Polynomial Optimization Over Unions of Sets

Linghao Zhang

Advisor: Jiawang Nie

Department of Mathematics

University of California San Diego

In partial fulfillment of the requirements

for the Mathematics Honors Program

June 2024

Acknowledgements

I would like to express my gratitude to Professor Jiawang Nie for his invaluable guidance throughout my thesis and for educating me on various mathematical subjects. Each discussion with Professor Nie was significant to me, and his mentorship was essential to the completion of this thesis.

I would also like to extend my gratitude to the UCSD School of Physical Sciences for 2022-2023 Physical Sciences Dean's Undergraduate Award for Excellence, 2023-2024 Selma and Robert Silagi Award for Undergraduate Excellence, and the Undergraduate Summer Research Award.

Finally, I would like to thank my family and friends for their support.

Abstract

This paper studies the polynomial optimization problem whose feasible set is a union of several basic closed semialgebraic sets. We propose a unified hierarchy of Moment-SOS relaxations to solve it globally. Under some assumptions, we prove the asymptotic or finite convergence of the unified hierarchy. Special properties for the univariate case are discussed. The application for computing (p, q)-norms of matrices is also presented.

Contents

Chapter 1

Introduction

1.1 Description of the Problem

We consider the optimization problem

$$\begin{cases} \min & f(x) \\ s.t. & x \in K := \bigcup_{l=1}^{m} K_l, \end{cases} \tag{1.1.1}$$

where each K_l is the basic closed semialgebraic set given as

$$K_l = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} c_i^{(l)}(x) = 0 \; (i \in \mathcal{E}^{(l)}), \\ c_j^{(l)}(x) \geq 0 \; (j \in \mathcal{I}^{(l)}) \end{array} \right\}.$$

Here, all functions f, $c_i^{(l)}$, $c_j^{(l)}$ are polynomials in $x := (x_1, \ldots, x_n)$; all $\mathcal{E}^{(l)}$ and $\mathcal{I}^{(l)}$ are finite labeling sets. We aim at finding the global minimum value f_{min} of (1.1.1) and a global minimizer x^* if it exists. It is worthy to note that solving (1.1.1) is equivalent to solving m standard polynomial optimization problems by minimizing $f(x)$ over each K_l separately, for $l = 1, \ldots, m$. When K is nonempty and compact, f_{min} is achievable at a feasible point, and (1.1.1) has a minimizer. When K is unbounded, a minimizer may or may not exist. We refer to [28, Section 5.1] for the existence of optimizers when the feasible set is unbounded.

The optimization (1.1.1) contains a broad class of problems. For the case $m = 1$, if all functions are linear, then (1.1.1) is a linear program (LP); if f is quadratic and all $c_i^{(l)}, c_j^{(l)}$ are linear, then (1.1.1) is a quadratic program (QP); if all $f, c_i^{(l)}, c_j^{(l)}$ are quadratic, then (1.1.1) is a quadratically constrained quadratic program (QCQP). Polynomial optimization has wide applications, including combinatorial optimization [4, 18], optimal control [8], stochastic and robust optimization [34, 35, 44], generalized Nash equilibrium problems [30, 31, 33], and tensor optimization [5, 26, 32, 36].

When the feasible set K is a single basic closed semialgebraic set (i.e., $m = 1$) instead of a union of several ones, the problem (1.1.1) becomes a standard polynomial optimization problem. There exists much work for solving standard

polynomial optimization problems. A classical approach for solving them glob-
ally is the hierarchy of Moment-SOS relaxations [12]. Under the archimedean-
ness for constraining polynomials, this hierarchy gives a sequence of convergent
lower bounds for the minimum value f_{min}. The Moment-SOS hierarchy has
finite convergence if the linear independence constraint qualification, the strict
complementarity and the second order sufficient conditions hold at every global
minimizer [25]. When the equality constraints define a finite set, this hierarchy
is also tight [14, 16, 24]. We refer to the books and surveys [8, 9, 13, 17, 28] for
introductions to polynomial optimization.

1.2   Contributions

When $m > 1$, the difficulty for solving the optimization problem (1.1.1)
increases. A straightforward approach to solve (1.1.1) is to minimize $f(x)$ over
each K_l separately, for $l = 1, \ldots, m$. By doing this, we reduce the problem
(1.1.1) into m standard polynomial optimization problems.

In this paper, we propose a unified Moment-SOS hierarchy for solving (1.1.1).
The standard kth order moment relaxation for minimizing $f(x)$ over the subset
K_l is (for $l = 1, \ldots, m$)

$$\begin{cases} \min & \langle f, y^{(l)} \rangle \\ s.t. & \mathcal{V}^{(2k)}_{c_i^{(l)}}[y^{(l)}] = 0 \ (i \in \mathcal{E}^{(l)}), \\ & L^{(k)}_{c_j^{(l)}}[y^{(l)}] \succeq 0 \ (j \in \mathcal{I}^{(l)}), \\ & M_k[y^{(l)}] \succeq 0, \\ & y_0^{(l)} = 1, y^{(l)} \in \mathbb{R}^{\mathbb{N}^n_{2k}}. \end{cases} \tag{1.2.1}$$

We refer to Section 2.1 for the above notation. The unified moment relaxation
we propose in this paper is

$$\begin{cases} \min & \langle f, y^{(1)} \rangle + \cdots + \langle f, y^{(m)} \rangle \\ s.t. & \mathcal{V}^{(2k)}_{c_i^{(l)}}[y^{(l)}] = 0 \ (i \in \mathcal{E}^{(l)}), \\ & L^{(k)}_{c_j^{(l)}}[y^{(l)}] \succeq 0 \ (j \in \mathcal{I}^{(l)}), \\ & M_k[y^{(l)}] \succeq 0, \sum_{l=1}^{m} y_0^{(l)} = 1, \\ & y^{(l)} \in \mathbb{R}^{\mathbb{N}^n_{2k}}, l = 1, \ldots, m. \end{cases} \tag{1.2.2}$$

For $k = 1, 2, \ldots$, this gives a unified hierarchy of relaxations.

A major advantage of (1.2.2) is that it gives a unified convex relaxation
for solving (1.1.1) instead of solving it over each K_l separately. It gives a se-
quence of lower bounds for the minimum value f_{min} of (1.1.1). Under the
archimedeanness, we can prove the asymptotic convergence of this unified hier-
archy. Moreover, under some further local optimality conditions, we can prove
its finite convergence. We, in addition, study the special properties for the uni-
variate case. When $n = 1$, there are nice representations for polynomials that

are nonnegative over intervals. The resulting unified Moment-SOS relaxations can be expressed in a more mathematically concise manner. We also present numerical experiments to demonstrate the efficiency of our unified Moment-SOS hierarchy.

An application of (1.1.1) is to compute the (p,q)-norm of a matrix A:

$$\|A\|_{p,q} := \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_q} = \max_{\|x\|_q=1} \|Ax\|_p,$$

where p, q are positive integers. When p and q are both even, this is a standard polynomial optimization problem. If one of them is odd, the norm $\|A\|_{p,q}$ can be expressed as the optimal value of a problem like (1.1.1). For instance, when $p = 4$ and $q = 3$, we can formulate this problem as

$$\begin{cases} \max & (\|Ax\|_4)^4 \\ s.t. & |x_1|^3 + \cdots + |x_n|^3 = 1. \end{cases} \tag{1.2.3}$$

The feasible set of the above can be expressed in the union form as in (1.1.1). It is interesting to note that the number of sets in the union is 2^n, so the difficulty of (1.2.3) increases substantially as n gets larger. More details are given in Chapter 5.

The paper is organized as follows. Chapter 2 introduces the notation and some preliminary results about polynomial optimization. Chapter 3 gives the unified hierarchy of Moment-SOS relaxations; the asymptotic and finite convergence are proved under certain assumptions. Chapter 4 studies some special properties of univariate polynomial optimization. Chapter 5 gives numerical experiments and applications. Chapter 6 draws conclusions and makes some discussions for future work.

Chapter 2

Preliminaries

2.1 Notation

The symbol $\mathbb{N}$ (resp., $\mathbb{R}$) stands for the set of nonnegative integers (resp., real numbers). For an integer $m > 0$, denote $[m] := \{1, 2, \ldots, m\}$. For a scalar $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer greater than or equal to t, and $\lfloor t \rfloor$ denotes the largest integer less than or equal to t. For a polynomial p, $\deg(p)$ denotes its total degree and $\mathrm{vec}(p)$ denotes its coefficient vector. For two vectors a and b, the notation $a \perp b$ means they are perpendicular. The superscript ^{T} denotes the transpose of a matrix or vector. For a symmetric matrix X, $X \succeq 0$ (resp., $X \succ 0$) means that X is positive semidefinite (resp., positive definite). The symbol S_+^n stands for the set of all n-by-n real symmetric positive semidefinite matrices. For two symmetric matrices X and Y, the inequality $X \succeq Y$ (resp., $X \succ Y$) means that $X - Y \succeq 0$ (resp., $X - Y \succ 0$). For $x := (x_1, \ldots, x_n)$ and a power vector $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, denote $|\alpha| := \alpha_1 + \cdots + \alpha_n$ and the monomial $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. For a real number $q \geq 1$, the q-norm of x is denoted as $\|x\|_q := (|x_1|^q + \cdots + |x_n|^q)^{1/q}$. The notation

$$\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\}$$

denotes the set of monomial powers with degrees at most d. The symbol $\mathbb{R}^{\mathbb{N}_d^n}$ denotes the space of all real vectors labeled by $\alpha \in \mathbb{N}_d^n$. The column vector of all monomials in x and of degrees up to d is denoted as

$$[x]_d := \begin{bmatrix} 1 & x_1 & \cdots & x_n & x_1^2 & x_1 x_2 & \cdots & x_n^d \end{bmatrix}^T.$$

The notation $\mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ stands for the ring of polynomials in x with real coefficients. Let $\mathbb{R}[x]_d$ be the set of real polynomials with degrees at most d. Denote by $\mathscr{P}(K)$ the cone of polynomials that are nonnegative on K and let

$$\mathscr{P}_d(K) := \mathscr{P}(K) \cap \mathbb{R}[x]_d.$$

In the following, we review some basics of polynomial optimization. For a

tuple $h := (h_1, \ldots, h_s)$ of polynomials in $\mathbb{R}[x]$, let

$$\mathrm{Ideal}[h] := h_1 \cdot \mathbb{R}[x] + \cdots + h_s \cdot \mathbb{R}[x].$$

