Combining SOS and Moment Relaxations with Branch and Bound to Extract Solutions to Global Polynomial Optimization Problems

Hesameddin Mohammadi, Matthew M. Peet

***Abstract*— We consider the problem of extracting approximate solutions to global polynomial optimization problems using a combination of branch and bound and Sum-of-Squares/Moment Relaxations. Specifically, Sum of Squares/Moment relaxations are used to find greatest lower bounds on the polynomial optimization problem. These bounds are used to select branches which allow us to shrink the feasible domain to within an error tolerance determined by the degree the of SOS/Moment relaxation (and hence the error of the lower bound). For a fixed degree, the complexity of the algorithm is linear in the number of branches and polynomial in the number of constraints. We illustrate the algorithm for a 6-variable, 7-constraint polynomial optimization problem which is not zero-dimensional - a case where existing Moment-based approaches for extracting the global minimizer fail.**

I. Introduction

Consider problems such as economic dispatch [1], optimal power flow [2] and optimal decentralized control [3]. Each of these problems can be formulated in the framework of *Global Polynomial Optimization* (GPO), which has the following form:

$$\begin{aligned} f^* := \min_{\mathbf{x}\in\mathbb{R}^n}. \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \geq 0 \quad \text{for } i = 1,\cdots,s \\ & h_j(\mathbf{x}) = 0 \quad \text{for } j = 1,\cdots,t, \end{aligned} \tag{1}$$

where f, g_i, and h_i are real-valued polynomials in decision variables $\mathbf{x}$. This class of optimization encompasses many types of problems, such as linear programming, quadratic programming, integer programming and mixed integer non-linear programs (MINLP) [4]. For example, if we choose the constraint $g_1(x_1) = x_1(x_1-1) = 0$, then this is equivalent to the integer constraint $x_1 \in \{0,1\}$. In addition to classical optimization, GPO can be used to analyze dynamical systems by, e.g., verifying polytopic invariants of dynamical systems defined by polynomial vector fields as in [5].

Although GPO is an important class of problems, decision problems associated with GPO are known to be NP-hard [6]. This is in contrast to other well-known classes of optimization such as linear programming, for which there exist efficient polynomial time algorithms [7], [8]. The primary reason that GPO is hard is that objective and feasible sets are not convex, implying that common approaches such as cutting plane and descent methods may not yield optimal solutions. Of course, in certain special cases, efficient algorithms exist, such as, e.g. for univariate polynomials [9], [10], and for quadratic GPO [11]. Additionally, for the unconstrained case, there exist some algebraic approaches such as the use of Groebner bases which parameterize critical points of the objective [12].

Recently, new approaches to GPO have been proposed which do not solve the optimization problem directly, but rather search for the Greatest Lower Bound (GLB), $\lambda^* \leq f^*$ which can be expressed in the following framework:

$$\begin{aligned} \lambda^* := \max_{\lambda\in\mathbb{R}}. \quad & \lambda \\ \text{subject to} \quad & f(\mathbf{x}) - \lambda > 0 \quad ,\forall \mathbf{x} \in S, \end{aligned} \tag{2}$$

where $S := \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geq 0,\ h_j(\mathbf{x}) = 0\}$ is the feasible set of Problem (1). Although there are many approaches to solviong the GLB problem, in this paper we focus on the following two algorithms: the Sum of Squares (SOS) approach [13], and its dual method (sometimes referred to as primal), the Moment Approach [14]. Both these approaches are well-studied and have efficient implementations as Matlab toolboxes such as SOSTOOLS [15] and Gloptipoly [16]. In both cases, these methods yield a hierarchy of semidefinite programs with an associated sequence of optimal values $\{p_k^*\}_{k\in\mathbb{N}}$ (SOS) and $\{d_k^*\}_{k\in\mathbb{N}}$ (Moment). Both sequences are increasing and under mild conditions each satisfy $p_k^* \leq d_k^* \leq f^*$ [17]. Furthermore, under additional conditions, it can be shown that $\lim_{k\to\infty} p_k^* = \lim_{k\to\infty} d_k^* = f^*$ [17]. Because calculation of these lower bounds is expressed as semidefinite programming, for a given k, p_k^* and d_k^* can be computed in polynomial time [18]. Although in this paper we focus on the SOS/Moment approaches, a discussion of alternative polynomial-time approaches for solving Problem (2) can be found in our recent survey in [19].

A natural question when applying SOS/Moment methods is under what conditions there exists a k such that $d_k^* = f^*$ or $d_k^* = f^*$ (which we will refer to as finite convergence) and whether solutions to the GLB problem can be used to find feasible or optimal solutions to GPO (extracting the minimizer). For the moment problem, these two questions are strongly related in that if the ideal $< h_1,\ldots,h_t >$ is radical and zero dimensional (typical for integer programming [20]), it has been shown that the moment relaxation obtains finite convergence and extraction of the minimizer is possible. Unfortunately, however, for other classes of problems, finite convergence is not guaranteed (although it does not seem to be uncommon [21]).

In the absence of finite convergence, the primary focus of the GLB problem is typically to reduce the error, defined as the difference between the GLB and f^*. Indeed, results in this area have established bounds on the error of SOS relaxation which scales as $\|p_k^* - f^*\| \cong \frac{1}{\sqrt[c]{\log(k)}}$ for some constant c which is a function of the polynomials g_i and h_j [22].

Unfortunately, however, one of the challenges of using the SOS/Moment methods is that the size of the resulting SDP scales poorly in k. For example, the $k'th$ moment problem results in an SDP with a number of variables which scales as $\binom{n+k}{n}$ [23], where n is the number of decision variables. Finally, note that this scaling is significant even in the case of finite convergence since the k for which $d_k^* = f^*$ may be large.

The goal of this paper is to combine SOS and Moment GLB problems with a Branch and Bound methodology to create a sequence of algorithms which, for any desired accuracy, η, yields a polynomial-time algorithm which returns a point $\hat{\mathbf{x}}$ which is guaranteed to be η-feasible and η-suboptimal for the GPO problem. By η-feasible we mean $g_i(\hat{\mathbf{x}}) \geq -\eta$ and $\|h_i(\hat{\mathbf{x}})\| \leq \eta$ and by η-suboptimal we mean $f^* - f(\hat{\mathbf{x}}) \leq \eta$. Our approach is to enclose the feasible set S within a hyper-rectangle defined by $C_l := \{\mathbf{x} : (x_i - \underline{c}_{i,l})(\bar{c}_{i,l} - x_i) \geq 0, i = 1, \cdots, n\}$. Then, at each iteration, we branch the hyper-rectangle by its longest edge $(\bar{c}_i - \underline{c}_i)$ and intersect each branch with S to create two new GLB problems. Then, for a fixed k we compute p_k^* from the SOS/Moment problems to obtain a GLB λ_l to f over $S = S \cap C_l$. We then choose the branch with the least GLB and further branch and GLB this hypercube to obtain λ_{l+1}. Through successive iterations, the feasible set is reduced in volume as $\frac{1}{2^l}$. Then, because for a fixed k we know that the GLB is accurate with error bound $\|p_k^* - f^*\| \cong \frac{1}{\sqrt[c]{\log(k)}} = \eta_k$, we are guaranteed that each iteration will yield a branch which contains a point which is η-suboptimal. Furthermore, because neither the degree nor the number of constraints changes at each iteration, each iteration is of fixed polynomial time. Finally, the number of iterations necessary to achieve η-feasibility is logarithmic in η. To illustrate the effectiveness of the proposed algorithm, we conducted numerical tests on an example problem wherein Gloptipoly fails to achieve finite convergence.

