2017

Conic Programming Approaches for Polynomial Optimization: Theory and Applications

Xiaolong Kuang
Lehigh University

Conic Programming Approaches for Polynomial Optimization: Theory and Applications

by

Xiaolong Kuang

Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy
in
Industrial and Systems Engineering

Lehigh University
August 2017

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Date

____
Dissertation Advisor

Committee Members:

____
Luis F. Zuluaga, Committee Chair

____
Tamás Terlaky

____
Katya Scheinberg

____
Samuel A. Burer

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Luis F. Zuluaga for the continuous support of my Ph.D study and related research. His proactive attitude towards research shapes my understanding on how to conduct independent research and be a good PhD student. I appreciate that Dr. Zuluaga gives me the freedom to search and re-search the topics that truly interest me, which improves my critical thinking and independent research ability. In the meanwhile, I want to thank him for referring me to the internship position at IBM Research, giving me the opportunities to do research with top scientists and researchers, which broadens my horizon and brings up a lot of research opportunities to work on for my PhD dissertation. His guidance in research as well as many other aspects in life are priceless, and Dr. Zuluaga is more than just a PhD advisor but a lifelong friend and mentor to me.

Besides my advisor, I would like to thank the rest of the committee members, Prof. Tamás Terlaky, Prof. Katya Scheinberg and Prof. Samuel A. Burer, for their participation in my PhD proposal, general exam and dissertation defense presentation and their insightful comments on this dissertation. Their encouragement and advice drive me to dive deeper in my research. I would also like to thank Dr. Bissan Ghaddar and Dr. Joe Naoum-Sawaya for their guidance and help when I was interning with them at IBM Research in Ireland. It was a really enjoyable experience to work with them during the internship and the continuing collaboration in the following years. In the meanwhile, I would like to thank Prof. Alberto J. Lamadrid in Economics Department to help me in applying optimization

theory to solve real-life problems in economics.

Finally, I want to thank my parents and my dear friends in supporting me to finish my PhD study. There were a lot of ups and downs for my PhD life in the last five years. My parents always kept me company when I felt frustrated and at the same time I shared a lot of happy moments in my PhD with them. My friends make my life colorful and we encourage each other to move forward. Without them, I couldn't image how the life would be.

Contents

List of Tables

List of Figures

Abstract

Historically, polynomials are among the most popular class of functions used for empirical modeling in science and engineering. Polynomials are easy to evaluate, appear naturally in many physical (real-world) systems, and can be used to accurately approximate any smooth function. It is not surprising then, that the task of solving polynomial optimization problems; that is, problems where both the objective function and constraints are multivariate polynomials, is ubiquitous and of enormous interest in these fields. Clearly, polynomial optimization problems encompass a very general class of non-convex optimization problems, including key combinatorial optimization problems.

The focus of the first three chapters of this document is to address the solution of polynomial optimization problems in theory and in practice, using a conic optimization approach. Convex optimization has been well studied to solve quadratic constrained quadratic problems. In the first part, convex relaxations for general polynomial optimization problems are discussed. Instead of using the matrix space to study quadratic programs, we study the convex relaxations for POPs through a lifted *tensor* space, more specifically, using the completely positive tensor cone and the completely positive semidefinite tensor cone. We show that tensor relaxations theoretically yield no-worse global bounds for a class of polynomial optimization problems than relaxation for a QCQP reformulation of the POPs. We also propose an approximation strategy for tensor cones and show empirically the advantage of the tensor relaxation.

In the second part, we propose an alternative SDP and SOCP hierarchy to obtain global

bounds for general polynomial optimization problems. Comparing with other existing SDP and SOCP hierarchies that uses higher degree sum of square (SOS) polynomials and scaled diagonally sum of square polynomials (SDSOS) when the hierarchy level increases, these proposed hierarchies, using fixed degree SOS and SDSOS polynomials but more of these polynomials, perform numerically better. Numerical results show that the hierarchies we proposed have better performance in terms of tightness of the bound and solution time compared with other hierarchies in the literature.

The third chapter deals with Alternating Current Optimal Power Flow problem via a polynomial optimization approach. The Alternating Current Optimal Power Flow (ACOPF) problem is a challenging non-convex optimization problem in power systems. Prior research mainly focuses on using SDP relaxations and SDP-based hierarchies to address the solution of ACOPF problem. In this Chapter, we apply existing SOCP hierarchies to this problem and explore the structure of the network to propose simplified hierarchies for ACOPF problems. Compared with SDP approaches, SOCP approaches are easier to solve and can be used to approximate large scale ACOPF problems.