The $2k$th *truncation* of $\mathrm{Ideal}[h]$ is

$$\mathrm{Ideal}[h]_{2k} := h_1 \cdot \mathbb{R}[x]_{2k-\deg(h_1)} + \cdots + h_s \cdot \mathbb{R}[x]_{2k-\deg(h_s)}.$$

The real variety of h is

$$V_{\mathbb{R}}(h) = \{x \in \mathbb{R}^n : h(x) = 0\}.$$

A polynomial $\sigma \in \mathbb{R}[x]$ is said to be a sum of squares (SOS) if there are polynomials $q_1, \ldots, q_t \in \mathbb{R}[x]$ such that $\sigma = q_1^2 + \cdots + q_t^2$. The convex cone of all SOS polynomials in x is denoted as $\Sigma[x]$. We refer to [8, 13, 17, 28] for more details. For a tuple of polynomials $g := (g_1, \ldots, g_t)$, its *quadratic module* is (let $g_0 := 1$)

$$\mathrm{QM}[g] := \Big\{ \sum_{i=0}^{t} \sigma_i g_i \,\Big|\, \text{each } \sigma_i \in \Sigma[x] \Big\}.$$

For a positive integer k, the degree-$2k$ truncation of $\mathrm{QM}[g]$ is

$$\mathrm{QM}[g]_{2k} := \Big\{ \sum_{i=0}^{t} \sigma_i g_i \,\Big|\, \sigma_i \in \Sigma[x], \deg(\sigma_i g_i) \leq 2k \Big\}.$$

The quadratic module $\mathrm{QM}[g]$ is said to be *archimedean* if there exists $q \in \mathrm{QM}[g]$ such that the set

$$S(q) := \{x \in \mathbb{R}^n \mid q(x) \geq 0\}$$

is compact.

Theorem 2.1.1. *[39] If* $\mathrm{QM}[g]$ *is archimedean and a polynomial* $f > 0$ *on* $S(g)$, *then* $f \in \mathrm{QM}[g]$.

A vector $y := (y_\alpha)_{\alpha \in \mathbb{N}_{2k}^n}$ is said to be a *truncated multi-sequences* (tms) of degree $2k$. For $y \in \mathbb{R}^{\mathbb{N}_{2k}^n}$, the *Riesz functional* determined by y is the linear functional $\mathscr{L}_y$ acting on $\mathbb{R}[x]_{2k}$ such that

$$\mathscr{L}_y \Big(\sum_{\alpha \in \mathbb{N}_{2k}^n} p_\alpha x^\alpha \Big) := \sum_{\alpha \in \mathbb{N}_{2k}^n} p_\alpha y_\alpha.$$

For convenience, we denote

$$\langle p, y \rangle := \mathscr{L}_y(p), \quad p \in \mathbb{R}[x]_{2k}.$$

The *localizing matrix* and *localizing vector* of p generated by y are respectively

$$L_p^{(k)}[y] := \mathscr{L}_y(p(x) \cdot [x]_{s_1}[x]_{s_1}^T),$$
$$\mathscr{V}_p^{(2k)}[y] := \mathscr{L}_y(p(x) \cdot [x]_{s_2}).$$

In the above, the linear operator is applied component-wisely and

$$s_1 := \lceil k - \deg(p)/2 \rceil, \quad s_2 := 2k - \deg(p).$$

We remark that $L_p^{(k)}[y] \succeq 0$ if and only if $\mathscr{L}_y \geq 0$ on $\mathrm{QM}[p]_{2k}$, and $\mathscr{V}_p^{(2k)}[y] = 0$ if and only if $\mathscr{L}_y = 0$ on $\mathrm{Ideal}[p]_{2k}$. More details for this can be found in [13, 17, 28]. The localizing matrix $L_p^{(k)}[y]$ satisfies the following equation

$$\left\langle p(x) \left(v^T [x]_s \right)^2, y \right\rangle = v^T \left(L_p^{(k)}[y] \right) v$$

for the degree $s := k - \lceil \deg(p)/2 \rceil$ and for every vector v of length $\binom{n+s}{s}$. For instance, when $n = 3$, $k = 3$ and $p = x_1 x_2 x_3 - x_3^3$,

$$L_p^{(3)}[y] = \begin{bmatrix} y_{111} - y_{003} & y_{211} - y_{103} & y_{121} - y_{013} & y_{112} - y_{004} \\ y_{211} - y_{103} & y_{311} - y_{203} & y_{221} - y_{113} & y_{212} - y_{104} \\ y_{121} - y_{013} & y_{221} - y_{113} & y_{131} - y_{023} & y_{122} - y_{014} \\ y_{112} - y_{004} & y_{212} - y_{104} & y_{122} - y_{014} & y_{113} - y_{005} \end{bmatrix}.$$

In particular, for $p = 1$, we get the *moment matrix* $M_k[y] := L_1^{(k)}[y]$. Similarly, the localizing vector $\mathscr{V}_p^{(2k)}[y]$ satisfies

$$\left\langle p(x) \left(v^T [x]_t \right), y \right\rangle = \left(\mathscr{V}_p^{(2k)}[y] \right)^T v$$

for $t := 2k - \deg(p)$. For instance, when $n = 3$, $k = 2$ and $p = x_1^2 + x_2^2 + x_3^2 - 1$,

$$\mathscr{V}_p^{(4)}[y] = \begin{bmatrix} y_{200} + y_{020} + y_{002} - y_{000} \\ y_{300} + y_{120} + y_{102} - y_{100} \\ y_{210} + y_{030} + y_{012} - y_{010} \\ y_{201} + y_{021} + y_{003} - y_{001} \\ y_{400} + y_{220} + y_{202} - y_{200} \\ y_{310} + y_{130} + y_{112} - y_{110} \\ y_{301} + y_{121} + y_{103} - y_{101} \\ y_{220} + y_{040} + y_{022} - y_{020} \\ y_{211} + y_{031} + y_{013} - y_{011} \\ y_{202} + y_{022} + y_{004} - y_{002} \end{bmatrix}.$$

It is worthy to note that if $L_{g_i}^{(k)}[y] \succeq 0$ and $f \in \mathrm{QM}[g]_{2k}$, then $\langle f, y \rangle \geq 0$. This can be seen as follows. For $f = \sum_{i=0}^{t} g_i \sigma_i$ with $\sigma_i = \sum_j p_{ij}^2 \in \Sigma[x]$ and $\deg(g_i \sigma_i) \leq 2k$, we have

$$\langle f, y \rangle = \left\langle \sum_{i=0}^{t} g_i \sigma_i, y \right\rangle = \sum_{i,j} \mathrm{vec}(p_{ij})^T \left(L_{g_i}^{(k)}[y] \right) \mathrm{vec}(p_{ij}) \geq 0.$$

A tms $y \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ is said to *admit* a Borel measure μ if

$$y_\alpha = \int x^\alpha d\mu \quad \text{for all } \alpha \in \mathbb{N}_{2k}^n.$$

Such μ is called a *representing measure* for y. The *support* of μ is the smallest closed set $S \subseteq \mathbb{R}^n$ such that $\mu(\mathbb{R}^n \setminus S) = 0$, denoted as $supp(\mu)$. The measure μ is said to be supported in a set K if $supp(\mu) \subseteq K$.

2.2 Moment relaxation

Consider the polynomial optimization problem

$$\begin{cases} \min & f(x) \\ s.t. & c_i(x) = 0 \ (i \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}), \end{cases} \qquad (2.2.1)$$

where f, c_i, c_j are polynomials in x. The kth order moment relaxation for (2.2.1) is

$$\begin{cases} \min & \langle f, y \rangle \\ s.t. & \mathscr{V}_{c_i}^{(2k)}[y] = 0 \ (i \in \mathcal{E}), \\ & L_{c_j}^{(k)}[y] \succeq 0 \ (j \in \mathcal{I}), \\ & M_k[y] \succeq 0, \\ & y_0 = 1, y \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases}$$

Suppose the tms y^* is a minimizer of above. Denote the degree

$$d := \max_{i \in \mathcal{E} \cup \mathcal{I}} \{ \lceil \deg(c_i)/2 \rceil \}.$$

We can extract minimizers if y^* satisfies the *flat truncation* condition: there exists an integer $k \geq t \geq \max\{d, \deg(f)/2\}$ such that

$$\operatorname{rank} M_{t-d}[y^*] = \operatorname{rank} M_t[y^*]. \qquad (2.2.2)$$

Interestingly, if (2.2.2) holds, we can extract $r := \operatorname{rank} M_t[y^*]$ minimizers for the optimization problem (2.2.1).

The following result is based on work by Curto and Fialkow [3] and Henrion and Lasserre [6]. The form of the result as presented here can be found in book [28, Section 2.7].

Theorem 2.2.1. *[3, 6] If y^* satisfies (2.2.2), then there exist $r := \operatorname{rank} M_t[y^*]$ distinct feasible points $u_1, \ldots, u_r$ for (2.2.1) and positive scalars $\lambda_1, \ldots, \lambda_r$ such that*

$$y^*|_{2t} = \lambda_1[u_1]_{2t} + \cdots + \lambda_r[u_r]_{2t}.$$

In the above, the notation $y^*|_{2t}$ stands for its subvector of entries that are labeled by $\alpha \in \mathbb{N}_{2t}^n$.