II. Notation and Preliminaries

$\mathbb{N}$ is the natural numbers and $\mathbb{N}^n$ is the set of n-tuples of $\mathbb{N}$. For $a, b \in \mathbb{N}^n$, we define lexicographical ordering inductively as $a \leq b$ if $a_1 < b_1$ or $a_1 = b_1$ and $a' \leq b'$ where $a' = [a_2, \cdots, a_n] \in \mathbb{N}^{n-1}$, $b' = [b_2, \cdots, b_n] \in \mathbb{N}^{n-1}$. For $n = 1$, the natural numbers provide the ordering. Then let A^n be the elements of $\mathbb{N}^n$ arranged in lexicographical order so that $A^n(i) < A^n(i+j)$ for all $i, j \in \mathbb{N}$.

Now let $\mathbb{R}[\mathbf{x}]$ denote the ring of multivariate polynomials with real coefficients in real variables $\mathbf{x} := (x_1, x_2, \ldots, x_n)$. For variables $\mathbf{x} \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$, we denote monomials using the multi-index notation $\mathbf{x}^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$. Now for variables $\mathbf{x} \in \mathbb{R}^n$, let $Z(\mathbf{x})$ denote the infinite vector of all monomials arranged in lexicographical order so that $Z(\mathbf{x})_i = \mathbf{x}^{A^n(i)}$. Then for any $k \in \mathbb{N}$, let $Z_k(\mathbf{x})$ be the vector of monomials of degree k or less which are defined by the first $\Lambda(k) := \binom{k+n}{k}$ elements of $Z(\mathbf{x})$. We denote the set of sums of squares polynomials, (SOS) by $\Sigma_{\mathbf{S}} := \{s \in \mathbb{R}[\mathbf{x}] : s(x) = \sum_{i=1}^l (p_i(x))^2, p_i \in \mathbb{R}[\mathbf{x}], l \in \mathbb{N}\}$. Let $H_{(k)}[\mathbf{x}] := \{h \in \mathbb{R}[\mathbf{x}] : \deg(h) \leq k\}$, be the polynomials of degree k or less where $\deg(h)$ denotes the degree of h.

Define the set of infinite sequences $\ell_\infty := \{z : \mathbb{N} \to \mathbb{R} : \max_i |z_i| < \infty\}$ with the standard norm. For any polynomial f, we denote by $\overline{f} \in \ell_\infty$ to be the ordered vector of coefficients of f s.t. $f(x) = \overline{f}^T Z(\mathbf{x})$.

Define the ball of radius k as $\mathbb{N}^n_{(k)} := \{b \in \mathbb{N}^n : |b|_{\ell_1} \leq k\}$, where $|b|_{\ell_1} := \sum_{i=1}^n |b_i|$. We denote by $\mathbb{S}^n$ and $\mathbb{S}^{n+}$, the set of all $n \times n$ symmetric matrices and cone of positive semidefinite matrices, respectively.

Quadratic Modules and the Archimedean Property Given polynomials $g_i \in \mathbb{R}[\mathbf{x}]$, $i = 0, \ldots, s$, where $g_0(\mathbf{x}) = 1$, suppose the semialgebraic set

$$S := \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, s\} \tag{3}$$

is a nonempty and compact subset of $\mathbb{R}^n$.

Definition 1: Given polynomials $g_i \in \mathbb{R}[\mathbf{x}]$, $i = 0, \ldots, s$, where $g_0(\mathbf{x}) = 1$, for all $k \in \mathbb{N}$, we define $M^{(k)}$ as the following degree-k bounded quadratic module:

$$M^{(k)} := \{p | p = \sum_{i=0}^{s} \sigma_i g_i \quad \sigma_i \in \Sigma_{\mathbf{S}} \quad \deg(\sigma_i g_i) \leq k.\}$$

Clearly, Definition (1) implies $M^{(1)} \subset M^{(2)} \subset \ldots$.

We define M to be the quadratic module generated by polynomials $(g_0, g_1, \ldots, g_s)$ as:

$$M := \bigcup_{i=1}^{\infty} M^{(i)}. \tag{4}$$

Note that although $M^{(K)} \subset M \cap H_{(k)}$, the reverse does not necessarily hold. Moreover, it has been shown that if $n, k \geq 5$ and $s = 0$, then $M \cap H_{(k)} \not\subseteq M^{(K)}$ [24]. A quadratic module M is called archimedean if $\exists p \in M$ s.t. $p(\mathbf{x}) = R^2 - \sum_{i=1}^n x_i^2$. Note that the archimedean property is a property of the quadratic module and not an inherent property of corresponding semialgebraic set, S. Specifically, if S is compact, then there exists $R > 0$ such that

$$R^2 - \sum_{i=1}^{n} x_i^2 \geq 0, \quad \forall \mathbf{x} \in S.$$

Thus, if M is not archimedean, one can add a redundant inequality constraint $g(\mathbf{x}) := R^2 - \sum_{i=1}^n x_i^2 \geq 0$ to the definition of S, in order to form an archimedean quadratic module corresponding to S.

III. Problem Statement

In this paper, we consider Global Polynomial Optimization (GPO) problems of the form:

$$\begin{aligned} f^* := \min_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \geq 0 \quad \text{for } i = 0, \cdots, s \\ & h_j(\mathbf{x}) = 0 \quad \text{for } j = 1, \cdots, t, \end{aligned} \tag{5}$$

where $h_j(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$. Note that the question of infeasibility of Problem (5) ($f^* = \infty$) is a separate question which can be addressed using the SOS methodology coupled with Positivstellensatz results [25]. However, for the purpose of this paper, we assume that Problem (5) is feasible. Our goal, then, is for any given $\varepsilon > 0$, to determine a point $\mathbf{x}_\varepsilon \in \mathbb{R}^n$ with the following properties:

- $\|f(\mathbf{x}_\varepsilon) - f^*\| \le \varepsilon$
- $\|h_j(\mathbf{x}_\varepsilon)\| \le \varepsilon$
- $g_i(\mathbf{x}_\varepsilon) \ge -\varepsilon$.

In the following, we will simplify the presentation through elimination of equality constraints of the form $h(\mathbf{x}) = 0$. Each equality constraint will be replaced by two inequality constraints of the form $g_1(\mathbf{x}) := h(\mathbf{x}) \ge 0$ and $g_2(\mathbf{x}) := -h(\mathbf{x}) \ge 0$. This does not change the problem statement, however, as $h(\mathbf{x}_\varepsilon) \ge -\varepsilon$ and $-h(\mathbf{x}_\varepsilon) \ge -\varepsilon$ implies $\|h(\mathbf{x}_\varepsilon)\| \le \varepsilon$.