The last chapter also relates to the use of conic optimization techniques, but in this case to pricing in markets with non-convexities. Indeed, it is an application of conic optimization approach to solve a pricing problem in energy systems. Prior research in energy market pricing mainly focus on linear costs in the objective function. Due to the penetration of renewable energies into the current electricity grid, it is important to consider quadratic costs in the objective function, which reflects the ramping costs for traditional generators. This study address the issue how to find the market clearing prices when considering quadratic costs in the objective function.

Chapter 1

Convex Relaxations for Polynomial Optimization

1.1 Introduction

Polynomials appear in a wide variety of areas in science. It is not surprising then that polynomial optimization has recently been a very active field of research [cf., 5]. Here, the interest is the class of non-convex, non-linear POPs. Clearly, a non-convex quadratic program (QP) belongs to this class of problems, and its study has been widely addressed in the literature. For example, to address the solution of QPs, semidefinite programming (SDP) [cf., 113] relaxations have been actively used to find good bounds and approximate solutions for general [see, e.g. 30, 80, 117] and important instances of this problem such as the max-cut problem and the stable set problem (see e.g., [32, 33, 47, 90]). In [58], less computationally expensive second order cone programming (SOCP) [cf., 4] relaxations have also been proposed to approximate non-convex QPs.

The early work linking convex optimization and polynomial optimization in [86, 102] reveals the possibility to use conic optimization to obtain global or near-global solutions for non-convex POPs in which higher than second-order polynomials are used. In the seminal work of Parrilo [89] and Lasserre [66], SDP is used to obtain the global or near-global optimum for POPs. Besides SDP approximations, other convex approximations to address the solution of POPs have been investigated using linear programming (LP) and SOCP techniques [2, 67, 68, 91, 121]. These techniques are at the core of the well-known area of Polynomial Optimization [cf., 5].

Alternatively, it has been shown that several NP-hard optimization problems can be expressed as linear programs over the convex cone of copositive matrices and its dual cone, the cone of completely positive matrices, including standard quadratic problems [20], stable set problems [32, 39], graph partitioning problems [93], and quadratic assignment problems [94]. In [27], Burer shows the much more general result that every quadratic problem with linear and binary constraints can be rewritten as such a problem. Completely positive relaxations for general quadratically constrained quadratic programs (QCQPs) have been studied in [8, 29]. In [11], CP reformulation for QCQPs and quadratic program with complementarity constraints (QPCCs) are discussed without any assumptions on the feasible regions. Although copositive/completely positive cones are not tractable in general, recent advances on obtaining algorithms ([3, 26, 35], etc.) to approximate copositive/completely positive cones provide an alternative way to globally solve quadratic POPs. Recently, Bomze shows in [18] that copositive relaxation provides stronger bounds than Lagrangian dual bounds in quadratically and linearly constrained QPs.

A natural thought is whether one can extend the copositive programming or completely positive programming reformulations for QPs to POPs. In [9], Arima et al. proposed the moment cone relaxation for a class of polynomial optimization problems (POPs) to extend the results on the completely positive cone programming relaxation for the quadratic optimization. Recently, Pena et al. show in [92] that under certain conditions general POPs can be reformulated as a conic program over the cone of completely positive tensors, which is a natural extension of the cone of completely positive matrices in quadratic problems. This tensor representation was originally proposed in [38], and is now the focus of active research [see, e.g., 53, 56, 78, 106]. In [92], it is also shown that the conditions for the equivalence of POPs and the completely positive conic programs, when applied to QPs, lead to conditions that are weaker than the ones introduced in [27].

In [64], we study completely positive (CP) and completely positive semidefinite (CPSD) tensor relaxations for POPs. Our main contributions are: 1) We extend the results for QPs in [18] to general POPs by using CP and CPSD tensor cones. In particular, we show that CP tensor relaxations provide tighter bounds than Lagrangian relaxations for general POPs. 2) We provide tractable approximations for CP and CPSD tensor cones that can be used to globally approximate general POPs. 3) We prove that CP and CPSD tensor relaxations yield tighter bounds than completely positive and positive semidefinite matrix relaxations for quadratic reformulations of some classes of POPs. 4) We provide preliminary numerical results on more general cases of POPs and show that the approximation of CP tensor cone programs can yield tighter bounds than relaxations based on doubly nonnegative (DNN) matrices [cf., 14] for completely positive matrix relaxation to the reformulated quadratic programs.

The remainder of the chapter is organized as follows. We briefly introduce the basic concepts of tensor cone and tensor representation of polynomials in Section 1.2. Lagrangian, completely positive semidefinite tensor, and completely positive tensor relaxations for PO problems are discussed in Section 1.3. In Section 1.4, we discuss a *quadratic approach* to general POPs; that is, when auxiliary decision variables are introduced to the problem to reformulate it as a Quadratically Constrained Quadratic Program (QCQP). Then, the completely positive relaxations is applied to the resulting QCQPs and the bounds are compared with those obtained from the tensor relaxations for a class of POPs. In Section 1.5, Linear Matrix Inequality (LMI) approximation strategies for the positive semidefinite and completely positive tensor cones are developed and a comparison of tensor relaxations with matrix relaxations obtaining using the quadratic approach is done by obtaining numerical results on several POPs. Lastly, Section 1.6 summarizes the results and provides future working directions.