2.3 Optimality conditions

Suppose u is a local minimizer of (2.2.1). Denote the active labeling set

$$J(u) := \{ j \in \mathcal{I} : c_j(u) = 0 \}.$$

The *linear independence constraint qualification condition* (LICQC) holds at u if the gradient set $\{\nabla c_i(u)\}_{i \in \mathcal{E} \cup J(u)}$ is linearly independent. When LICQC holds, there exists a Lagrange multiplier vector

$$\lambda := (\lambda_i)_{i \in \mathcal{E}} \cup (\lambda_j)_{j \in \mathcal{I}}$$

satisfying the Karush-Kuhn-Tucker (KKT) conditions

$$\nabla f(u) = \sum_{i \in \mathcal{E}} \lambda_i \nabla c_i(u) + \sum_{j \in \mathcal{I}} \lambda_j \nabla c_j(u), \tag{2.3.1}$$

$$0 \le c_j(u) \perp \lambda_j \ge 0, \quad \text{for all } j \in \mathcal{I}. \tag{2.3.2}$$

The equation (2.3.1) is known as the *first order optimality condition* (FOOC), and (2.3.2) is called the *complementarity condition* (CC). If, in addition, $\lambda_j + c_j(u) > 0$ for all $j \in \mathcal{I}$, the *strict complementarity condition* (SCC) is said to hold at u. For the λ_i satisfying (2.3.1)-(2.3.2), the Lagrange function is

$$L(x) := f(x) - \sum_{i \in \mathcal{E}} \lambda_i c_i(x) - \sum_{j \in \mathcal{I}} \lambda_j c_j(x).$$

The Hessian of the Lagrangian is

$$\nabla^2 L(x) := \nabla^2 f(x) - \sum_{i \in \mathcal{E}} \lambda_i \nabla^2 c_i(x) - \sum_{j \in \mathcal{I}} \lambda_j \nabla^2 c_j(x).$$

If u is a local minimizer and LICQC holds, the *second order necessary condition* (SONC) holds at u:

$$v^T \left(\nabla^2 L(u) \right) v \ge 0 \quad \text{for all} \quad v \in \bigcap_{i \in \mathcal{E} \cup J(u)} \nabla c_i(u)^\perp,$$

where $\nabla c_i(u)^\perp := \{ v \in \mathbb{R}^n \mid \nabla c_i(u)^T v = 0 \}$. Stronger than SONC is the *second order sufficient condition* (SOSC):

$$v^T \left(\nabla^2 L(u) \right) v > 0 \quad \text{for all} \quad 0 \ne v \in \bigcap_{i \in \mathcal{E} \cup J(u)} \nabla c_i(u)^\perp.$$

If a feasible point u satisfies FOOC, SCC, and SOSC, then u must be a strict local minimizer. We refer to the book [1] for optimality conditions in nonlinear programming.

Chapter 3

A Unified Moment-SOS Hierarchy

In this chapter, we give a unified hierarchy of Moment-SOS relaxations to solve (1.1.1). Under some assumptions, we prove this hierarchy has asymptotic or finite convergence.

3.1 Unified Moment-SOS relaxations

For convenience of description, we denote the equality and inequality constraining polynomial tuples for K_l as

$$c_{eq}^{(l)} := (c_i^{(l)})_{i \in \mathcal{E}^{(l)}}, \quad c_{in}^{(l)} := (c_j^{(l)})_{j \in \mathcal{I}^{(l)}}.$$

Recall that $\text{Ideal}[c_{eq}^{(l)}]$ denotes the ideal generated by $c_{eq}^{(l)}$ and $\text{QM}[c_{in}^{(l)}]$ denotes the quadratic module generated by $c_{in}^{(l)}$. We refer to Section 2.1 for the notation. The minimum value of (1.1.1) is denoted as f_{min} and its feasible set is K. We look for the largest scalar γ that is a lower bound of f over K, i.e., $f - \gamma \in \mathscr{P}(K)$. Since

$$K = K_1 \cup K_2 \cup \cdots \cup K_m,$$

we have $f - \gamma \geq 0$ on K if and only if $f - \gamma \geq 0$ on K_l for every $l = 1, \ldots, m$. Note that $f - \gamma \geq 0$ on K_l is ensured by the membership (for some degree $2k$)

$$f - \gamma \in \text{Ideal}[c_{eq}^{(l)}]_{2k} + \text{QM}[c_{in}^{(l)}]_{2k}.$$

The kth order SOS relaxation for solving (1.1.1) is therefore

$$\begin{cases} \max & \gamma \\ s.t. & f - \gamma \in \bigcap_{l=1}^{m} \left[\text{Ideal}[c_{eq}^{(l)}]_{2k} + \text{QM}[c_{in}^{(l)}]_{2k} \right]. \end{cases} \quad (3.1.1)$$

The dual optimization of (3.1.1) is then the moment relaxation

$$
\begin{cases}
\min & \langle f , y^{(1)} \rangle + \cdots + \langle f , y^{(m)} \rangle \\
\text{s.t.} & \mathscr{V}^{(2k)}_{c^{(l)}_i}[y^{(l)}] = 0 \ (i \in \mathcal{E}^{(l)}), \\
& L^{(k)}_{c^{(l)}_j}[y^{(l)}] \succeq 0 \ (j \in \mathcal{I}^{(l)}), \\
& M_k[y^{(l)}] \succeq 0, \sum_{l=1}^{m} y^{(l)}_0 = 1, \\
& y^{(l)} \in \mathbb{R}^{\mathbb{N}^n_{2k}}, l = 1, \ldots, m.
\end{cases}
\tag{3.1.2}
$$

The integer k is called the *relaxation order*. For $k = 1, 2, \ldots$, the sequence of primal-dual pairs (3.1.1)-(3.1.2) is called the *unified* Moment-SOS hierarchy. For each k, we denote by $f_{sos,k}$ and $f_{mom,k}$ the optimal values of (3.1.1) and (3.1.2) respectively. We remark that the moment relaxation (3.1.2) can be equivalently written in terms of Riesz functional. Let $\mathscr{L}^{(l)}$ denote the Riesz functional given by $y^{(l)}$, then (3.1.2) is equivalent to

$$
\begin{cases}
\min & \mathscr{L}^{(1)}(f) + \cdots + \mathscr{L}^{(m)}(f) \\
\text{s.t.} & \mathscr{L}^{(l)} = 0 \text{ on } \mathrm{Ideal}[c^{(l)}_{eq}]_{2k}, \\
& \mathscr{L}^{(l)} \geq 0 \text{ on } \mathrm{QM}[c^{(l)}_{in}]_{2k}, \\
& \mathscr{L}^{(1)}(1) + \cdots + \mathscr{L}^{(m)}(1) = 1, \\
& l = 1, \ldots, m.
\end{cases}
$$

Proposition 3.1.1. *For each relaxation order k, it holds that*

$$
f_{sos,k} \leq f_{mom,k} \leq f_{min}.
\tag{3.1.3}
$$

Moreover, both sequences $\{f_{sos,k}\}_{k=1}^{\infty}$ and $\{f_{mom,k}\}_{k=1}^{\infty}$ are monotonically increasing.

Proof. By the weak duality, we have $f_{sos,k} \leq f_{mom,k}$. For every $\epsilon > 0$, there exist $l' \in [m]$ and $u \in K_{l'}$ such that $f(u) \leq f_{min} + \epsilon$. Let $y := (y^{(1)}, \ldots, y^{(m)})$ be such that $y^{(l')} = [u]_{2k}$ and $y^{(l)} = 0$ for all $l \in [m] \setminus \{l'\}$. Then, y is feasible for (3.1.2) and

$$
f_{mom,k} \leq \langle f , y^{(1)} \rangle + \cdots + \langle f , y^{(m)} \rangle = \langle f , y^{(l')} \rangle = f(u) \leq f_{min} + \epsilon.
$$

Since $\epsilon > 0$ can be arbitrary, $f_{mom,k} \leq f_{min}$. Therefore, we get (3.1.3). Clearly, if γ is feasible for (3.1.1) with an order k, then γ must also be feasible for (3.1.1) with all larger values of k, since the feasible set gets larger as k increases. So the sequence of lower bounds $\{f_{sos,k}\}_{k=1}^{\infty}$ is monotonically increasing. On the other hand, when k increases, the feasible set of (3.1.2) shrinks, so the minimum value of (3.1.2) increases. Therefore, $\{f_{mom,k}\}_{k=1}^{\infty}$ is also monotonically increasing. $\square$

3.2   Extraction of minimizers

We show how to extract minimizers of (1.1.1) from the unified moment relaxation. This is a natural extension from the case $m = 1$ in Section 2.2.

Suppose the tuple $y^* := (y^{(*,1)}, \ldots, y^{(*,m)})$ is a minimizer of (3.1.2). Denote the degree

$$d_l := \max_{i \in \mathcal{E}^{(l)} \cup \mathcal{I}^{(l)}} \{\lceil \deg(c_i^{(l)})/2 \rceil\}.$$

We can extract minimizers by checking the *flat truncation* condition: there exists an integer $t \geq \max_{l \in [m]} \{d_l, \deg(f)/2\}$ such that

$$\operatorname{rank} M_{t-d_l}[y^{(*,l)}] = \operatorname{rank} M_t[y^{(*,l)}] \quad \text{for each } l \in \mathcal{A}, \qquad (3.2.1)$$

where the labeling set

$$\mathcal{A} := \{l \in [m] : y_0^{(l)} > 0\}.$$

Interestingly, if (3.2.1) holds, we can extract

$$r := \sum_{l \in \mathcal{A}} \operatorname{rank} M_t[y^{(*,l)}] \qquad (3.2.2)$$

minimizers for the optimization problem (1.1.1).

Algorithm 3.2.1. To solve the polynomial optimization (1.1.1), do the following:

Step 0 Let $k := \max_{l \in [m]} \{d_l, \lceil \deg(f)/2 \rceil\}.$

Step 1 Solve the relaxation (3.1.2). If it is infeasible, output that (1.1.1) is infeasible and stop. Otherwise, solve it for a minimizer $y^* := (y^{(*,1)}, \ldots, y^{(*,m)})$.

Step 2 Check if the flat truncation (3.2.1) holds or not. If (3.2.1) holds, then the relaxation (3.1.2) is tight and for each $l \in \mathcal{A}$, the truncation $y^{(*,l)}|_{2t}$ admits a finitely atomic measure $\mu^{(l)}$ such that each point in $supp(\mu^{(l)})$ is a minimizer of (1.1.1). Moreover, $f_{min} = f_{mom,k}$.

Step 3 If (3.2.1) fails, let $k := k + 1$ and go to Step 1.

The conclusion in Step 2 is justified by the following.

Theorem 3.2.2. *Let $y^* := (y^{(*,1)}, \ldots, y^{(*,m)})$ be a minimizer of (3.1.2). Suppose (3.2.1) holds for all $l \in \mathcal{A}$. Then, the moment relaxation (3.1.2) is tight and for each $l \in \mathcal{A}$, the truncation*

$$y^{(*,l)}|_{2t} := (y_\alpha^{(*,l)})_{\alpha \in \mathbb{N}_{2t}^n}$$

admits a r_l-atomic measure $\mu^{(l)}$, where $r_l = \operatorname{rank} M_t[y^{(*,l)}]$, and each point in $supp(\mu^{(l)})$ is a minimizer of (1.1.1). Therefore, the total number of minimizers is r as in (3.2.2).