IV. Background on Semidefinite representations

In this section, we define the two methods we use to solve the GLB problem - namely the SOS approach and Moment relaxations (Since the fundamentals of these methods are known, if the reader does not wish to examine the details of implementation, they may wish to proceed to the following section). Both these methods rely on Positivstellensatz results to parameterize the set of polynomials which are positive over a given semialgebraic set. First, we define S to be the semialgebraic set as in (3). We define M to be the corresponding quadratic module, as in (4), which is assumed to be archimedean. The relevant Positivstellensatz result is then stated as follows [22].

Theorem 1: (Putinar's Positivestellensatz) Suppose the quadratic module M is archimedean, then for any $f \in \mathbb{R}[\mathbf{x}]$,

$$\underset{\forall \mathbf{x} \in S}{f(\mathbf{x}) > 0} \quad \Rightarrow \quad f \in M.$$

Now define the following GLB problems: (6), (7) and (8):

$$\begin{aligned} \lambda_6^* := \max_{\lambda \in \mathbb{R}}.\quad & \lambda \\ \text{subject to} \quad & f(\mathbf{x}) - \lambda > 0 \quad, \forall \mathbf{x} \in S. \end{aligned} \tag{6}$$

$$\begin{aligned} \lambda_7^* := \max_{\lambda \in \mathbb{R}}.\quad & \lambda \\ \text{subject to} \quad & f(\mathbf{x}) - \lambda \ge 0 \quad, \forall \mathbf{x} \in S, \end{aligned} \tag{7}$$

$$\begin{aligned} \lambda_8^* := \max_{\lambda \in \mathbb{R}}.\quad & \lambda \\ \text{subject to} \quad & f(\mathbf{x}) - \lambda \in M. \end{aligned} \tag{8}$$

Lemma 1: If M is archimedean, then $f^* = \lambda_6^* = \lambda_7^* = \lambda_8^*$.

Proof: By definition $f^* = \lambda_7^*$. Clearly $\lambda_7^* \ge \lambda_6^*$. $\lambda_7^* \ge \lambda_6^*$ can be established by contradiction, which implies $\lambda_7^* = \lambda_6^*$. Now Theorem (1) implies $\lambda_8^* \ge \lambda_6^*$ and non-negativity of elements of M over S implies $\lambda_7^* \ge \lambda_8^*$. We conclude that $f^* = \lambda_6^* = \lambda_8^* = \lambda_7^*$. ■

Lemma 1 implies that the optimal value of Problem (5) can be obtained by solving the GLB problems in either Problem (7) or Problem (8).

A. SOS Tightenings

In this subsection, we derive a hierarchy of SOS GLB problems and define the corresponding SDPs which can be used to solve these GLB problem. The SOS approach is predicated on a correspondence between SOS polynomials of degree less than or equal to $2d$ and positive semidefinite matrices of size $\Lambda(d)$, such that for any matrix $\Omega \in \mathbb{S}^{\Lambda(d)+}$ there is a SOS polynomial σ, $\deg(\sigma) \le 2d$, where $\sigma := Z_d(\mathbf{x})^T \Omega Z_d(\mathbf{x})$. Likewise, for any $\sigma \in \Sigma_{\mathbf{S}} \cap H_{(2d)}[\mathbf{x}]$, there is at least one $\Omega \in \mathbb{S}^{\Lambda(d)+}$, where $\sigma = Z_d(\mathbf{x})^T \Omega Z_d(\mathbf{x})$. This correspondence can be used to solve a hierarchy of GLB problems $P(k)$, $\forall k \in \mathbb{N}$, whose optimal values provide lower bounds to the optimal value f^* of Problem (8) [13]. Specifically, let p_k^* be the optimal value of Problem $P(k)$ defined as

$$\begin{aligned} p_k^* := \quad \max_{\lambda \in \mathbb{R}}.\quad & \lambda \\ \text{subject to} \quad & f(\mathbf{x}) - \lambda \in M^{(k)}. \end{aligned} \tag{9}$$

Since we have $M^{(1)} \subset M^{(2)} \subset \cdots \subset \bigcup_{i=1}^{\infty} M^{(i)} = M$, one can say: $p_1^* \le p_2^* \le \cdots \le p^*$, where $p^* := \lambda_8^*$ is the optimal value of Problem (8). Furthermore, it has been shown that $\lim_{k \to \infty} p_k^* = p^*$ and indeed, bounds on the convergence rate of these approximations exist as a function of M, f and k (See [17], [22]).

To solve the sequence of problems $P(k)$, we observe that any element of $M^{(k)}$ can be represented using $s+1$ semidefinite matrices of size at most $\Lambda(\lceil \frac{k}{2} \rceil)$. Specifically, let $t_i := \deg(g_i)$, $d_i := \lfloor (k - t_i)/2 \rfloor$, and $g_i(\mathbf{x}) = \sum_{\gamma_i \in \mathbb{N}^n_{(t_i)}} g_{\gamma_i} \mathbf{x}^{\gamma_i}$ for $i = 0, \ldots, s$. Then, one can parameterize the set $M^{(k)}$ as

$$\begin{aligned} M^{(k)} = \Big\{ P | P(\mathbf{x}) &= \sum_{i=0}^{s} Z_{d_i}(\mathbf{x})^T \Omega_i Z_{d_i}(\mathbf{x}) g_i(\mathbf{x}) \\ &= \sum_{i=0}^{s} \sum_{\alpha_i, \beta_i \in \mathbb{N}^n_{(d_i)}} (\Omega_{i(\alpha_i,\beta_i)} \mathbf{x}^{\alpha_i + \beta_i} \sum_{\gamma_i \in \mathbb{N}^n_{(t_i)}} g_{\gamma_i} \mathbf{x}^{\gamma_i}) \\ &= \sum_{i=0}^{s} \sum_{\alpha_i, \beta_i \in \mathbb{N}^n_{(d_i)}} \sum_{\gamma_i \in \mathbb{N}^n_{(t_i)}} (\Omega_{i(\alpha_i,\beta_i)} g_{\gamma_i} \mathbf{x}^{\alpha_i + \beta_i + \gamma_i}) \,, \Omega_i \in \mathbb{S}^{\Lambda(d_i)+} \Big\} \end{aligned}$$

Pardoning the complex notation, this representation shows that all the coefficients of elements of $M^{(k)}$ are bilinear functions of the coefficients of the g_i's and the elements of the positive semidefinite matrices Ω_i. Attempting to simplify, we can rewrite this expression as

$$M^{(k)} = \Big\{ P | P(\mathbf{x}) = \sum_{\omega \in \mathbb{N}^n_{(k)}} (\sum_{i=0}^{s} C_\omega(\Omega_i, g_i)) \mathbf{x}^\omega \,, \Omega_i \in \mathbb{S}^{\Lambda(d_i)+} \Big\}$$

where the bilinear functions $C_\omega(\Omega_i, g_i)$ are defined as

$$C_\omega(\Omega_i, g_i) := \sum_{\substack{\alpha_i, \beta_i \in \mathbb{N}^n_{(d_i)}, \gamma_i \in \mathbb{N}^n_{(t_i)} \\ \alpha_i + \beta_i + \gamma_i = \omega}} \Omega_{i(\alpha_i,\beta_i)} g_{\gamma_i}$$

Finally, we obtain a semidefinite program corresponding to each of problems $P(k)$ as follows.

$$p_k^* = \max_{\substack{\lambda\in\mathbb{R},\, \Omega_i\in\mathbb{S}^{\Lambda(d_i)}\\ i=0,\dots,s}} \quad \lambda \tag{10}$$
$$\text{subject to} \quad \sum_{i=0}^{s} C_\omega(\Omega_i, g_i) = \overline{f}_\omega, \quad \forall\omega\in\mathbb{N}^n_{(k)}\backslash\{\overline{0}\}$$
$$\sum_{i=0}^{s} C_{\overline{0}}(\Omega_i, g_i) = \overline{f}_{\overline{0}} - \lambda$$
$$\Omega_i \succeq 0 \qquad\qquad i=0,\dots,s,$$

where $f(\mathbf{x}) = \sum_{\omega\in\mathbb{N}^n_{(k)}} \overline{f}_\omega \mathbf{x}^\omega$ and $\overline{0}$ denotes the zero of $\mathbb{R}^n$.