1.2 Preliminaries

1.2.1 Basic Concepts and Notations

We first introduce basic concepts and notations used throughout the chapter. Following [92], we start by defining tensors.

Definition 1. *Let $\mathcal{T}_{n,d}$ denote the set of* tensors *of dimension n and order d in $\mathbb{R}^n$, that is*

$$\mathcal{T}_{n,d} = \underbrace{\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n}_{d},$$

where $\otimes$ is the tensor product.

A tensor $T \subseteq \mathcal{T}_{n,d}$ is symmetric if the entries are independent of the permutation of its indices. We denote $\mathcal{S}_{n,d} \subseteq \mathcal{T}_{n,d}$ as the set of symmetric tensors of dimension n and order d. For any $T^1, T^2 \in \mathcal{T}_{n,d}$, let $\langle \cdot, \cdot \rangle_{n,d}$ denote the tensor inner product defined by

$$\langle T^1, T^2 \rangle_{n,d} = \sum_{\{i_1,\dots,i_d\} \in \{1,\dots,n\}^d} T^1_{(i_1,\dots,i_d)} T^2_{(i_1,\dots,i_d)}.$$

Definition 2. *For any $x \in \mathbb{R}^n$, let the mapping $\mathbb{R}^n \to \mathcal{S}_{n,d}$ be defined by*

$$M_d(x) = \underbrace{x \otimes \cdots \otimes x}_{d}.$$

Definition 1 and 2 are natural extensions of matrix notations to higher order. For example, $\mathcal{T}_{n,2}$ is the set $n \times n$ matrices, while $\mathcal{S}_{n,2}$ is the set of $n \times n$ symmetric matrices, $\langle \cdot, \cdot \rangle_{n,2}$ is the Frobenius inner product and $M_2(x) = xx^T$ for any $x \in \mathbb{R}^n$. In general, $M_d(x)$ is the symmetric tensor whose $(i_1, ..., i_d)$ entry is $x_{i_1} \cdots x_{i_d}$.

Proposition 1. *Let $\mathbb{E}_{n,d}$ be all 1 tensor with dimension n and order d and $e \in \mathbb{R}^n$ be the all one vector, then*

$$\langle \mathbb{E}_{n,d}, M_d(x) \rangle_{n,d} = (e^T x)^d, \forall x \in \mathbb{R}^n.$$

Proof. By the definition of $M_d(\cdot)$ and $\langle \cdot, \cdot \rangle_{n,d}$,

$$\langle \mathbb{E}_{n,d}, M_d(x) \rangle_{n,d} = \sum_{k_1+k_2+\cdots+k_n=d} \binom{d}{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} = (e^T x)^d,$$

where $\binom{d}{k_1,k_2,\ldots,k_n}$ is the multinomial coefficient. □

Proposition 2. *For* $x \in \mathbb{R}^n, y \in \mathbb{R}^n$,

$$\langle M_d(x), M_d(y) \rangle_{n,d} = (x^T y)^d.$$

Proof. Let $x, y \in \mathbb{R}^n$ be given and $z \in \mathbb{R}^n$ be defined as $z_i = x_i y_i, i = 1, ..., n$, and let $e \in \mathbb{R}^n$ be the all one vector, from the definition of $M_d(\cdot)$ and $\langle \cdot, \cdot \rangle_{n,d}$,

$$\begin{aligned}
\langle M_d(x), M_d(y) \rangle_{n,d} &= \sum_{\{i_1,\ldots,i_d\} \in \{1,\ldots,n\}^d} M_d(x)_{(i_1,\ldots,i_d)} M_d(y)_{(i_1,\ldots,i_d)} \\
&= \sum_{\{i_1,\ldots,i_d\} \in \{1,\ldots,n\}^d} x_{i_1} x_{i_2} \cdots x_{i_d} \cdot y_{i_1} y_{i_2} \cdots y_{i_d} \\
&= \sum_{\{i_1,\ldots,i_d\} \in \{1,\ldots,n\}^d} (x_{i_1} y_{i_1})(x_{i_2} y_{i_2}) \cdots (x_{i_d} y_{i_d}) \\
&= \langle \mathbb{E}_{n,d}, M_d(z) \rangle_{n,d} \\
&= (e^T z)^d \qquad \text{(from Proposition 1)} \\
&= (x^T y)^d.
\end{aligned}$$

□

Analogous to positive semidefinite and copositive matrices of order 2, positive semidefinite and copositive tensors can be defined as follows.