Proof. By the assumption, the $y^{(*,l)} \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ satisfies (3.2.1) and

$$L_{c_j^{(l)}}^{(k)}[y^{(*,l)}] \succeq 0 \ (j \in \mathcal{I}^{(l)}), \quad M_k[y^{(*,l)}] \succeq 0.$$

Then, by Theorem 2.2.1, there exists r_l distinct points $u_1^{(l)}, \ldots, u_{r_l}^{(l)} \in K_l$ and positive scalars $\lambda_1^{(l)}, \ldots, \lambda_{r_l}^{(l)}$ such that

$$y^{(*,l)}|_{2t} = \lambda_1^{(l)}[u_1^{(l)}]_{2t} + \cdots + \lambda_{r_l}^{(l)}[u_{r_l}^{(l)}]_{2t}.$$

The constriant $\sum_{l=1}^{m} y_0^{(l)} = 1$ implies that $\sum_{l=1}^{m} \sum_{i=1}^{r_l} \lambda_i^{(l)} = 1$, so

$$\sum_{l=1}^{m} \sum_{i=1}^{r_l} \lambda_i^{(l)} f(u_i^{(l)}) = \sum_{l=1}^{m} \langle f, y^{(*,l)}|_{2t} \rangle = \sum_{l=1}^{m} \langle f, y^{(*,l)} \rangle = f_{mom,k} \leq f_{min}.$$

For each $u_i^{(l)} \in K_l$, we have $f(u_i^{(l)}) \geq f_{min}$, so

$$\sum_{l=1}^{m} \sum_{i=1}^{r_l} \lambda_i^{(l)} f(u_i^{(l)}) \geq \sum_{l=1}^{m} \sum_{i=1}^{r_l} \lambda_i^{(l)} f_{min} = f_{min}.$$

Hence, $f_{mom,k} = f_{min}$ and

$$\sum_{l=1}^{m} \sum_{i=1}^{r_l} \lambda_i^{(l)} \left[f(u_i^{(l)}) - f_{min} \right] = 0.$$

Since each $\lambda_i^{(l)} > 0$, then each $f(u_i^{(l)}) = f_{min}$, i.e., each $u_i^{(l)}$ is a minimizer of (1.1.1). $\qquad\square$

In Step 2, the flat truncation condition (3.2.1) is used to extract minimizers. When it holds, a numerical method is given in [6] for computing the minimizers. We refer to [28, Section 2.7] for more details.

3.3   Convergence analysis

Recall that $f_{min}, f_{sos,k}$ and $f_{mom,k}$ denote the optimal values of (1.1.1), (3.1.1) and (3.1.2) respectively. The unified Moment-SOS hierarchy (3.1.1)-(3.1.2) is said to have *asymptotic convergence* if $f_{sos,k} \to f_{min}$ as $k \to \infty$. If $f_{sos,k} = f_{min}$ for some k, this unified hierarchy is said to be *tight* or to have *finite convergence*. The following theorem is a natural extension from the case $m = 1$.

Theorem 3.3.1 (Asymptotic convergence)**.** *If* $\mathrm{Ideal}[c_{eq}^{(l)}] + \mathrm{QM}[c_{in}^{(l)}]$ *is archimedean for every* $l = 1, \ldots, m$, *then the Moment-SOS hierarchy (3.1.1)-(3.1.2) has asymptotic convergence:*

$$\lim_{k \to \infty} f_{sos,k} = \lim_{k \to \infty} f_{mom,k} = f_{min}.$$

Proof. For $\epsilon > 0$, let $\gamma = f_{min} - \epsilon$. Then

$$f(x) - \gamma = f(x) - f_{min} + \epsilon > 0$$

on K_l. Since $\text{Ideal}[c_{eq}^{(l)}] + \text{QM}[c_{in}^{(l)}]$ is archimedean for every l, by Theorem 2.1.1,

$$f(x) - \gamma \in \text{Ideal}[c_{eq}^{(l)}]_{2k} + \text{QM}[c_{in}^{(l)}]_{2k}$$

for all k large enough. So

$$f_{min} - \epsilon = \gamma \leq f_{sos,k} \leq f_{min} \implies f_{min} - \epsilon \leq \lim_{k \to \infty} f_{sos,k} \leq f_{min}.$$

Since $\epsilon > 0$ can be arbitrary, $\lim_{k \to \infty} f_{sos,k} = f_{min}$. By (3.1.3), we get the desired conclusion. □

Recall the linear independence constraint qualification condition (LICQC), the strict complementarity condition (SCC), and the second order sufficient condition (SOSC) introduced in Section 2.3. The following is the conclusion for the finite convergence of the unified Moment-SOS hierarchy of (3.1.1)-(3.1.2).

Theorem 3.3.2 (Finite convergence). *Assume* $\text{Ideal}[c_{eq}^{(l)}] + \text{QM}[c_{in}^{(l)}]$ *is archimedean for every* $l = 1, \ldots, m$. *If the LICQC, SCC, and SOSC hold at every global minimizer of (1.1.1) for each* K_l, *then the Moment-SOS hierarchy (3.1.1)-(3.1.2) has finite convergence, i.e.,*

$$f_{sos,k} = f_{mom,k} = f_{min}$$

for all k *large enough.*

Proof. We denote by $f_{min,l}$ the minimum value of f on the set K_l. Let

$$\mathcal{B} := \{l : f_{min,l} = f_{min}\}.$$

(i) For the case $l \notin \mathcal{B}$, $f_{min,l} > f_{min}$,

$$f(x) - f_{min} \geq f_{min,l} - f_{min} > 0$$

on K_l. Since $\text{Ideal}[c_{eq}^{(l)}] + \text{QM}[c_{in}^{(l)}]$ is archimedean, there exists k_0 such that

$$f - (f_{min} - \epsilon) \in \text{Ideal}[c_{eq}^{(l)}]_{2k_0} + \text{QM}[c_{in}^{(l)}]_{2k_0}$$

for all $\epsilon > 0$.

(ii) For the case $l \in \mathcal{B}$, $f_{min,l} = f_{min}$. Since the LICQC, SCC, and SOSC hold at every global minimizer x^* of (1.1.1), there exists a degree k_0 such that for all $\epsilon > 0$, we have

$$f - (f_{min} - \epsilon) \in \text{Ideal}[c_{eq}^{(l)}]_{2k_0} + \text{QM}[c_{in}^{(l)}]_{2k_0}.$$

This is shown in the proof of Theorem 1.1 of [25].

Combining cases (i) and (ii), we know that $\gamma = f_{min} - \epsilon$ is feasible for (3.1.1) with the order k_0. Hence, $f_{sos,k_0} \geq \gamma = f_{min} - \epsilon$. Since $\epsilon > 0$ can be arbitrary, we get $f_{sos,k_0} \geq f_{min}$. By Proposition 3.1.1, we get $f_{sos,k} = f_{mom,k} = f_{min}$ for all $k \geq k_0$. □

Chapter 4

Univariate Polynomial Optimization

In this chapter, we consider the special case of univariate polynomial optimization, i.e., $n = 1$. The following results for the univariate case are extensions from the single interval case, and are presented here to provide a complete and thorough understanding for convenience of readers. The problem (1.1.1) can be expressed as

$$\begin{cases} \min & f(x) := f_0 + f_1 x + \cdots + f_d x^d \\ s.t. & x \in \bigcup_{l=1}^{m} K_l, \end{cases} \tag{4.0.1}$$

where $K_l = [a_l, b_l]$ with $a_l < b_l$ for $l = 1, \ldots, m$. We still denote by f_{min} the minimum value of (4.0.1). For convenience, we only consider compact intervals. The discussions for unbounded intervals like $(-\infty, b]$ and $[a, +\infty)$ are similar (see [28, Chapter 3]).

Let $y := (y_0, \ldots, y_d) \in \mathbb{R}^{d+1}$ be a univariate tms of degree d with $d = 2d_0 + 1$ or $d = 2d_0$. The d_0th order moment matrix of y is

$$M_{d_0}[y] = \begin{bmatrix} y_0 & y_1 & \cdots & y_{d_0} \\ y_1 & y_2 & \cdots & y_{d_0+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{d_0} & y_{d_0+1} & \cdots & y_{2d_0} \end{bmatrix}.$$

For convenience of notation, we also denote that

$$G_{d_0}[y] := \begin{bmatrix} y_2 & y_3 & \cdots & y_{d_0+1} \\ y_3 & y_4 & \cdots & y_{d_0+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{d_0+1} & y_{d_0+2} & \cdots & y_{2d_0} \end{bmatrix},$$

$$N_{d_0}[y] := \begin{bmatrix} y_1 & y_2 & \cdots & y_{d_0+1} \\ y_2 & y_3 & \cdots & y_{d_0+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{d_0+1} & y_{d_0+2} & \cdots & y_{2d_0+1} \end{bmatrix}.$$

It is well-known that polynomials that are nonnegative over an interval can be expressed in terms of sum of squares. The following results were known to Lukács [19], Markov [22], Pólya and Szegö [37], Powers and Reznick [38]. For each $h \in \mathbb{R}[x]_d$ that is nonnegative on the interval $[a_l, b_l]$, we have:

(i) If $d = 2d_0 + 1$ is odd, then there exist $p, q \in \mathbb{R}[x]_{d_0}$ such that

$$h = (x - a_l)p^2 + (b_l - x)q^2. \tag{4.0.2}$$

(ii) If $d = 2d_0$ is even, then there exist $p \in \mathbb{R}[x]_{d_0}, q \in \mathbb{R}[x]_{d_0-1}$ such that

$$h = p^2 + (x - a_l)(b_l - x)q^2. \tag{4.0.3}$$

The optimization problem (4.0.1) can be solved by the unified Moment-SOS hierarchy of (3.1.1)-(3.1.2). For the univariate case, they can be simplified. We discuss in two different cases of d.