We conclude that each GLB Problem $P(k)$ can be formulated as an SDP with matrix variables of dimension $\Lambda(k)$. Note that in practice, construction of the corresponding SDP can be automated through the Matlab toolbox SOSTOOLS.

B. Moment relaxations

In this subsection, we derive the SDP formulation of the moment relaxations and the corresponding hierarchy of GLB problems $D(k)$. To properly explain the moment approach, however, we first review a few basics of Measure Theory which may not be familiar to the reader. First, let $\mathbb{K}$, denote the set of Borel subsets of $\mathbb{R}^n$. A Borel subset of $\mathbb{R}^n$ is a set which is formed from a set by taking all countable combinations of three operations: a) union; b) intersection; and c) relative complement, on open subsets of $\mathbb{R}^n$. We denote the set of finite and signed Borel measures on $\mathbb{K}$ by $\mathscr{M}(\mathbb{K})$. A measure on $\mathbb{K}$ is any function $\mu:\mathbb{K}\to\mathbb{R}^+\cup\{+\infty,0\}$ which satisfies the following properties:

- $\mu(\emptyset)=0$
- $\mu(\bigcup_{i=1}^\infty B_i) = \sum_{i=1}^\infty \mu(B_i)$ for all sequences $\{B_i\}$ of pairwise disjoint elements of $\mathbb{K}$.

We say that $\mu\in\mathscr{M}(\mathbb{K})$ has support on S, if $\mathbb{R}^n\backslash S$ is the largest open set with $\mu(\mathbb{R}^n\backslash S)=0$. Measure μ is called a *probability measure* if $\mu(\mathbb{R}^n)=1$. In addition, for $\mathbf{x_0}\in\mathbb{R}^n$, we can define the delta measure as

$$\delta_{\mathbf{x_0}}(B) := \begin{cases} 1 & \text{if } \mathbf{x_0}\in B \\ 0 & \text{if } \mathbf{x_0}\notin B \end{cases}.$$

Now, for any given $B\subset\mathbb{R}^n$, define the indicator function of B as

$$\mathbf{1}_B(\mathbf{x}) =: \begin{cases} 1 & \text{if } \mathbf{x}\in B \\ 0 & \text{if } \mathbf{x}\notin B \end{cases}$$

A function $g:\mathbb{R}^n\to\mathbb{R}$ is called simple, if for some finite set of indices $I\subset\mathbb{N}$, $a_i\in\mathbb{R}$ and hypercubes $A_i\subset\mathbb{R}^n$, we have $g:=\sum_{i\in I} a_i \mathbf{1}_{A_i}$. Let g be a simple function and μ a measure on $\mathbb{K}$; then we define $\int_{\mathbb{R}^n} g\, d\mu := \sum_{i\in I} a_i\mu(A_i)$. It can be shown that $\int_{\mathbb{R}^n} g\, d\mu$ is an inherent property of g and does not depend on the particular representation of g by a_i's and A_i's.

We say that $f:\mathbb{R}^n\to\mathbb{R}$ is measurable if for any arbitrary Borel subset of $\mathbb{R}$, U, $f(U)^{-1}\subset\mathbb{K}$, where $f(U)^{-1} := \{\mathbf{x}\in\mathbb{R}^n | f(\mathbf{x})\in U\}$. Then for any measurable function $f:\mathbb{R}^n\to[0,\infty)$ and a given measure μ of $\mathbb{R}^n$, one can define $\int_{\mathbb{R}^n} f\, d\mu$ (the integral of f with respect to μ) as:

$$\int_{\mathbb{R}^n} f\, d\mu := \sup\{\int_{\mathbb{R}^n} g\, d\mu | 0\le g\le f\,,\ g \text{ is simple}\}$$

For any measurable function $g:\mathbb{R}^n\to\mathbb{R}$, we define $g^+(\mathbf{x}) := \begin{cases} g(\mathbf{x}) & \text{if } g(\mathbf{x})\ge 0 \\ 0 & \text{if } g(\mathbf{x})\le 0 \end{cases}$ and $g^- := g^+ - g$, where obviously $g=g^+-g^-$ and $g^+\ge 0$, $g^-\ge 0$. Then, we define

$$\int_{\mathbb{R}^n} g\, d\mu := \int_{\mathbb{R}^n} g^+\, d\mu - \int_{\mathbb{R}^n} g^-\, d\mu\,.$$

Function g is called integrable iff both of the above integrals exist and are bounded. Then for any given $S\subset\mathbb{K}$, the integral of g over S will be defined as:

$$\int_S g\, d\mu := \int_{\mathbb{R}^n} g\, \mathbf{1}_s\, d\mu\,.$$

Function g is called integrable over S, if the above integral exists, where g does not have to be integrable in general. It is not also difficult to show that these definitions are consistent with our natural expectations of integrals, meaning that it satisfies the following properties:

For all integrable functions f,g, $c\in\mathbb{R}$, $S\in\mathbb{K}$ and $\mu\in\mathscr{M}(\mathbb{K})$

a) $\int_s c(f+g)\, d\mu = c\int_s f\, d\mu + c\int_s g\, d\mu$
b) $\int_s 1\, d\mu = \mu(S)$
c) $f\le g \implies \int_s f\, d\mu \le \int_s g\, d\mu$
d) $\int_s f\, d\delta_{\mathbf{x_0}} = \begin{cases} f(\mathbf{x_0}) & \text{if } \mathbf{x_0}\in S \\ 0 & \text{if } \mathbf{x_0}\notin S \end{cases} \quad \forall\mathbf{x_0}\in\mathbb{R}^n$

The key connection between our main problem and measures is expressed in the following lemma [14]:

Lemma 2: Given $S\in\mathbb{K}$ and $f:\mathbb{R}^n\to\mathbb{R}$ integrable over S:

$$f(\mathbf{x})\ge 0, \quad \forall\mathbf{x}\in S \iff \int_S f\, d\mu \ge 0, \quad \forall\mu\in\mathscr{M}(\mathbb{K})$$

Proof: $\Rightarrow$ is obvious. $\Leftarrow$ is proven using delta measures at the maxima of f on S. Note in addition, that the lemma holds when we take the restriction to probability measures with support on S. ■