Definition 3. *Define the* $\mathcal{K}$-semidefinite (or set-semidefinite) symmetric tensor cone *of dimension* n *and order* d *as:*

$$\mathcal{C}_{n,d}(\mathcal{K}) = \{T \in \mathcal{S}_{n,d} : \langle T, M_d(x) \rangle_{n,d} \geq 0, \forall x \in \mathcal{K}\}.$$

For $\mathcal{K} = \mathbb{R}^n$, $\mathcal{C}_{n,d}(\mathbb{R}^n)$ *denotes the* positive semidefinite (PSD) tensor cone. *For* $\mathcal{K} = \mathbb{R}^n_+$, $\mathcal{C}_{n,d}(\mathbb{R}^n_+)$ *denotes the* copositive tensor cone.

Similar to the one-to-one correspondence of $n \times n$ PSD matrices to nonnegative homogeneous quadratic polynomials of n variables, there is also a one-to-one correspondence of PSD tensors with dimension n and order d to nonnegative homogeneous polynomials with n variables and degree d [cf., 78]. Note that there is no nonnegative homogeneous polynomial with odd degree. Thus it follows that there is no PSD tensor with odd order. Next we discuss the *dual cones* of $\mathcal{C}_{n,d}(\mathbb{R}^n_+)$ and $\mathcal{C}_{n,d}(\mathbb{R}^n)$, following the discussion in [78] and [92].

Definition 4. *Given any cone* $\mathcal{C}$ *of symmetric tensors, the* dual cone *of* $\mathcal{C}$ *is*

$$\mathcal{C}^* = \{Y \in \mathcal{S}_{n,d} : \langle X, Y \rangle \geq 0, \forall X \in \mathcal{C}\},$$

and if $\mathcal{C}^* = \mathcal{C}$, *then cone* $\mathcal{C}$ *is* self-dual.

The dual cones of the positive semidefinite tensor cone and copositive tensor cone have been studied in [78] and [92]. More formally,

Proposition 3.

(a) $\mathcal{C}^*_{n,d}(\mathbb{R}^n_+) = \operatorname{conv}\{M_d(x) : x \in \mathbb{R}^n_+\}$.

(b) $\mathcal{C}^*_{n,2d}(\mathbb{R}^n) = \operatorname{conv}\{M_{2d}(x) : x \in \mathbb{R}^n\}$.

Similar to the completely positive matrix cone $\mathcal{C}^*_{n,2}(\mathbb{R}^n_+)$, we call $\mathcal{C}^*_{n,d}(\mathbb{R}^n_+)$ the *completely positive (CP) tensor cone*. It is well known that the positive semidefinite matrix cone is self-dual, however, in general, the positive semidefinite tensor cone is not self-dual [cf., 78]. Thus, here we name $\mathcal{C}^*_{n,2d}(\mathbb{R}^n)$ as the *completely positive semidefinite (CPSD) tensor cone*. Before formally stating that $\mathcal{C}^*_{n,2d}(\mathbb{R}^n) \neq \mathcal{C}_{n,2d}(\mathbb{R}^n)$ in general, we first introduce the *homogeneous sum of square (SOS) tensor cone* of dimension d and order $2d$ as

$$\mathcal{C}_{n,2d}(\mathcal{SOS}) = \{T_{n,2d} : \langle T_{n,2d}, M_{2d}(x) \rangle = \sum_i \lambda_i \left(\langle T^i_{n,d}, M_d(x) \rangle \right)^2, \ \textit{for some } \lambda_i \geq 0\}.$$

Similarly, there is a one-to-one corresponding relationship between homogeneous SOS tensors with dimension n and order $2d$ and homogeneous SOS polynomials with dimension n and degree $2d$. Next we discuss the relationships between nonnegative and sum of square polynomials from the perspective of tensor representation and reveal the relationship between SOS and CPSD tensors.

Proposition 4 ([78, Prop. 5.8 (i)])**.**

$$\mathcal{C}^*_{n,2d}(\mathbb{R}^n) \subseteq \mathcal{C}_{n,2d}(\mathcal{SOS}) \subseteq \mathcal{C}_{n,2d}(\mathbb{R}^n).$$

Proof. Let $T \in \mathcal{C}^*_{n,2d}(\mathbb{R}^n)$, by Proposition 3, $T = \sum_i \lambda_i M_{2d}(y^i), y^i \in \mathbb{R}^n, \lambda_i \geq 0, \sum_i \lambda_i = 1$. Then $\forall x \in \mathbb{R}^n$,

$$\begin{aligned}
\langle T, M_{2d}(x)\rangle_{n,2d} &= \left\langle \sum_i \lambda_i M_{2d}(y^i), M_{2d}(x) \right\rangle_{n,2d} \\
&= \sum_i \lambda_i \left\langle M_{2d}(y^i), M_{2d}(x) \right\rangle_{n,2d} \\
&= \sum_i \lambda_i (x^T y^i)^{2d} \qquad\qquad \text{(from Proposition 2)} \\
&= \sum_i \left[\sqrt{\lambda_i}(x^T y^i)^d\right]^2.
\end{aligned}$$