4.1   The case d is odd $(d = 2d_0 + 1)$

When the degree $d = 2d_0 + 1$ is odd, by the representation (4.0.2), f_{min} equals the maximum value of the SOS relaxation

$$\begin{cases} \max & \gamma \\ \text{s.t.} & f - \gamma = (x - a_l)[x]_{d_0}^T X_l[x]_{d_0} + (b_l - x)[x]_{d_0}^T Y_l[x]_{d_0}, \\ & X_l \in S_+^{d_0+1}, Y_l \in S_+^{d_0+1}, l = 1, \ldots, m. \end{cases} \tag{4.1.1}$$

Its dual optimization is the moment relaxation

$$\begin{cases} \min & \langle f, y^{(1)} \rangle + \cdots + \langle f, y^{(m)} \rangle \\ \text{s.t.} & y_0^{(1)} + \cdots + y_0^{(m)} = 1, \\ & b_l M_{d_0}[y^{(l)}] \succeq N_{d_0}[y^{(l)}] \succeq a_l M_{d_0}[y^{(l)}], \\ & y^{(l)} = (y_0^{(l)}, y_1^{(l)}, \ldots, y_{2d_0+1}^{(l)}), \\ & l = 1, \ldots, m. \end{cases} \tag{4.1.2}$$

In the above,

$$\langle f, y^{(l)} \rangle = f_0 y_0^{(l)} + \cdots + f_{2d_0+1} y_{2d_0+1}^{(l)}.$$

Denote by f_{sos} and f_{mom} the optimal values of (4.1.1) and (4.1.2) respectively. For all (X_l, Y_l) that is feasible for (4.1.1) and for all $y^{(l)}$ that is feasible for (4.1.2), we have

$$\langle f, y^{(1)} \rangle + \cdots + \langle f, y^{(m)} \rangle \geq \gamma.$$

This is because

$$b_l M_{d_0}[y^{(l)}] - N_{d_0}[y^{(l)}] = L_{b_l - x}^{(d_0+1)}[y^{(l)}] \succeq 0,$$

$$N_{d_0}[y^{(l)}] - a_l M_{d_0}[y^{(l)}] = L_{x - a_l}^{(d_0+1)}[y^{(l)}] \succeq 0,$$

which implies that

$$
\begin{aligned}
&\langle f, y^{(1)} \rangle + \cdots + \langle f, y^{(m)} \rangle - \gamma \\
&= \langle f - \gamma, y^{(1)} \rangle + \cdots + \langle f - \gamma, y^{(m)} \rangle \\
&= \sum_{l=1}^{m} \left[\left\langle L_{x-a_l}^{(d_0+1)}[y^{(l)}], X_l \right\rangle + \left\langle L_{b_l-x}^{(d_0+1)}[y^{(l)}], Y_l \right\rangle \right] \\
&\geq 0.
\end{aligned}
$$

Indeed, we have the following theorem.

Theorem 4.1.1. *For the relaxations (4.1.1) and (4.1.2), we always have*

$$f_{sos} = f_{mom} = f_{min}.$$

Proof. By the representation (4.0.2), for $\gamma = f_{min}$, the subtraction $f - f_{min}$ can be represented as in (4.1.1) for each $l = 1, \ldots, m$, so $f_{sos} = f_{min}$. By the weak duality, we have $f_{sos} \leq f_{mom} \leq f_{min}$. Hence, they are all equal. □

The optimizers for (4.0.1) can be obtained by the following algorithm.

Algorithm 4.1.2. *[28, Algorithm 3.3.6]* Assume $d = 2d_0 + 1$ and $(y^{(1)}, \ldots, y^{(m)})$ is a minimizer for the moment relaxation (4.1.2). For each l with $y_0^{(l)} > 0$ and $r = \text{rank } M_{d_0}[y^{(l)}]$, do the following:

Step 1 Solve the linear system

$$
\begin{bmatrix}
y_0^{(l)} & y_1^{(l)} & \cdots & y_{r-1}^{(l)} \\
y_1^{(l)} & y_2^{(l)} & \cdots & y_r^{(l)} \\
\vdots & \vdots & \ddots & \vdots \\
y_{2d_0-r+1}^{(l)} & y_{2d_0-r+2}^{(l)} & \cdots & y_{2d_0}^{(l)}
\end{bmatrix}
\begin{bmatrix}
g_0^{(l)} \\
g_1^{(l)} \\
\vdots \\
g_{r-1}^{(l)}
\end{bmatrix}
=
\begin{bmatrix}
y_r^{(l)} \\
y_{r+1}^{(l)} \\
\vdots \\
y_{2d_0+1}^{(l)}
\end{bmatrix}.
$$

Step 2 Compute r distinct roots $t_1^{(l)}, \ldots, t_r^{(l)}$ of the polynomial

$$g^{(l)}(x) := g_0^{(l)} + g_1^{(l)} x + \cdots + g_{r-1}^{(l)} x^{r-1} - x^r.$$

Step 3 The roots $t_1^{(l)}, \ldots, t_r^{(l)}$ are minimizers of the optimization problem (4.0.1).

The conclusion in Step 3 is justified by Theorem 4.2.3. The following is an exposition for the above algorithm.

Example 4.1.3. Consider the constrained optimization problem

$$\begin{cases} \min & x + 2x^6 - x^7 \\ s.t. & x \in [-2, -1] \cup [1, 2]. \end{cases}$$

The moment relaxation is

$$\begin{cases} \min & \langle f , y^{(1)} \rangle + \langle f , y^{(2)} \rangle \\ s.t. & -M_3[y^{(1)}] \succeq N_3[y^{(1)}] \succeq -2M_3[y^{(1)}], \\ & 2M_3[y^{(2)}] \succeq N_3[y^{(2)}] \succeq M_3[y^{(2)}], \\ & y_0^{(1)} + y_0^{(2)} = 1. \end{cases}$$

The minimizer $y^* = (y^{(*,1)}, y^{(*,2)})$ of the above is obtained as

$$y^{(*,1)} = 0.4191 \cdot (1, -1, 1, -1, 1, -1, 1, -1),$$

$$y^{(*,2)} = 0.5809 \cdot (1, 0, -2, -6, -14, -30, -62, -126) +$$
$$0.6058 \cdot (0, 1, 3, 7, 15, 31, 63, 127).$$

Applying Algorithm 4.1.2, we get $g_0^{(1)} = -1, g_0^{(2)} = -2, g_1^{(2)} = 3$ and the polynomials

$$g^{(1)}(x) = -1 - x, \quad g^{(2)}(x) = -2 + 3x - x^2.$$

Therefore, the minimizers are the distinct roots $-1, 1, 2$ and the global minimum value $f_{min} = 2$.

4.2   The case d is even $(d = 2d_0)$

When the degree $d = 2d_0$ is even, by the representation (4.0.3), f_{min} equals the maximum value of

$$\begin{cases} \max & \gamma \\ s.t. & f - \gamma = [x]_{d_0}^T X_l [x]_{d_0} + (x - a_l)(b_l - x)([x]_{d_0-1}^T Y_l [x]_{d_0-1}), \quad (4.2.1) \\ & X_l \in S_+^{d_0+1}, Y_l \in S_+^{d_0}, l = 1, \dots, m. \end{cases}$$

Its dual optimization is the moment relaxation

$$\begin{cases} \min & \langle f , y^{(1)} \rangle + \cdots + \langle f , y^{(m)} \rangle \\ s.t. & (a_l + b_l)N_{d_0-1}[y^{(l)}] \succeq a_l b_l M_{d_0-1}[y^{(l)}] + G_{d_0}[y^{(l)}], \\ & M_{d_0}[y^{(l)}] \succeq 0, y_0^{(1)} + \cdots + y_0^{(m)} = 1, \quad (4.2.2) \\ & y^{(l)} = (y_0^{(l)}, y_1^{(l)}, \dots, y_{2d_0}^{(l)}), \\ & l = 1, \dots, m. \end{cases}$$

We still denote by f_{sos} and f_{mom} the optimal values of (4.2.1) and (4.2.2) respectively. The same conclusion in Theorem 4.1.1 also holds here. The optimizers for (4.0.1) can be obtained by the following algorithm.

Algorithm 4.2.1. *[28, Algorithm 3.3.6]* Assume $d = 2d_0$ and $(y^{(1)}, \ldots, y^{(m)})$ is a minimizer for the moment relaxation (4.2.2). For each l with $y_0^{(l)} > 0$ and $r = \operatorname{rank} M_{d_0}[y^{(l)}]$, do the following:

Step 1 If $r \le d_0$, solve the linear system

$$
\begin{bmatrix}
y_0^{(l)} & y_1^{(l)} & \cdots & y_{r-1}^{(l)} \\
y_1^{(l)} & y_2^{(l)} & \cdots & y_r^{(l)} \\
\vdots & \vdots & \ddots & \vdots \\
y_{2d_0-r}^{(l)} & y_{2d_0-r+1}^{(l)} & \cdots & y_{2d_0-1}^{(l)}
\end{bmatrix}
\begin{bmatrix}
g_0^{(l)} \\
g_1^{(l)} \\
\vdots \\
g_{r-1}^{(l)}
\end{bmatrix}
=
\begin{bmatrix}
y_r^{(l)} \\
y_{r+1}^{(l)} \\
\vdots \\
y_{2d_0}^{(l)}
\end{bmatrix}.
$$

Step 2 If $r = d_0 + 1$, compute the smallest value of $y_{2d_0+1}^{(l)}$ satisfying

$$
b_l M_{d_0}[y^{(l)}] \succeq N_{d_0}[y^{(l)}] \succeq a_l M_{d_0}[y^{(l)}],
$$

then solve the linear system

$$
\begin{bmatrix}
y_0^{(l)} & y_1^{(l)} & \cdots & y_{d_0}^{(l)} \\
y_1^{(l)} & y_2^{(l)} & \cdots & y_{d_0+1}^{(l)} \\
\vdots & \vdots & \ddots & \vdots \\
y_{d_0}^{(l)} & y_{d_0+1}^{(l)} & \cdots & y_{2d_0}^{(l)}
\end{bmatrix}
\begin{bmatrix}
g_0^{(l)} \\
g_1^{(l)} \\
\vdots \\
g_{d_0}^{(l)}
\end{bmatrix}
=
\begin{bmatrix}
y_{d_0+1}^{(l)} \\
y_{d_0+2}^{(l)} \\
\vdots \\
y_{2d_0+1}^{(l)}
\end{bmatrix}.
$$

Step 3 Compute r distinct roots $t_1^{(l)}, \ldots, t_r^{(l)}$ of the polynomial

$$
g^{(l)}(x) := g_0^{(l)} + g_1^{(l)} x + \cdots + g_{r-1}^{(l)} x^{r-1} - x^r.
$$

Step 4 The roots $t_1^{(l)}, \ldots, t_r^{(l)}$ are minimizers of the optimization problem (4.0.1).

The conclusion in Step 4 is justified by Theorem 4.2.3. The following is an exposition for the above algorithm.