Definition 2: Let S be the semialgebraic set as in (3), and a probability measure $\mu\in\mathscr{M}(\mathbb{K})$ that has support on S be given. we define the moment vector of μ over S, in the following way:

$$(\psi(\mu)_\alpha) := (\int_S \mathbf{x}^\alpha\, d\mu)\,,\ \alpha\in\mathbb{N}^n\,, \text{with lexicographical order.}$$

Immediately, it follows that for any $f\in\mathbb{R}[\mathbf{x}]$:

$$\int_S f\, d\mu = \sum_{\alpha\in\mathbb{N}^n} \overline{f}_\alpha\, \psi(\mu)_\alpha = \overline{f}\cdot\psi(\mu).$$

Theorem 2: (Riesz-Haviland) Let $\mathbf{y}=(y_\alpha)\in\ell_\infty$, $\alpha\in\mathbb{N}^n$ and $S\subset\mathbb{R}^n$ be closed.

$$\exists\mu\in\mathscr{M}(\mathbb{K}) \text{ s.t } y=\psi(\mu) \iff$$

$$\mathbf{y}\cdot\overline{f}\ge 0,\ \forall f\in\mathbb{R}[\mathbf{x}],\ f \text{ nonnegative on } S$$

Using Lemma (2), one can easily reformulate Problem (7), in the following way:

$$\max_{\lambda\in\mathbb{R}}. \quad \lambda \tag{11}$$
$$\text{subject to} \quad \int_S (f-\lambda)\, d\mu \geq 0 \quad ,\forall \mu \in \mathscr{M}(\mathbb{K})$$
$$\mu \text{ is a probabitliy measure and has support on } S$$

Then, by the use of Theorem (2), we rewrite Problem (11):

$$\begin{aligned} \max_{\mathbf{y}\in\ell_\infty,\lambda\in\mathbb{R}}. \quad & \lambda \\ \text{s.t.} \quad & \mathbf{y}\cdot\overline{f-\lambda} \geq 0, \\ & \mathbf{y}\cdot\overline{g} \geq 0 \quad ,\forall\, g\in\mathbb{R}[\mathbf{x}] \text{ nonnegative on } S \\ & \mathbf{y}_{\overline{0}} = 1 \end{aligned} \tag{12}$$

where $\mathbf{y}_{\overline{0}} = 1$ means the corresponding measure of $\mathbf{y}$ has to be a probability measure.

Let $y \in \ell_\infty$ and $\mathbf{y}\cdot\overline{g} \geq 0 \quad ,\forall\, g \in \mathbb{R}[\mathbf{x}]$ positive on S. If $h \in \mathbb{R}[\mathbf{x}]$ be nonnegative on S, $h+\varepsilon$ is positive on $S, \forall\varepsilon \in \mathbb{R}^+$

$$\mathbf{y}\cdot\overline{h+\varepsilon} \geq 0 \implies \mathbf{y}\cdot\overline{h} \geq -\mathbf{y}\cdot\overline{\varepsilon} = \varepsilon\, y_{\overline{0}} \implies \mathbf{y}\cdot\overline{h} \geq 0.$$

Therefore, one can substitute the inequality constraint $\mathbf{y}.\overline{g} \geq 0$ in Problem (12) with strict inequality $\mathbf{y}.\overline{g} > 0$ to take advantage of Theorem (2) and rewrite Problem (12) in the following way:

$$\begin{aligned} \max_{\mathbf{y}\in\ell_\infty,\lambda\in\mathbb{R}}. \quad & \lambda \\ \text{subject to} \quad & \mathbf{y}\cdot\overline{f-\lambda} \geq 0, \\ & \mathbf{y}\cdot\overline{g} \geq 0 \quad ,\forall\, g\in M \\ & \mathbf{y}_{\overline{0}} = 1 \end{aligned} \tag{13}$$

Comparing Problem (13) and (8), the duality between moment and SOS methods is obtained [17]. Verifying $\mathbf{y}.\overline{g} \geq 0$, the second constraint in Problem (13), is difficult because $g \in M$ can be of any degree. In order to overcome this issue, one may define the following hierarchy of moment relaxations, which is analogous to SOS tightenings. Let d_k^* be the optimal value of Problem $D(k)$:

$$\begin{aligned} d_k^* := \max_{\mathbf{y}\in\ell_\infty,\lambda\in\mathbb{R}}. \quad & \lambda \\ \text{subject to} \quad & \mathbf{y}.\overline{f-\lambda} \geq 0, \\ & \mathbf{y}.\overline{g} \geq 0 \quad ,\forall\, g\in M^{(k)} \\ & \mathbf{y}_{\overline{0}} = 1 \end{aligned} \tag{14}$$

Since, $M^{(i)} \subset M^{(i+1)}$, $\forall i \in \mathbb{N}$, one can say that (d_k^*) is an increasing sequence, lower bounding p^*. Additionally, $p_k^* \leq d_k^*$ is implied by the weak duality between $P(k)$ and $D(k)$. Moreover, since $\lim_{k\to\infty} p_k^* = p^*$, we have $\lim_{k\to\infty} d_k^* = p^*$. It has been proved that if the interior of S is not empty, strong duality holds [17], meaning that $p_k^* = d_k^*$. Similar to SOS tightenings, moment relaxations have semidefinite representations. According to the structure of $M^{(k)}$, and linearity of $\mathbf{y}\cdot\overline{g}$ on g, one can write:

$$\begin{aligned} \mathbf{y}.\overline{g} \geq 0,\, \forall g\in M^{(k)} \iff & \mathbf{y}.\overline{g_i s_i^2} \geq 0, \\ & \forall\ s_i \in H_{(d_i)}[\mathbf{x}],\ i=0,\dots,s. \end{aligned} \tag{15}$$

Definition 3: Let $g(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$, $d \in \mathbb{N}$ and $\mathbf{y} \in \ell_\infty$ be given. $\Phi_{\mathbf{y},g,d} \in \mathbb{S}^{\Lambda(d)}$, the localizing matrix of $\mathbf{y}$ with respect to g, can be defined as follows:

$$\Phi_{\mathbf{y},g,d}[\alpha,\beta] := \mathbf{y}\cdot\overline{\mathbf{x}^{\alpha+\beta} g(\mathbf{x})}, \quad \forall\alpha,\beta\in\mathbb{N}^n_{(d)},$$

where elements of $\Phi_{\mathbf{y},g,d}$ are indexed by $\alpha,\beta \in \mathbb{N}^n_{(d)}$ with lexicographical order.

Let $Q_d(\mathbf{x}) := Z_d(\mathbf{x})Z_d(\mathbf{x})^T$, $d\in\mathbb{N}$ be a $\Lambda(d)\times\Lambda(d)$ matrix, formed of the monomials with degree less than $2d$.