Take $z^i_k = x_k y^i_k$, then $x^T y^i = e^T z^i$ where $e \in \mathbb{R}^n$ is an all one vector. Therefore,

$$\begin{aligned}
\langle T, M_{2d}(x)\rangle_{n,2d} &= \sum_i \left[\sqrt{\lambda_i}(e^T z^i)^d\right]^2 \\
&= \sum_i \left[\sqrt{\lambda_i}\langle \mathbb{E}_{n,d}, M_d(z^i)\rangle_{n,d}\right]^2, \qquad \text{(from Proposition 1)}
\end{aligned}$$

therefore $\mathcal{C}^*_{n,2d}(\mathbb{R}^n) \subseteq \mathcal{C}_{n,2d}(\mathcal{SOS})$. By the definition of homogeneous SOS tensor cone, it is clear that $\mathcal{C}_{n,2d}(\mathcal{SOS}) \subseteq \mathcal{C}_{n,2d}(\mathbb{R}^n)$. □

The proof of Proposition 4 can be seen as an alternative proof for Proposition 5.8 (i) in [78] using the tensor notations introduced here. Well studied sum of square polynomial optimization reveals that a nonnegative multivariate homogeneous polynomial is a homogeneous sum of square polynomial if it is quadratic, that is $\mathcal{C}^*_{n,2}(\mathbb{R}^n) = \mathcal{C}_{n,2}(\mathcal{SOS}) = \mathcal{C}_{n,2}(\mathbb{R}^n)$. This statement coincides with the self-duality of the PSD matrix cone. Luo et al. showed in [78] that $\mathcal{C}^*_{n,2d}(\mathbb{R}^n) \subsetneq \mathcal{C}_{n,2d}(\mathcal{SOS})$ for $d \geq 2$. On the other hand, the Motzkin polynomial together with isomorphism between homogeneous polynomials and tensors shows that $\mathcal{C}_{n,2d}(\mathcal{SOS}) \subsetneq \mathcal{C}_{n,2d}(\mathbb{R}^n)$ when $d \geq 2$.

1.2.2 Tensor Representation of Polynomial Optimization

In section 1.2.1, we discussed that some homogeneous polynomials can be expressed as tensor inner product with $M_d(x)$. Next, we introduce a tensor representation for general polynomials that are not necessarily homogeneous. Define $\mathbb{R}[x]$ as the ring of polynomials with real coefficients in $\mathbb{R}^n$, and let $\mathbb{R}_d[x] := \{p \in \mathbb{R}[x] : \deg(p) \leq d\}$ denote the set of polynomials with dimension n and degree at most d. For simplicity, we use $M_d(1, x), x \in \mathbb{R}^n$ to represent $M_d((1, x^T)^T), x \in \mathbb{R}^n$ throughout this paper. For any $p(x) \in \mathbb{R}_d[x]$, we can write $p(x)$ as

$$p(x) = \langle T_d(p), M_d(1, x)\rangle_{n+1,d}, \tag{1.1}$$

where $T_d(\cdot)$ is the mapping of coefficients of $p(x)$ in terms of $M_d(1, x)$ in $\mathcal{S}_{n+1,d}$. Following [92], define $T_d : \mathbb{R}_d[x] \to \mathcal{S}_{n+1,d}$ as

$$T_d\left(\sum_{\beta\in\mathbb{Z}^n_+:|\beta|\leq d} p_\beta x^\beta\right)_{i_1,\ldots,i_d} := \frac{\alpha_1!\cdots\alpha_n!}{|\alpha|!}p_\alpha,$$

where α is the (unique) exponent such that $x^\alpha := x_1^{\alpha_1}\cdots x_n^{\alpha_n} = x_{i_1}\cdots x_{i_d}$ (i.e., α_k is the number of times k appears in the multi-set $\{i_1, \ldots, i_d\}$) and $|\alpha| = \sum_{i=1}^n \alpha_i$. For any polynomial $p(x) \in \mathbb{R}_d[x]$, let $\tilde{p}(x)$ denote the homogenous component of $p(x)$ with highest degree, then it follows

$$\tilde{p}(x) = \langle T_d(p), M_d(0, x)\rangle_{n+1,d}. \tag{1.2}$$

Equation (1.1) and (1.2) allow us to represent any multivariate polynomials with their tensor forms and provide the possibility to study the boundedness of general polynomials with their tensor representations.

Theorem 1. *Let $\mu \in \mathbb{R}$, we have*

(a) Let $p(x) \in \mathbb{R}_d[x]$. Then $p(x) \geq \mu$ for all $x \in \mathbb{R}^n_+$ if and only if $T_d(p-\mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)$.