Example 4.2.2. Consider the constrained optimization problem

$$
\begin{cases}
\min & 4x^2 + 12x^3 + 13x^4 + 6x^5 + x^6 \\
s.t. & x \in [-4, -2] \cup [-1, 2].
\end{cases}
$$

The moment relaxation is

$$
\begin{cases}
\min & \langle f , y^{(1)} \rangle + \langle f , y^{(2)} \rangle \\
s.t. & -6N_2[y^{(1)}] \succeq 8M_2[y^{(1)}] + G_3[y^{(1)}], \\
& N_2[y^{(2)}] \succeq -2M_2[y^{(2)}] + G_3[y^{(2)}], \\
& M_3[y^{(1)}] \succeq 0, M_3[y^{(2)}] \succeq 0, \\
& y_0^{(1)} + y_0^{(2)} = 1.
\end{cases}
$$

The minimizer $y^* = (y^{(*,1)}, y^{(*,2)})$ of the above is obtained as

$$y^{(*,1)} = 0.0110 \cdot (1, -2, 4, -8, 16, -32, 64),$$
$$y^{(*,2)} = 0.9890 \cdot (1, 0, 0, 0, 0, 0, 0) +$$
$$0.2190 \cdot (0, -1, 1, -1, 1, -1, 1).$$

Applying Algorithm 4.2.1, we get $g_0^{(1)} = -2, g_0^{(2)} = 0, g_1^{(2)} = -1$ and the polynomials

$$g^{(1)}(x) = -2 - x, \quad g^{(2)}(x) = -x - x^2.$$

Therefore, the minimizers are the distinct roots $-2, -1, 0$ and the global minimum value $f_{min} = 0$.

The performance of the moment relaxations (4.1.2) and (4.2.2) can be summarized as follows.

Theorem 4.2.3. *Suppose f is a univariate polynomial of degree $d = 2d_0 + 1$ or $d = 2d_0$. Then, all the optimal values f_{min}, f_{sos}, f_{mom} are achieved for each corresponding optimization problem and they are all equal to each other. Suppose $y^* := (y^{(*,1)}, \ldots, y^{(*,m)})$ is a minimizer of (4.1.2) when $d = 2d_0 + 1$ or of (4.2.2) when $d = 2d_0$. Then, the tms*

$$z^* := y^{(*,1)} + \cdots + y^{(*,m)}$$

must admit a representing measure μ^* supported in K, and each point in the support of μ^* is a minimizer of (4.0.1). If f is not a constant polynomial, then f has at most $2m + \lceil (d-1)/2 \rceil$ minimizers and the representing measure μ^* for z^* must be r-atomic with

$$r \le 2m + \lceil (d-1)/2 \rceil.$$

Proof. Since each interval $[a_l, b_l]$ is compact, K is also compact. So the minimum value f_{min} is achievable, and it equals the largest $\gamma \in \mathbb{R}$ such that $f - \gamma$ is nonnegative on $[a_l, b_l]$ for every $l = 1, \ldots, m$, so $f_{min} = f_{sos}$ (see Theorem 4.1.1). Each of the moment relaxations (4.1.2) and (4.2.2) has a strictly feasible point, e.g., the tms $\hat{y}^{(l)} = \int_{a_l}^{b_l} [x]_{2d_0+1} \, dx$ is strictly feasible and

$$b_l M_{d_0}[\hat{y}^{(l)}] \succ N_{d_0}[\hat{y}^{(l)}] \succ a_l M_{d_0}[\hat{y}^{(l)}].$$

The tms $\tilde{y}^{(l)} = \int_{a_l}^{b_l} [x]_{2d_0} \, dx$ is strictly feasible and

$$(a_l + b_l) N_{d_0-1}[\tilde{y}^{(l)}] \succ a_l b_l M_{d_0-1}[\tilde{y}^{(l)}] + G_{d_0}[\tilde{y}^{(l)}], \quad M_{d_0}[\tilde{y}^{(l)}] \succ 0.$$

By the strong duality, $f_{sos} = f_{mom}$, and both (4.1.1) and (4.2.1) achieve their optimal values. By [28, Theorem 3.3.4], $y^{(*,l)}$ must admit a representing measure $\mu^{(l)}$ supported in $[a_l, b_l]$. Hence, z^* must admit a representing measure $\mu^* =$

$\mu^{(1)} + \cdots + \mu^{(m)}$ supported in K. The optimization problem (4.0.1) is then equivalent to the linear convex conic optimization

$$
\begin{cases}
\min & \int f \, d\mu \\
s.t. & \mu(K) = 1, \mu \in \mathscr{B}(K),
\end{cases}
\tag{4.2.3}
$$

where $\mathscr{B}(K)$ denotes the convex cone of all Borel measures whose supports are contained in K. We claim that if a Borel measure μ^* is a minimizer of (4.2.3), then each point in the support of μ^* is a minimizer of (4.0.1). Suppose $E \subseteq K$ is the set of minimizers of (4.0.1). For any $x^* \in E$, let δ_{x^*} denote the unit Dirac measure supported at x^*. Then, we have

$$
f_{min} = \int_K f_{min} \, d\mu^* \leq \int_K f(x) \, d\mu^* \leq \int_K f(x) \, d\delta_{x^*} = f(x^*) = f_{min}.
$$

Hence,

$$
0 = \int_K [f(x) - f_{min}] \, d\mu^* = \int_{supp(\mu^*)} [f(x) - f_{min}] \, d\mu^*.
$$

Thus, $f = f_{min}$ on $supp(\mu^*)$. This implies that $supp(\mu^*) \subseteq E$. So, every point in $supp(\mu^*)$ is a minimizer of (4.0.1).

Note that f has degree d. If f is not a constant polynomial, it can have at most $d - 1$ critical points. Moreover, the local maximizers and minimizers alternate. Thus, at most $\lceil (d-1)/2 \rceil$ of these critical points are local minimizers. On each interval $[a_l, b_l]$, two endpoints are possibly local minimizers. Since there are m intervals in total, f has at most $2m + \lceil (d - 1)/2 \rceil$ local minimizers on K. In the above, we have proved that each point in $supp(\mu^*)$ is a minimizer of (4.0.1). So the representing measure μ^* for z^* must be r-atomic with $r \leq 2m + \lceil (d - 1)/2 \rceil$. □

We refer to Algorithm 4.1.2 (when $d = 2d_0 + 1$) and Algorithm 4.2.1 (when $d = 2d_0$) for how to determine the support of the representing measure $\mu^{(l)}$ for $y^{(*,l)}$. The points in the support are all minimizers of (4.0.1). The upper bound for the number of minimizers is already sharp when $m = 1$. For instance, consider the optimization

$$
\begin{cases}
\min & x(1 - x)(x + 1) \\
s.t. & x \in [-1, 1].
\end{cases}
$$

There are 3 global minimizers $-1, 0, 1$ and $2m + \lceil (d - 1)/2 \rceil = 2 + 1 = 3$.

We would like to remark that the representations for nonnegative univariate polynomials have broad applications. In particular, it can be applied to the shape design of transfer functions for linear time invariant (LTI) single-input-single-output (SISO) systems [29]. Since the transfer function is rational, the optimization problem can be formulated in terms of coefficients of polynomials. We can then solve it by using representations of nonnegative univariate polynomials. For instance, we look for a transfer function such that it is close to a

piecewise constant shape. That is, we want the transfer function to be close to given constant values $\xi_1, \ldots, \xi_m$ in a union of m disjoint intervals $[a_l, b_l]$ with

$$a_1 < b_1 < a_2 < b_2 < \cdots < a_m < b_m.$$

As in [29], the transfer function can be written as $p_1(x)/p_2(x)$. Specifically, we want to get p_1, p_2 such that

$$p_1(x), p_2(x) \geq 0, \quad \forall x \geq 0,$$

$$(1 - \alpha)\xi_l \leq \frac{p_1(x)}{p_2(x)} \leq (1 + \beta)\xi_l, \quad \forall x \in [a_l, b_l], \ l = 1, \ldots, m.$$

The above is equivalent to the linear conic constraints

$$p_1(x), p_2(x) \in \mathscr{P}_d([0, \infty)),$$
$$p_1 - (1 - \alpha)\xi_l p_2 \in \mathscr{P}_d([a_l, b_l]), \ l = 1, \ldots, m,$$
$$(1 + \beta)\xi_l p_2 - p_1 \in \mathscr{P}_d([a_l, b_l]), \ l = 1, \ldots, m.$$

We refer to [29] for more details.

Chapter 5

Numerical Experiments

In this chapter, we present numerical experiments for how to solve polynomial optimization over the union of several basic closed semialgebraic sets. Algorithm 3.2.1 is applied to solve it. All computations are implemented using MATLAB R2022a on a MacBook Pro equipped with Apple M1 Max processor and 16GB RAM. The unified moment relaxation (3.1.2) is solved by the software `Gloptipoly` [7], which calls the SDP package `SeDuMi` [41]. For neatness, all computational results are displayed in four decimal digits.

5.1 Some examples

Example 5.1.1. Consider the constrained optimization problem

$$\begin{cases} \min\limits_{x \in \mathbb{R}^4} & (x_1^2 + x_2^2 + x_3^2 + x_4^2 + 1)^2 - 4(x_1^2 x_2^2 + x_2^2 x_3^2 + x_3^2 x_4^2 + x_4^2 + x_1^2) \\ s.t. & x \in K_1 \cup K_2 \cup K_3 \cup K_4, \end{cases}$$

where

$$K_1 = \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 \le 0\},$$
$$K_2 = \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_4^2 \le 0\},$$
$$K_3 = \{x \in \mathbb{R}^4 : x_1^2 + x_3^2 + x_4^2 \le 0\},$$
$$K_4 = \{x \in \mathbb{R}^4 : x_2^2 + x_3^2 + x_4^2 \le 0\}.$$

The objective function is a dehomogenization of the Horn's form [40]. For $k = 2$, we get $f_{mom,2} = 0$, and the flat truncation (3.2.1) is met for all $l \in \mathcal{A} = \{1, 4\}$. So, $f_{mom,2} = f_{min}$. The obtained four minimizers are

$$(0, 0, 0, \pm 1) \in K_1, \quad (\pm 1, 0, 0, 0) \in K_4.$$

For $k = 2$, the unified moment relaxation (3.1.2) took around 0.6 second, while solving the individual moment relaxations (1.2.1) for all K_1, K_2, K_3, K_4 took around 0.9 second.