Lemma 3: Given $\mathbf{y}\in\ell_\infty$, $g(\mathbf{x})\in\mathbb{R}[\mathbf{x}]$ and $d\in\mathbb{N}$:

$$\mathbf{y}\cdot\overline{g\,s^2} \geq 0,\, \forall\, s\in H_{(d)}[\mathbf{x}] \iff \Phi_{\mathbf{y},g,d} \succeq 0$$

Proof: One can write:

$$\begin{aligned} \mathbf{y}\cdot\overline{g\,s^2} &= \mathbf{y}\cdot\overline{g(\mathbf{x})\overline{s}^T_{(d)} Q_d(\mathbf{x})\overline{s}_{(d)}} \\ &= \mathbf{y}\cdot\overline{\overline{s}^T_{(d)}\left(g(\mathbf{x})\,Q_d(\mathbf{x})\right)\overline{s}_{(d)}} = \mathbf{y}\cdot\overline{\sum_{\alpha,\beta\in\mathbb{N}^n_{(d)}} \overline{s}_\alpha\overline{s}_\beta g(\mathbf{x})\mathbf{x}^{\alpha+\beta}} \\ &= \sum_{\alpha,\beta\in\mathbb{N}^n_{(d)}} \overline{s}_\alpha\overline{s}_\beta \mathbf{y}\cdot\overline{g(\mathbf{x})\mathbf{x}^{\alpha+\beta}} = \overline{s}^T_{(d)}\,\Phi_{\mathbf{y},g,d}\,\overline{s}_{(d)}. \end{aligned}$$

Hence, it can be implied that,

$$\mathbf{y}.\overline{g\,s^2} \geq 0 \iff \overline{s}^T_{(d)}\,\Phi_{\mathbf{y},g,d}\,\overline{s}_{(d)} \geq 0$$

Now, since $\overline{s}_{(d)}$ can be any element of $\mathbb{R}^{\Lambda(d)}$, the lemma is proved. ∎

By the use of Lemma (3) and Equation (15), we can reformulate $D(k)$, Problem (14), and define an equivalent semidefinite program in the following way:

$$\begin{aligned} d_k^* := \max_{\mathbf{y}\in\ell_\infty,\lambda\in\mathbb{R}}. \quad & \lambda \\ \text{subject to} \quad & \mathbf{y}.\overline{f-\lambda} \geq 0, \\ & \Phi_{\mathbf{y},g_i,d_i} \succeq 0, \quad i=0,\dots,s \\ & \mathbf{y}_{\overline{0}} = 1 \end{aligned} \tag{16}$$

It also has been shown that SOS/moment semidefinite programs, Problem (10) and Problem (16) respectively, are dual in the sense of Lagrangian duality [17], [14].

V. Proposed Methodology

In this section, we propose a method which relies on a combination of Branch and Bound, SOS, and Moment Relaxations to find approximate minimizers to Problem(5).

Setup To begin, we choose a degree, k. Next, we replace each equality constraint with a pair of inequality constraints. Then, let S be the semialgebraic set as defined in (3) and let $M^{(k)}$ be the corresponding degree-k bounded quadratic module, as in Definition (1). We assume that M, as in (4), is archimedean, implying that there exists $R\in\mathbb{R}^+$, such that $p(\mathbf{x}) := R^2 - \sum_{i=1}^n x_i^2 \in M$. Finally, we initialize the algorithm with $i=1$, $a = -R\cdot\mathbf{1}$, and $b = R\cdot\mathbf{1}$.

To assist in understanding the algorithm, we define the following subroutine whose inputs are the edges of a hypercube and whose output is a GLB on that hypercube.

Modified GLB Problem Subroutine: `l=MGLB(a,b)` Given a and $b \in \mathbb{R}^n$ and degree bound k, we define $w_i(\mathbf{x}) := (b_i - x_i)(x_i - a_i)$, the semialgebraic set

$$S_{ab} := \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geq 0,\, w_j(\mathbf{x}) \geq 0\} \tag{17}$$

and the corresponding module

$$M_{ab}^{(k)} := \Big\{ p \,:\, p = \sum_{i=0}^{s} \sigma_i g_i + \sum_{i=s+1}^{s+n} \sigma_i w_i, \quad \sigma_i \in \Sigma_{\mathbf{S}}, \quad \deg(\sigma_i g_i) \leq k,\ \deg(\sigma_i w_i) \leq k. \Big\}$$

This allows us to formulate and solve the modified SOS GLB problem

$$p^*_{a,b,k} := \max_{\lambda \in \mathbb{R}}.\ \lambda \quad \text{subject to} \quad f(\mathbf{x}) - \lambda \in M_{ab}^{(k)}. \tag{18}$$

and its dual GLB moment problem $d^*_{a,b,k}$ as described in the proceeding section. Then, set the GLB as $l_{a,b,k} = \max\{d^*_{a,b,k} = p^*_{a,b,k}\}$ and return this value.

Step 1: Branch and Bound Given a and b which define the hypercube $C(a,b) := \{\mathbf{x} \in \mathbb{R}^n : x_i \in [a_i, b_i]\}$, we find $i^* = \arg\max(b_i - a_i)$ and use this to define two new hypercubes as $a_1 = a$, $b_2 = b$,

$$b_{1,i} = \begin{cases} \frac{b_i - a_i}{2} & i = i^* \\ b_i & \text{otherwise,} \end{cases} \qquad a_{2,i} = \begin{cases} \frac{b_i - a_i}{2} & i = i^* \\ a_i & \text{otherwise.} \end{cases}$$

Clearly $C(a_1,b_1)$ and $C(a_2,b_2)$ are disjoint and $C(a_1,b_1) \cup C(a_2,b_2) = C(a,b)$. Now, we call the MGLB subroutine to obtain l_1=`MGLB(`a_1`,`b_1`)` and l_2=`MGLB(`a_2`,`b_2`)` which define greatest lower bounds to f^* over $S \cap C_1$ and $S \cap C_2$ respectively. Finally, we increment $i = i+1$, set $GLB_i = \max l_1, l_2$ and

$$a = \begin{cases} a_1 & l_1 \geq l_2 \\ a_2 & \text{otherwise} \end{cases}, \quad b = \begin{cases} b_1 & l_1 \geq l_2 \\ b_2 & \text{otherwise} \end{cases}$$

Step 2: Estimate Feasible Point Now, we propose a feasible point at the centroid of the current hypercube as $\hat{\mathbf{x}} := \frac{a+b}{2}$. Note that this may be a poor choice of estimate. A better heuristic would be to linearize the constraints on the hypercube and solve the resulting feasibility LP. Given the estimate $\hat{\mathbf{x}} := \frac{a+b}{2}$, we compute the error in both the constraints and the objective with respect to the lower bound. Note that this error may decrease at subsequent iterations.

Step 3: Goto Step 1 Repeat Steps 1 and 2 until error bounds converge. Note that Lemma 4 (which follows) shows that $min\{l_1, l_2\} \geq l_0$=`MGLB(`a`,`b`)`. This implies that the $GLB_i \leq GLB_{i+1}$ (non-decreasing at each iteration (branch).

A. Convergence

We use the following lemmas to show that the sequence of Greatest Lower Bounds in non-decreasing.

