[4] D.A. Cox, J.B. Little and D. O'Shea, *Using Algebraic Geometry*, Graduate Textx in Mathematics, vol. 185, Springer-Verlag, 1998.

[5] R.E. Curto, *Flat Extensions of Positive Moments Matrices: Recursively Generated Relations*, Memory of American Mathematical Society **136** (1998), no. 648.

[6] I.M. Gal'fand, M. Kapranov and A. Zelevinsky, *Disciminants, Resultants, and Multidimensional Determinants*, Berhäuser, 1994.

[7] B. Hanzon and D. Jibetean, *Global Minimization of a Multivariate Polynomial Using Matrix Methods*, Journal of Global Optimization **27** (2003), 1-23.

[8] S. Kim and M. Kojima, *Secon Order Cone Programming Relaxation of Nonconvex Quadratic Optimization Problems*, Optimization Methods and Software **15** (2001), 201-224.

[9] s. Kim and M. Kojima, *Exact Solutions of some Nonconvex Quadratic Optimization Problems via SDP and SOCP Relaxations*, Computational Optimization and Applications, to appear.

[10] M. Kojima, S. Kim and H. Waki, *A General Framwork for Convex Relaxation of Polynomial Optimization Problems over Cones*, Journal of Operations Research Society of Japan **46** (2003), 125-144.

[11] M. Laurent, *A Comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre Relaxations for* $0-1$ *Programming*, Report PNA-R0108, CWI, Amsterdam, The Netherlands, 2001.

[12] J.B. Lasserre, *Global Optimization with Polynomials and the Problems of Moments*, SIAM Journal on Optimization **11** (2001), 796-817.

[13] J.B. Lasserre, *An Expicit Equivalent Convex Positive Semidefinite Program for nonlinear 0-1 Programs*, SIAM Journal on Optimization **12** (2002), 756-769.

[14] J.B. Lasserre, *Semidefinite Programming VS. LP Relaxations for Polynomial Programming*, Mathematics of Operations Research **27** (2002), 347-360.

[15] L. Lovász and A. Schrijver, *Cones of Matrices and Set Functions and* $0-1$ *Optimization*, SIAM Journal on Optimization **1** (1991), 166-190.

[16] T.Y. Li, *numerical solution of multivariate polynomial systems by homotopy continuation methods*, Acta Numerica **6** (1997), 399-436.

[17] H.M. Möller and H.J. Stetter, *Multivariate Polynomial Equations with Multiple Zeros Solved by Matrix Eigenproblems*, Numerische Mathematike **70** (1995), 311-329.

[18] Y.E. Nestorov and A. Nemirovski, *Interior Point Polynomial Methods in Convex Programming*, Studies in Applied Mathematics, vol. 13, SIAM, Philadelphia, PA, 1994.

[19] P.A. Parrilo and B. Sturmfels, *Minimizing Polynomial Functions*, DIMACS Series in Discrete Mathmatics and Theoretical Computer Science.

[20] P.A. Parrilo, *Semidefinite Programming Relaxation for Semialgebraic Problems*, Mathematical Programming **96** (2003), 293-320.

[21] M. Putinar, *Positive Polynomials on Compact Semialgebraic Sets*, Indiana University Journal of Mathematics **42** (1993), 969-984.

[22] L.Q. Qi and K.L. Teo, *Multivariate Polynomial Minimization and Its Application in Signal Processing*, Journal of Global Optimization **26** (2003), 419-433.

[23] B. Reznick, *Some Concrete Aspects of Hilbert's 17th Problem*, Contemporary Mathematics **253** (2000), 251-272.

[24] H.D. Sherali and W.P. Adams, *A Hierarchy of Relaxations Between the Continuous and Convex Hull representations for Zero-One Programming Problems*, SIAM Journal on Discrete Mathematics **3** (1990), 411-430.

[25] H.D. Sherali and W.P. Adams, *A Hierarchy of Relaxations and Convex Hull Characerizations for Zero-One Programming Problems*, Discrete Applied Mathematics **52** (1994), 83-106.

[26] H.D. Sherali and C.H. Tuncbilek, *A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique*, Journal of Global Optimization **2** (1992), 101-112.

[27] H.D. Sherali and C.H. Tuncbilek, *A Reformulation-Convexification Approach for Solving nonconvex Quadratic Programming Problems*, Journal of Global Optimization **7** (1995), 1-31.

[28] H.D. Sherali and C.H. Tuncbilek, *New Reformulation-Linearization/Convexification Relaxation for univariate and Multivariate Polynomial Programming Problems*, Operations Research Letter **21** (1997), 1-9.

[29] N.Z. Shor, *Class of Global Minimum Bounds of Polynomial Functions*, Cybernetics **23** (1987), 731-734.

[30] L. Vandenberghe and S. Boyd, *Semidefinite Programming*, SIAM Review **38** (1996), 49-95.

[31] J. Verschelde, *Polynomial homotopies for dense, sparse and determinantal systems*, MSRI Berkeley Preprint #1999-041.