Example 5.1.2. Consider the constrained optimization problem

$$\begin{cases} \min\limits_{x\in\mathbb{R}^3} & x_1^3 + x_2^3 + x_3^3 - x_1^2 x_2 - x_1 x_2^2 - x_1^2 x_3 - x_1 x_3^2 \\ & \qquad\qquad -x_2^2 x_3 - x_2 x_3^2 + 3x_1 x_2 x_3 \\ s.t. & x \in K_1 \cup K_2 \cup K_3, \end{cases}$$

where

$$K_1 = \{x \in \mathbb{R}^3 : x_1 \geq 0, x_1^2 + x_2^2 + x_3^2 = 1\},$$

$$K_2 = \{x \in \mathbb{R}^3 : x_2 \geq 0, x_1^2 + x_2^2 + x_3^2 = 1\},$$

$$K_3 = \{x \in \mathbb{R}^3 : x_3 \geq 0, x_1^2 + x_2^2 + x_3^2 = 1\}.$$

The objective function is obtained from Robinson's form [40] by changing x_i^2 to x_i for each i. For $k = 2$, we get $f_{mom,2} = -1.3185$, and the flat truncation (3.2.1) is met for all $l \in \mathcal{A} = \{1, 2, 3\}$. So, $f_{mom,2} = f_{min}$. The obtained three minimizers are

$$(0.2783, 0.2783, -0.9193) \in K_1 \cap K_2, \quad (0.2783, -0.9193, 0.2783) \in K_1 \cap K_3,$$

$$(-0.9193, 0.2783, 0.2783) \in K_2 \cap K_3.$$

For $k = 2$, the unified moment relaxation (3.1.2) took around 0.6 second, while solving the individual moment relaxations (1.2.1) for all K_1, K_2, K_3 took around 1.1 seconds.

Example 5.1.3. Consider the constrained optimization problem

$$\begin{cases} \min\limits_{x\in\mathbb{R}^3} & x_1 x_2 x_3 + x_1^2 x_2^2 (x_1^2 + x_2^2) + x_3^6 - 3x_1^2 x_2^2 x_3^2 \\ s.t. & x \in K_1 \cup K_2 \cup K_3, \end{cases}$$

where

$$K_1 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0, x_2 x_3 \geq 0\},$$

$$K_2 = \{x \in \mathbb{R}^3 : x_1^2 + x_3^2 - x_2^2 = 0, x_1 x_3 \geq 0\},$$

$$K_3 = \{x \in \mathbb{R}^3 : x_2^2 + x_3^2 - x_1^2 = 0, x_1 x_2 \geq 0\}.$$

The objective function is obtained from Motzkin's form [40] by adding the term $x_1 x_2 x_3$. For $k = 3$, we get $f_{mom,3} = -1.0757$, and the flat truncation (3.2.1) is met for all $l \in \mathcal{A} = \{2, 3\}$. So, $f_{mom,3} = f_{min}$. The obtained four minimizers are

$$(-1.0287, -1.6390, -1.2760) \in K_2, \quad (1.0287, -1.6390, 1.2760) \in K_2,$$

$$(-1.6390, -1.0287, -1.2760) \in K_3, \quad (1.6390, 1.0287, -1.2760) \in K_3.$$

For $k = 3$, the unified moment relaxation (3.1.2) took around 0.7 second, while solving the individual moment relaxations (1.2.1) for all K_1, K_2, K_3 took around 1.2 seconds.

Example 5.1.4. Consider the constrained optimization problem

$$\begin{cases} \min\limits_{x \in \mathbb{R}^3} & x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 + 4x_1 x_2 x_3 \\ s.t. & x \in K_1 \cup K_2 \cup K_3, \end{cases}$$

where

$$K_1 = \{x \in \mathbb{R}^3 : x_1 = x_2^2, x_3 = x_2^2\},$$

$$K_2 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 \leq 4, x_1 x_2 = -x_3, x_1 x_3 \leq 0\},$$

$$K_3 = \{x \in \mathbb{R}^3 : -1 \leq x_1 \leq 0, -1 \leq x_2 \leq 0, -1 \leq x_3 \leq 0\}.$$

The objective function is a dehomogenization of the Choi-Lam form [40]. For $k = 2$, we get $f_{mom,2} = -1$, and the flat truncation (3.2.1) is met for all $l \in \mathcal{A} = \{1, 2, 3\}$. So, $f_{mom,2} = f_{min}$. The obtained four minimizers are

$$(1, -1, 1) \in K_1, \quad (-1, 1, 1) \in K_2, \quad (1, 1, -1) \in K_2, \quad (-1, -1, -1) \in K_3.$$

For $k = 2$, the unified moment relaxation (3.1.2) took around 0.6 second, while solving the individual moment relaxations (1.2.1) for all K_1, K_2, K_3 took around 1.1 seconds.

A class of problems like (1.1.1) has absolute values in the constraints. For example, we consider that

$$K = \left\{ x : h(x) + \sum_{i=1}^{\ell} |g_i(x)| \geq 0 \right\}.$$

We can equivalently express K as

$$K = \bigcup_{s_1, \dots, s_l = \pm 1} \left\{ x : h(x) + \sum_{i=1}^{\ell} s_i \cdot g_i(x) \geq 0, s_i \cdot g_i(x) \geq 0 \right\}. \qquad (5.1.1)$$

Example 5.1.5. Consider the constrained optimization problem

$$\begin{cases} \min\limits_{x \in \mathbb{R}^2} & x_1^4 + x_2^4 - x_1^2 x_2^2 - 2x_1^2 - 3x_2^2 \\ s.t. & |x_1|^3 + |x_2|^3 \geq 4. \end{cases}$$

The constraining set can be equivalently expressed as $\bigcup\limits_{l=1}^{4} K_l$ with

$$K_1 = \{x \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1^3 + x_2^3 \geq 4\},$$

$$K_2 = \{x \in \mathbb{R}^2 : x_1 \geq 0, -x_2 \geq 0, x_1^3 - x_2^3 \geq 4\},$$

$$K_3 = \{x \in \mathbb{R}^2 : -x_1 \geq 0, x_2 \geq 0, -x_1^3 + x_2^3 \geq 4\},$$

$$K_4 = \{x \in \mathbb{R}^2 : -x_1 \geq 0, -x_2 \geq 0, -x_1^3 - x_2^3 \geq 4\}.$$

A contour of the objective over the feasible set is in Figure 5.1. For $k = 2$, we get $f_{mom,2} = -6.3333$, and the flat truncation (3.2.1) is met for all $l \in \mathcal{A} = \{1, 2, 3, 4\}$. So, $f_{mom,2} = f_{min}$. The obtained four minimizers are

$$(1.5275, 1.6330) \in K_1, \quad (1.5275, -1.6330) \in K_2,$$

$$(-1.5275, 1.6330) \in K_3, \quad (-1.5275, -1.6330) \in K_4.$$

For $k = 2$, the unified moment relaxation (3.1.2) took around 0.6 second, while solving the individual moment relaxations (1.2.1) for all K_1, K_2, K_3, K_4 took around 0.8 second.
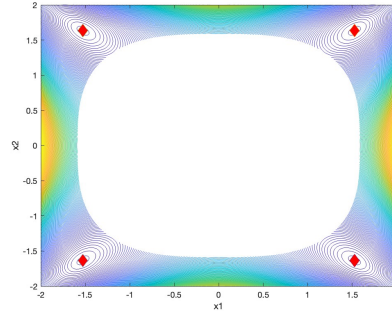

Figure 5.1: The contour is for the objective function in Example 5.1.5. The region outside the oval is the feasible set. The four diamonds are the minimizers.

5.2   An application

Now we show how to compute the (p, q)-norm of a matrix $A \in \mathbb{R}^{m \times n}$ for positive integers p, q. Recall that

$$\|A\|_{p,q} := \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_q} = \max_{\|x\|_q = 1} \|Ax\|_p.$$

When p and q are both even integers, this is a standard polynomial optimization problem. If one of them is odd, then we need to get rid of the absolute value constraints. When p is even and q is odd, we can equivalently express that

$$\left\{ \begin{array}{rl} (\|A\|_{p,q})^p = & \max \quad (a_1^T x)^p + \cdots + (a_m^T x)^p \\ & s.t. \quad |x_1|^q + \cdots + |x_n|^q = 1. \end{array} \right. \tag{5.2.1}$$

Here, the a_i^T is the ith row of A. When p is odd and q is even, we have

$$\left\{ \begin{array}{rl} \|A\|_{p,q} = & \max \quad x_{n+1} \\ & s.t. \quad (x_1)^q + \cdots + (x_n)^q = 1, \\ & \quad\;\; |a_1^T x|^p + \cdots + |a_m^T x|^p = (x_{n+1})^p. \end{array} \right. \tag{5.2.2}$$

Similarly, when p and q are both odd, we have

$$\begin{cases} \|A\|_{p,q} = & \max \quad x_{n+1} \\ & s.t. \quad |x_1|^q + \cdots + |x_n|^q = 1, \\ & \quad\quad |a_1^T x|^p + \cdots + |a_m^T x|^p = (x_{n+1})^p. \end{cases} \tag{5.2.3}$$

The constraining sets in the above optimization problems can be decomposed in the same way as in (5.1.1).

Example 5.2.1. Consider the following matrix

$$A = \begin{bmatrix} -8 & -8 & -3 & 1 \\ 4 & -7 & 7 & 6 \\ 6 & -7 & -7 & -4 \\ 8 & 0 & -9 & -6 \end{bmatrix}.$$

Some typical values of the norm $\|A\|_{p,q}$ and the vector x^* for achieving it are listed in Table 5.1. The norms $\|A\|_{p,q}$ are all computed successfully by the

Table 5.1: The (p,q)-norms for the matrix A in Example 5.2.1.

(p,q)	$\|A\|_{p,q}$	x^* for $\|A\|_{p,q} = \|Ax^*\|_p, \|x^*\|_q = 1$
$(2,3)$	21.6132	$(0.6568, -0.3937, -0.7542, -0.6097)$
$(3,2)$	15.5469	$(0.5606, -0.2097, -0.6742, -0.4327)$
$(3,3)$	19.0928	$(0.6782, -0.4598, -0.7329, -0.5820)$
$(3,4)$	21.2617	$(0.7446, -0.5841, -0.7824, -0.6700)$
$(4,3)$	18.0128	$(0.6825, -0.4605, -0.7305, -0.5794)$
$(4,4)$	20.0605	$(0.7465, -0.5863, -0.7809, -0.6682)$
$(4,5)$	21.4196	$(0.7895, -0.6633, -0.8166, -0.7261)$
$(5,4)$	19.3770	$(0.7471, -0.5848, -0.7810, -0.6683)$
$(5,5)$	20.6894	$(0.7896, -0.6635, -0.8165, -0.7260)$

unified moment relaxation (3.1.2) for the relaxation order $k = 2$ or 3.