Lemma 4: Let $a \leq c < d \leq b \in \mathbb{R}$, $g = (x-a)(b-x)$ and $h = (x-c)(d-x)$. Then there exist α, β and $\gamma \in \mathbb{R}$, such that.

$$g(x) = \alpha h(x) + \beta (x+\gamma)^2, \quad \alpha, \beta \geq 0$$

Proof: Using change of variable $z = x - a$, we can assume without loss of generality that $a = 0$. Now let $p^2 = c$, $q^2 = d - c$, and $r^2 := b - d$. First, we consider the case where $p^2, r^2 \neq 0$. This leads to two sub-cases:
Case 1 : $r^2 \neq p^2$. Let

$$\gamma = \frac{p^4 + p^2q^2 - \sqrt{p^2r^2(p^2+q^2)(q^2+r^2)}}{r^2 - p^2}, \beta = \frac{p^4 + p^2q^2}{\gamma^2 - p^4 - p^2q^2},$$

and $\alpha = \beta + 1$. Verifying the equality $g(x) = \alpha h(x) + \beta(x+\gamma)^2$ is straightforward. To show that $\beta, \alpha \geq 0$, we use the following.

$$\beta \geq 0 \iff \gamma^2 > p^4 + p^2q^2$$

$$\iff \left(p^4 + p^2q^2 - \sqrt{p^2r^2(p^2+q^2)(q^2+r^2)}\right)^2 > (p^4 + p^2q^2)(r^2 - p^2)^2$$

$$\iff \underbrace{(p^4 + p^2q^2)^2 + p^2r^2(p^2+q^2)(q^2+r^2) - (p^4+p^2q^2)(r^2-p^2)^2}_{L} >$$

$$\underbrace{2(p^4+p^2q^2)\sqrt{p^2r^2(p^2+q^2)(q^2+r^2)}}_{U} \iff \begin{cases} L > 0 \\ L^2 > U^2 \end{cases}$$

After simplification we have:

$$L^2 - U^2 = p^4q^4(p^2+q^2)^2(p^2-r^2)^2 > 0, \quad \text{and}$$
$$L = p^2\,(p^2+q^2)\,(p^2\,q^2 + 2\,p^2\,r^2 + q^2\,r^2) > 0$$

which completes the proof for Case 1.
Case 2 : $r^2 = p^2$. In this case, let

$$\gamma = -\frac{2p^2+q^2}{2}, \beta = \frac{4p^2(p^2+q^2)}{q^4}, \alpha = \beta + 1$$

Equality and positivity for this case can then be easily verified. Now, suppose $r^2 = p^2 = 0$. In this case, simply set $\beta = 0, \alpha = 1$. If $p^2 = 0, r^2 \neq 0$, set $\beta = \frac{b}{d} - 1$, $\alpha = \frac{b}{d}$, $\gamma = 0$. The case $p^2 \neq 0$, $r^2 = 0$ is similar to $p^2 = 0$, $r^2 \neq 0$,through the change of variable $z = b - x$. ∎

Now, using Lemma 4, one can prove the following.

Lemma 5: Given $\gamma \leq \alpha \leq \beta \leq \delta$, and $w_{i,1}(\mathbf{x}) = (\beta_i - x_i)(x_i - \alpha_i)$, $w_{i,2}(\mathbf{x}) = (\delta_i - x_i)(x_i - \gamma_i)$,for $i = 1, \cdots, n$, and polynomials $g_0, \cdots, g_s$, let

$$N_1^{(k)} := \{\sum_{i=0}^{s} \sigma_i g_i + \sum_{j=1}^{n} \omega_j w_{j,1}\,,\, \sigma_i,\, \omega_j \in \Sigma_{\mathbf{S}}\,,\, \deg(\sigma_i g_i)\,,\, \deg(\omega_j w_{j,1}) \leq k\}$$
$$N_2^{(k)} := \{\sum_{i=0}^{s} \sigma_i g_i + \sum_{j=1}^{n} \omega_j w_{j,2}\,,\, \sigma_i,\, \omega_j \in \Sigma_{\mathbf{S}}\,,\, \deg(\sigma_i g_i)\,,\, \deg(\omega_j w_{j,2}) \leq k\}.$$

be k-degree bounded quadratic modules. Then $N_2^{(k)} \subset N_1^{(k)}$.

Proof: Using Lemma 4, there exist $p_j, q_j, r_j \in \mathbb{R}$ such that

$$w_{j,2} = p_j^2 \cdot w_{j,1} + q_j^2 \cdot (x_j + r_j)^2\,,\ p_j, q_j, r_j \in \mathbb{R}.$$

Now if $h \in N_2^{(k)}$, there exist $\sigma_i, \omega_{j,2} \in \Sigma_{\mathbf{S}}$ such that $h = \sum_{i=0}^{s} \sigma_i g_i + \sum_{j=1}^{n} \omega_j w_{j,2}$. Then assuming WLOG that $g_0 = 1$, we plug in the expression for $w_{j,2}$ to get

$$h = \sum_{i=0}^{s} \sigma_i \cdot g_i + \sum_{j=1}^{n} \omega_j \cdot (p_j^2 \cdot w_{j,1} + q_j^2 \cdot (x_j + r_j)^2)$$

$$= \underbrace{(\sigma_0 + \sum_{j=1}^{n} q_j^2 \cdot \omega_j \cdot (x_j + r_j)^2)}_{\sigma_{0\ \text{new}}} + \sum_{i=1}^{s} \sigma_i \cdot g_i + \sum_{j=1}^{n} \underbrace{p_j^2\, \omega_j}_{\omega_{j\ \text{new}}} \cdot w_{j,1}$$

Then it can be seen that $\sigma_{0\text{ new}}, \omega_{j\text{ new}} \in \Sigma_{\mathbf{S}}$, $\deg(\sigma_{0\text{ new}} g_0) \leq k$, and $\deg(\omega_{j\text{ new}} \cdot w_{j,1}) \leq k$ which implies that $h \in N_1^{(k)}$ ■
Lemma 5 implies that if the MGLB problem has GLB l for hypercube defined by $a, b \in \mathbb{R}^n$ then for any hypercube contained in this hypercube, it will return a greater or equivalent GLB.

VI. Numerical Results

Consider the following GPO problem.

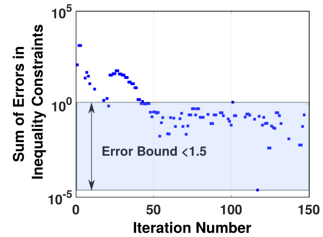
$$\min_{X \in \mathbb{R}^r} \quad f(\mathbf{x}) = 7x_1x_5^3 + 6x_1x_5^2x_6 + 9x_2x_4^3 + 4x_2x_4x_5 + 3x_2x_5x_6 + x_3x_4x_5$$

$$\begin{aligned}
\text{subject to} \quad & g_1(\mathbf{x}) = 10 - (x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2) \geq 0 \\
& g_2(\mathbf{x}) = x_1^3 + x_2^2x_4 + x_3x_5^2 \geq 0 \\
& g_3(\mathbf{x}) = x_2^2x_1 + x_5^3 + x_4x_1x_2 \geq 0 \\
& h_1(\mathbf{x}) = x_1 + x_2^2 - x_3^2 + x_4x_5 = 0 \\
& h_2(\mathbf{x}) = x_5x_1x_2x_3 - x_4^2 = 0
\end{aligned}$$

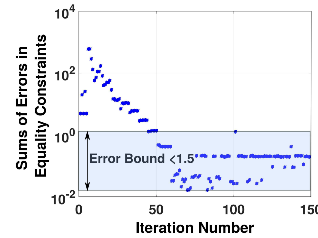
In this example we have 6 variables, an objective function of degree 4 and several equality and inequality constraints of degree 4 or less. The problem is not zero dimensional and the moment approach does not achieve finite convergence and existing methods of extracting solutions fail. At each branch, we use Mosek to solve the SDP related to both the SOS and Moment problems. For this instance, we set the degree to $k = 5$. As seen in Figure 1, the branch and bound algorithm converges relatively quickly to a certain level of error and then saturates. Iterations pas this point to not significantly improve accuracy of the feasible point. As predicted, this saturation and residual error (blue shaded region) is due the use of a fixed degree bound k. As k is decreased, the residual error increases and as k is decreased the residual error decreases. For this problem the final iteration returns the point $\hat{\mathbf{x}} = [-4.4385, -0.0146, 1.1865, 0.6787, 8.6084, 2.0752]$ for which all inequalities are feasible and the equality constraints have errors of .0035 and .2, respectively. The objective value is $f(\hat{\mathbf{x}}) = -23,909.35$.