(b) Let $p(x) \in \mathbb{R}_{2d}[x]$. Then $p(x) \geq \mu$ for all $x \in \mathbb{R}^n$ if and only if $T_{2d}(p-\mu) \in \mathcal{C}_{n+1,2d}(\mathbb{R}^{n+1})$.

Proof. For (a), assume $T_d(p-\mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)$. By Definition 3, $\langle T_d(p-\mu), M_d(1,x)\rangle_{n+1,d} \geq 0, \forall x \in \mathbb{R}^n_+$, then

$$p(x) - \mu = \langle T_d(p-\mu), M_d(1,x)\rangle_{n+1,d} \geq 0,\ \forall x \in \mathbb{R}^n_+. \tag{1.3}$$

For the other direction, assume $p(x) \geq \mu, \forall x \in \mathbb{R}^n_+$, then by (1.3), $\langle T_d(p-\mu), M_d(1,x)\rangle_{n+1,d} \geq 0, \forall x \in \mathbb{R}^n_+$. Thus, for any $(x_0, x) \in \mathbb{R}_{++} \times \mathbb{R}^n_+$,

$$\langle T_d(p-\mu), M_d(x_0,x)\rangle_{n+1,d} = x_0 \langle T_d(p-\mu), M_d(1, \frac{x}{x_0})\rangle_{n+1,d} \geq 0. \tag{1.4}$$

Furthermore, due to continuity, for $k > 0$,

$$\langle T_d(p-\mu), M_d(0,x)\rangle = \lim_{k\to+\infty} \langle T_d(p-\mu), M_d(1/k,x)\rangle \geq 0, \tag{1.5}$$

where the last inequality follows from (1.4). From (1.4), (1.5), and Definition 3, it follows that $T_d(p-\mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)$.

The proof of (b) is similar to the proof of (a). □

Corollary 1. *Let $\mu \in \mathbb{R}$, we have*

(a) Let $p(x) \in \mathbb{R}_d[x]$. Then $\inf\{p(x) : x \in \mathbb{R}^n_+\} = \sup\{\mu \in \mathbb{R} : T_d(p-\mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)\}$.

(b) Let $p(x) \in \mathbb{R}_{2d}[x]$. Then $\inf\{p(x) : x \in \mathbb{R}^n\} = \sup\{\mu \in \mathbb{R} : T_{2d}(p-\mu) \in \mathcal{C}_{n+1,2d}(\mathbb{R}^{n+1})\}$.

Theorem 1 and Corollary 1 generalize the key Lemma 2.1 and Corollary 2.1 in [18] for polynomials with degree higher than two using tensor representation. Moreover, Corollary 1 can be seen as a convexification of an unconstrained (possibly non-linear non-convex) POP to a linear conic program over CP and CSDP tensor cones. In the next section, we will discuss the convex relaxations for general constrained polynomial optimization problems.

1.3 Relaxations of POPs

Let $p_i \in \mathbb{R}_d[x], i = 0, \ldots, m$. Consider two general POPs with polynomial constraints:

$$\begin{aligned} z_+ = \inf \quad & p_0(x) \\ \text{s.t.} \quad & p_i(x) \leq 0, \ i = 1, \ldots, m \\ & x \in \mathbb{R}^n_+, \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} z = \inf \quad & p_0(x) \\ \text{s.t.} \quad & p_i(x) \leq 0, \ i = 1, \ldots, m, \end{aligned} \tag{1.7}$$

where $d = \max\{\deg(p_i) : i \in \{0, 1, \ldots, m\}\}$. Problems (1.6) and (1.7) represent general POPs, which encompass a large class of non-linear non-convex problems, including non-convex QPs with binary variables (i.e., binary constraints can be written in the polynomial form $x_i(1 - x_i) \leq 0$, $-x_i(1 - x_i) \leq 0$). Naturally, we have $z \leq z_+$ since the feasible set of problem (1.6) is a subset of problem (1.7). Next we show that the results of Bomze for quadratic problems [18] can be extended to POPs of form (1.6) and (1.7).