Chapter 6

Conclusions and Future Work

This paper proposes a unified Moment-SOS hierarchy for solving the polynomial optimization problem (1.1.1) whose feasible set K is a union of several basic closed semialgebraic sets K_l. Instead of minimizing the objective f separately over each individual set K_l, we give a unified hierarchy of Moment-SOS relaxations to solve (1.1.1). This hierarchy produces a sequence of lower bounds for the optimal value f_{min} of (1.1.1). When the archimedeanness is met for each constraining subset K_l, we show the asymptotic convergence of this unified hierarchy. Furthermore, if the linear independence constraint qualification, the strict complementarity and the second order sufficient conditions hold at every global minimizer for each K_l, we prove the finite convergence of the hierarchy. For the univariate case, special properties of the corresponding Moment-SOS relaxation are discussed. To the best of the authors' knowledge, this is the first unified hierarchy of Moment-SOS relaxations for solving polynomial optimization over unions of sets. Moreover, numerical experiments are provided to demonstrate the efficiency of this method. In particular, as applications, we show how to compute the (p, q)-norm of a matrix for positive integers p, q.

There exists relevant work on approximation and optimization about measures with unions of several individual sets. For instance, Korda et al. [11] considers the generalized moment problem (GMP) that exploits the ideal sparsity, where the feasible set is a basic closed semialgebraic set containing conditions like $x_i x_j = 0$. Because of this, the moment relaxation for solving the GMP involves several measures, each supported in an individual set. Lasserre et al. [15] proposes the multi-measure approach to approximate the moments of Lebesgue measures supported in unions of basic semialgebraic sets. Magron et al. [20] discusses the union problem in the context of piecewise polynomial systems. We would also like to compare the sizes of relaxations (1.2.1) and (1.2.2). To apply the individual relaxation (1.2.1), we need to solve it for m times. For the unified relaxation (1.2.2), we only need to solve it for one time. For a fixed relaxation

order k in (1.2.1), the length of the vector $y^{(l)}$ is $\binom{n+2k}{2k}$. For the same k in (1.2.2), there are m vectors of $y^{(l)}$, and each of them has length $\binom{n+2k}{2k}$. The comparison of the numbers of constraints is similar. Observe that (1.2.1) has $|\mathcal{E}^{(l)}|$ equality constraints, $|\mathcal{I}^{(l)}| + 1$ linear matrix inequality constraints, and one scalar equality constraint. Similarly, (1.2.2) has $|\mathcal{E}^{(1)}| + \cdots + |\mathcal{E}^{(m)}|$ equality constraints, $|\mathcal{I}^{(1)}| + \cdots + |\mathcal{I}^{(m)}| + m$ linear matrix inequality constraints, and one scalar equality constraint. It is not clear which approach is more computationally efficient. However, in our numerical examples, solving (1.2.2) is relatively faster.

There is much interesting future work to do. For instance, when the number of individual sets is large, the unified Moment-SOS relaxations have a large number of variables. How to solve the moment relaxation (3.1.2) efficiently is important in applications. For large scale problems, some sparsity patterns can be exploited. We refer to [10, 21, 42, 43] for related work. It is interesting future work to explore the sparsity for unified Moment-SOS relaxations. Moreover, how to solve polynomial optimization over a union of infinitely many sets is another interesting future work.

Bibliography

[1] BERTSEKAS, D.: *Nonlinear programming*, second edition, Athena Scientific, 1995.

[2] BOCHNAK, J., COSTE, M., ROY, M.F.: *Real algebraic geometry*, vol. 36, Springer Science & Business Media, 2013.

[3] CURTO, R., FIALKOW, L.: *Truncated K-moment problems in several variables*, J. Oper. Theory, 54, 189–226, 2005.

[4] DE KLERK, E.: *Aspects of semidefinite programming : interior point algorithms and selected applications*, Springer, 2002.

[5] DRESSLER, M., NIE, J., YANG, Z.: *Separability of Hermitian tensors and PSD decompositions*, Linear and Multilinear Algebra 70 (21), 6581–6608, 2022.

[6] HENRION, D., LASSERRE, J.B.: *Detecting global optimality and extracting solutions in gloptipoly*, in Positive polynomials in control, Springer, 2005, pp. 293–310.

[7] HENRION, D., LASSERRE, J.B., LÖFBERG, J.: *Gloptipoly 3: moments, optimization and semidefinite programming*, Optimization Methods & Software, 24 (2009), pp. 761–779.

[8] HENRION, D., KORDA, M., LASSERRE, J.B.: *The Moment-SOS Hierarchy*, World Scientific, Singapore, 2020.

[9] HENRION, D.: *Geometry of exactness of moment-SOS relaxations for polynomial optimization*, Preprint, 2023. `arXiv:2310.17229`

[10] KLEP, I., MAGRON, V., POVH, J.: *Sparse noncommutative polynomial optimization*, Mathematical Programming, pp.1-41, 2022.

[11] KORDA, M., LAURENT, M., MAGRON, V., STEENKAMP, A.: *Exploiting ideal-sparsity in the generalized moment problem with application to matrix factorization ranks*, Mathematical Programming, 2023.

[12] LASSERRE, J.B.: *Global optimization with polynomials and the problem of moments*, SIAM Journal on optimization, 11 (2001), pp. 796–817.

[13] LASSERRE, J.B.: *Introduction to polynomial and semi-algebraic optimization*, Cambridge University Press, Cambridge, 2015.

[14] LASSERRE, J.B., LAURENT, M., ROSTALSKI, P.: *Semidefinite characterization and computation of zero-dimensional real radical ideals*, Found. Comput. Math., 8 (2008), 607–647.

[15] LASSERRE, J.B., EMIN, Y.: *Semidefinite relaxations for Lebesgue and Gaussian measures of unions of basic semialgebraic sets*, Mathematics of Operations Research, 44(4):1477–1493, 2019.

[16] LAURENT, M.: *Semidefinite representations for finite varieties*, Math. Program., 109 (2007), 1–26.

[17] LAURENT, M.: *Sums of squares, moment matrices and optimization over polynomials*, Emerging Applications of Algebraic Geometry of IMA Volumes in Mathematics and its Applications, Vol. 149, pp. 157–270, Springer, 2009.

[18] LAURENT, M.: *Optimization over polynomials: Selected topics*, Proceedings of the International Congress of Mathematicians (S.Y. Jang, Y.R. Kim, D.-W. Lee, and I. Yie, eds.), pp. 843–869, 2014.

[19] LUKÁCS, F.: *Verschärfung des ersten mittelwertsatzes der integralrechnung für rationale polynome*, Mathematische Zeitschrift, 2 (1918), pp. 295–305.

[20] MAGRON, V., FORETS, M., HENRION, D.: *Semidefinite approximations of invariant measures for polynomial systems*, Discrete & Continuous Dynamical Systems - B, 24(12):6745–6770, 2019.

[21] MAI, N.H.A., MAGRON, V., LASSERRE, J.B.: *A sparse version of Reznick's Positivstellensatz*, Mathematics of Operations Research 48 (2), 812-833, 2023.

[22] MARKOV, A.: *Lectures on functions of minimal deviation from zero (russian), 1906*, Selected Works: Continued fractions and the theory of functions deviating least from zero, OGIZ, Moscow-Leningrad, (1948), pp. 244–291.

[23] NESTEROV, Y.: *Squared Functional Systems and Optimization Problems*, High performance optimization, pp. 405-440, 2000.

[24] NIE, J.: *Polynomial optimization with real varieties*, SIAM Journal on Optimization, 23 (2013), pp. 1634–1646.

[25] NIE, J.: *Optimality conditions and finite convergence of Lasserre's hierarchy*, Math. Program., 146(1-2), 97–121, 2014.

[26] NIE, J.: *Symmetric tensor nuclear norms*, SIAM J. Appl. Algebra Geometry, 1(1), 599–625, 2017.

[27] NIE, J.: *Tight relaxations for polynomial optimization and lagrange multiplier expressions*, Mathematical Programming, 178 (2019), pp. 1–37.

[28] NIE, J.: *Moment and polynomial optimization*, SIAM, 2023.

[29] NIE, J., DEMMEL, J.: *Shape optimization of transfer functions*, Multiscale optimization methods and applications, 313–326, 2006.

[30] NIE, J., TANG, X.: *Nash equilibrium problems of polynomials*, Mathematics of Operations Research, to appear, 2023.

[31] NIE, J., TANG, X.: *Convex generalized Nash equilibrium problems and polynomial optimization*, Mathematical Programming 198 (2), 1485–1518, 2023.

[32] NIE, J., TANG, X., YANG, Z., ZHONG, S.: *Dehomogenization for completely positive tensors*, Numerical Algebra, Control and Optimization 13 (2), 340-363, 2023.

[33] NIE, J., TANG, X., ZHONG, S.: *Rational generalized Nash equilibrium problems*, SIAM Journal on Optimization 33 (3), 1587–1620, 2023.

[34] NIE, J., YANG, L., ZHONG, S.: *Stochastic polynomial optimization*, Optimization Methods and Software 35 (2), 329-347, 2020.

[35] NIE, J., YANG, L., ZHONG, S., ZHOU, G.: *Distributionally robust optimization with moment ambiguity sets*, Journal of Scientific Computing 94 (1), 12, 2023.

[36] NIE, J., ZHANG, X.: *Real eigenvalues of nonsymmetric tensors*, Comp. Opt. and Appl., 70(1), 1–32, 2018.

[37] PÓLYA, G., SZEGÖ, G., ET AL.: *Problems and Theorems in Analysis: Series, integral calculus, theory of functions*, Springer, 1972.

[38] POWERS, V., REZNICK, B.: *Polynomials that are positive on an interval*, Transactions of the American Mathematical Society, 352 (2000), pp. 4677–4692.

[39] PUTINAR, M.: *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal, 42 (1993), pp. 969–984.

[40] REZNICK, B.: *Some concrete aspects of Hilbert's 17th problem*, Contemp. Math., 253, 251–272, AMS, 2000.

[41] STURM, J.F.: *Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones*, Optimization methods and software, 11 (1999), pp. 625–653.

[42] WANG, J., MAGRON, V., LASSERRE, J.B.: *TSSOS: A Moment-SOS Hierarchy That Exploits Term Sparsity*, SIAM Journal on Optimization 31 (1), 30-58, 2021.

[43] WANG, J., MAGRON, V., LASSERRE, J.B., MAI, N.H.A.: *CS-TSSOS:*
      Correlative and term sparsity for large-scale polynomial optimization,
      ACM Transactions on Mathematical Software, 48(4), pp.1-26, 2022.

[44] ZHONG, S., CUI, Y., NIE, J.: *Towards Global Solutions for Noncon-*
      vex Two-Stage Stochastic Programs: A Polynomial Lower Approximation
      Approach, Preprint, 2023. `arXiv:2310.04243`