VII. Conclusion

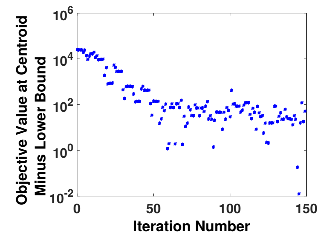
We have proposed a polynomial-time algorithm for extracting solutions to Global Polynomial Optimization problems based on a combination of Branch and Bound and Greatest Lower Bounds obtained from SOS/Moment algorithms. The complexity is exponential in the degree and linear in the number of branches. For a fixed degree, the algorithm returns a solution which is feasible and suboptimal with some tolerance ε. As the degree increases the error tolerance decreases - implying that for any ε, the result is a polynomial-time ε-approximation algorithm for the general problem of GPO (although unfortunately the degree which achieves that ε is not known apriori). For a fixed degree of semidefinite relaxations, our numerical case study demonstrates convergence to a level of residual error which can then be decreased by increasing the degree. In ongoing work, we seek to bound this residual error as a function of degree. In addition, we would like to ultimately reduce errors in feasibility by improving our heuristic selection criterion to ensure feasibility. Since the polynomials become progressively regular as the domain shrinks, one option is linearization of the constraints and solution of the resulting LP. It is unclear, however, whether we can obtain any bounds for the reliability of this approach.



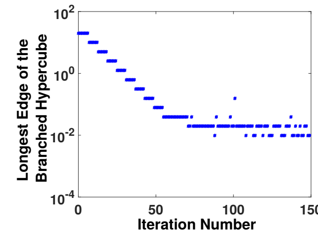
(a) Sum of errors in inequality constraints at center, $\sum_{g_i(\mathbf{x}_m) \leq 0} |g_i(\mathbf{x}_m)|$ vs. number of iteration



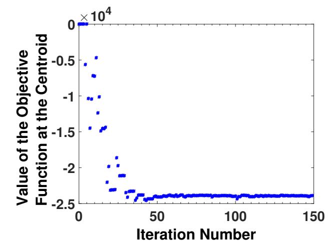
(b) Sum of errors in equality constraints at center, $\sum_j |h_j(\mathbf{x}_m)|$ vs. number of iteration



(c) Gap between lower bound and objective value at center point, $|f(\mathbf{x}_m) - l|$ vs. number of iteration



(d) Longest edge of the branched hypercube vs. number of iteration



(e) Objective function value at center point vs. number of iteration

Fig. 1. Numerical results of the example

References

[1] J. B. Park, Y. W. Jeong, J. R. Shin, and K. Y. Lee, "An improved particle swarm optimization for nonconvex economic dispatch problems," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 156–166, 2010.

[2] B. Ghaddar, J. Marecek, and M. Mevissen, "Optimal power flow as a polynomial optimization problem," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 539–546, 2016.

[3] J. Lavaei, "Optimal decentralized control problem as a rank-constrained optimization," in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*, pp. 39–45, IEEE, 2013.

[4] M. Laurent, *Emerging Applications of Algebraic Geometry*, ch. Sums of Squares, Moment Matrices and Optimization Over Polynomials, pp. 157–270. New York, NY: Springer New York, 2009.

[5] M. A. B. Sassi and A. Girard, "Computation of polytopic invariants for polynomial dynamical systems using linear programming," *Automatica*, vol. 48, no. 12, pp. 3114 – 3121, 2012.

[6] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*. Springer Science & Business Media, 2012.

[7] L. Khachiyan, "Polynomial algorithms in linear programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 20, no. 1, pp. 53 – 72, 1980.

[8] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica*, vol. 4, no. 4, pp. 373–395.

[9] H. D. Sherali and C. H. Tuncbilek, "New reformulation linearization/convexification relaxations for univariate and multivariate polynomial programming problems," *Operations Research Letters*, vol. 21, no. 1, pp. 1 – 9, 1997.

[10] N. Z. Shor, "Quadratic optimization problems," *Soviet Journal of Computer and Systems Sciences*, vol. 25, no. 6, pp. 1–11, 1987.

[11] Z.-Q. Luo and S. Zhang, "A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1716–1736, 2010.

[12] B. Sturmfels, *Solving systems of polynomial equations*. No. 97, American Mathematical Soc., 2002.

[13] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, Citeseer, 2000.

[14] J. B. Lasserre, *Moments, positive polynomials and their applications*, vol. 1. London: World Scientific, 2009.

[15] S. Prajna, A. Papachristodoulou, and P. A. Parrilo, "Introducing sostools: A general purpose sum of squares programming solver," in *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, vol. 1, pp. 741–746, IEEE, 2002.

[16] D. Henrion and J.-B. Lasserre, *Positive Polynomials in Control*, ch. Detecting Global Optimality and Extracting Solutions in GloptiPoly, pp. 293–310. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[17] M. Schweighofer, "Optimization of polynomials on compact semialgebraic sets," *SIAM Journal on Optimization*, vol. 15, no. 3, pp. 805–825, 2005.

[18] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.

[19] R. Kamyar and M. M. Peet, "Polynomial optimization with applications to stability analysis and control - alternatives to sum of squares," *Discrete and Continuous Dynamical Systems - Series B*, vol. 20(8), pp. 2383–2417, 2015.

[20] J. Lasserre, "Polynomials nonnegative on a grid and discrete optimization," *Transactions of the American Mathematical Society*, vol. 354, no. 2, pp. 631–649, 2002.

[21] J. Nie, "Optimality conditions and finite convergence of lasserre's hierarchy," *Mathematical Programming*, vol. 146, no. 1, pp. 97–121, 2013.

[22] J. Nie and M. Schweighofer, "On the complexity of putinar's positivstellensatz," *Journal of Complexity*, vol. 23, no. 1, pp. 135 – 150, 2007.

[23] J. B. Lasserre, "Moments and sums of squares for polynomial optimization and related problems," *Journal of Global Optimization*, vol. 45, no. 1, pp. 39–61, 2009.

[24] B. Reznick, "Some concrete aspects of hilbert's 17th problem," *Contemporary Mathematics*, vol. 253, pp. 251–272, 2000.

[25] G. Stengle, "A nullstellensatz and a positivstellensatz in semialgebraic geometry," *Mathematische Annalen*, vol. 207, no. 2, pp. 87–97, 1974.