1.3.1 Lagrangian relaxations

Let $u_i \geq 0$ be the Lagrangian multiplier of the inequality constraints $p_i(x) \leq 0$ for $i = 1, .., m$ and $v_i \geq 0$ for constraints $x_i \in \mathbb{R}_+$ for $i = 1, ..., n$, so the Lagrangian function for problem (1.6) is

$$L_+(x;u,v) := p_0(x) + \sum_{i=1}^{m} u_i p_i(x) - v^T x,$$

so the Lagrangian dual function of problem (1.6) is

$$\Theta_+(u,v) := \inf\{L_+(x;u,v) : x \in \mathbb{R}^n\},$$

with its optimal value

$$z_{LD,+} = \sup\{\Theta_+(u,v) : (u,v) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+\},$$

We also use a Semi-Lagrangian dual function to represent the nonnegative variable constraints of problem (1.6),

$$\Theta_{semi}(u) := \inf\{L(x;u) : x \in \mathbb{R}^n_+\},$$

with its optimal value

$$z_{semi} = \sup\{\Theta_{semi}(u) : u \in \mathbb{R}^m_+\},$$

Similarly, let $u_i \geq 0$ be the Lagrangian multiplier of the inequality constraints $p_i(x) \leq 0$ for $i = 1, ..., m$, so the Lagrangian function for problem (1.7) is

$$L(x;u) := p_0(x) + \sum_{i=1}^{m} u_i p_i(x),$$

so the Lagrangian dual function of problem (1.7) is

$$\Theta(u) := \inf\{L(x;u) : x \in \mathbb{R}^n\},$$

and the dual optimal value is

$$z_{LD} = \sup\{\Theta(u) : u \in \mathbb{R}^m_+\},$$

from weak duality theory, we have $z_{LD} \leq z$. Thus we have the following relationship:

$$\begin{aligned}\Theta_+(u,v) &= \inf\{L_+(x;u,v) : x \in \mathbb{R}^n\} \\ &\leq \inf\{L_+(x;u,v) : x \in \mathbb{R}^n_+\} \\ &= \inf\{L(x;u) - v^T x : x \in \mathbb{R}^n_+\} \\ &\leq \inf\{L(x;u) : x \in \mathbb{R}^n_+\} = \Theta_{semi}(u),\end{aligned}$$

where the second inequality holds because $x, v \in \mathbb{R}^n_+$ always implies $v^T x \geq 0$. Therefore, we have:

$$z_{LD,+} \leq z_{semi} \leq z_+,$$

where the latter inequality holds by weak duality.

1.3.2 CPSD tensor relaxation for free variables

Consider following *conic program*:

$$\begin{aligned}z_{SP} = \inf \quad & \langle T_d(p_0), X\rangle \\ \text{s.t.} \quad & \langle T_d(p_i), X\rangle \leq 0, \;\; i = 1, \ldots, m \\ & \langle T_d(1), X\rangle = 1 \\ & X \in \mathcal{C}^*_{n+1,d}(\mathbb{R}^{n+1}),\end{aligned} \tag{1.8}$$

and its conic dual problem is

$$z_{SD} = \sup\{\mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+\}. \tag{1.9}$$

For simplicity, we use $\langle \cdot, \cdot \rangle$ represent the tensor inner product of appropriate dimension and order.

Proposition 5. *Problem* (1.8) *is a relaxation of problem* (1.7) *with* $z_{SP} \leq z$.

Proof. Let $x \in \mathbb{R}^n$ be a feasible solution of problem (1.7). It follows that $X = M_d(1, x)$ is a feasible solution of problem (1.8) directly by applying (1.1). Also $p(x) = \langle T_d(p_0), X\rangle$ is a direct result of (1.1) with the same objective value. $\square$

Theorem 2. *For problem* (1.7), *its Lagrangian dual function optimal value satisfies,*

$$z_{LD} = \sup\{\mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}^m_+, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\}$$

and $z_{LD} = z_{SD} \leq z_{SP} \leq z$.

Proof. By Corollary 1 (b),

$$\begin{aligned}\Theta(u) &= \inf\{L(x; u) : x \in \mathbb{R}^n\} \\ &= \sup\{\mu : T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\},\end{aligned}$$

then

$$\begin{aligned}z_{LD} &= \sup\{\Theta(u) : u \in \mathbb{R}^m_+\} \\ &= \sup\{\mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}^m_+, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\}.\end{aligned}$$

From (1.9), we have

$$\begin{aligned}z_{SD} &= \sup\{\mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^{m} u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+\} \\ &= \sup\{\mu : T_d(p_0 + \sum_{i=1}^{m} u_i p_i - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+\} \\ &= \sup\{\Theta(u) : u \in \mathbb{R}^m_+\} \\ &= z_{LD}.\end{aligned}$$

Furthermore, $z_{SD} \leq z_{SP} \leq z$ holds directly from weak conic duality and Proposition 5. $\square$

From Theorem 2, the Lagrangian dual optimal value has no duality gap if and only if conic program itself has no duality gap and positive semidefinite tensor relaxation is tight.

1.3.3 CP and CPSD tensor relaxations for nonnegative variables

Consider following *conic programs*:

$$\begin{aligned} z_{CP} = \inf \quad & \langle T_d(p_0), X\rangle \\ \text{s.t.} \quad & \langle T_d(p_i), X\rangle \leq 0,\ i = 1, \dots, m \\ & \langle T_d(1), X\rangle = 1 \\ & X \in \mathcal{C}^*_{n+1,d}(\mathbb{R}^{n+1}_+), \end{aligned} \tag{1.10}$$

and

$$\begin{aligned} z_{SP,+} = \inf \quad & \langle T_d(p_0), X\rangle \\ \text{s.t.} \quad & \langle T_d(p_i), X\rangle \leq 0,\ i = 1, \dots, m \\ & \langle T_d(-x_i), X\rangle \leq 0,\ i = 1, \dots, n \\ & \langle T_d(1), X\rangle = 1 \\ & X \in \mathcal{C}^*_{n+1,d}(\mathbb{R}^{n+1}), \end{aligned} \tag{1.11}$$

and their conic dual problems

$$z_{CD} = \sup\{\mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^{m} u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+), u \in \mathbb{R}^m_+\}. \tag{1.12}$$

$$z_{SD,+} = \sup\{\mu : T_d(p_0 - \mu) + \sum_{i=1}^{m} u_i T_d(p_i) + \sum_{i=1}^{n} v_i T_d(-x_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+, v \in \mathbb{R}^n_+\}. \tag{1.13}$$

Proposition 6. *Problem* (1.10) *and problem* (1.11) *are relaxations for problem* (1.6) *with* $z_{CP} \leq z_+$ *and* $z_{SP,+} \leq z_+$.

Theorem 3. *For problem* (1.6), *its Semi-Lagrangian dual function optimal value and its Lagrangian dual function optimal value satisfy*

$$\begin{aligned} z_{semi} &= \sup\{\mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}^m_+, T_d(L(x; u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)\}, \\ z_{LD,+} &= \sup\{\mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^n_+, T_d(L_+(x; u, v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\}, \end{aligned}$$

and

(a) $z_{LD,+} \leq z_{semi} = z_{CD} \leq z_{CP} \leq z_+$.

(b) $z_{LD,+} = z_{SD,+} \leq z_{SP,+} \leq z_+$.

Proof. By Corollary 1,

$$\begin{aligned}\Theta_{semi}(u) &= \inf\{L(x;u) : x \in \mathbb{R}^n\}\\ &= \sup\{\mu : T_d(L_+(x;u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\},\\ \Theta_+(u,v) &= \inf\{L_+(x;u,v) : x \in \mathbb{R}^n\}\\ &= \sup\{\mu : T_d(L_+(x;u,v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\},\end{aligned}$$

then

$$\begin{aligned}z_{semi} &= \sup\{\Theta_{semi}(u) : u \in \mathbb{R}^m_+\}\\ &= \sup\{\mu : (\mu, u) \in \mathbb{R}, T_d(L(x;u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)\}.\\ z_{LD,+} &= \sup\{\Theta_+(u,v) : u \in \mathbb{R}^m_+, v \in \mathbb{R}^n_+\}\\ &= \sup\{\mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^n, T_d(L_+(x;u,v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\}.\end{aligned}$$

For (a), from (1.12), we have,

$$\begin{aligned}z_{CD} &= \sup\{\mu : T_d(p_0) - \mu T_d(1) + \sum_{i=1}^m u_i T_d(p_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+), u \in \mathbb{R}^m_+\}\\ &= \sup\{\mu : T_d(p_0(x) + \sum_{i=1}^m u_i p_i(x) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+), u \in \mathbb{R}^m_+\}\\ &= \sup\{\mu : (\mu, u) \in \mathbb{R} \times \mathbb{R}^m_+, T_d(L(x;u) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}_+)\}\\ &= \sup\{\Theta_{semi}(u) : u \in \mathbb{R}^m_+\}\\ &= z_{semi}.\end{aligned}$$

And $z_{CD} \leq z_{CP} \leq z_+$ is an immediate result of weak conic duality and Proposition 6. For (b), from (1.13), we have

$$\begin{aligned} z_{SD,+} &= \sup\{\mu : T_d(p_0 - \mu) + \sum_{i=1}^{m} u_i T_d(p_i) + \sum_{i=1}^{n} v_i T_d(-x_i) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+, v \in \mathbb{R}^n_+\} \\ &= \sup\{\mu : T_d(p_0(x) + \sum_{i=1}^{m} u_i p_i(x) - \sum_{i=1}^{n} v^T x - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1}), u \in \mathbb{R}^m_+, v \in \mathbb{R}^n_+\} \\ &= \sup\{\mu : (\mu, u, v) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^n_+, T_d(L_+(x; u, v) - \mu) \in \mathcal{C}_{n+1,d}(\mathbb{R}^{n+1})\} \\ &= \sup\{\Theta_+(u, v) : u \in \mathbb{R}^m_+, v \in \mathbb{R}^n_+\} \\ &= z_{LD,+}. \end{aligned}$$

And $z_{SD,+} \leq z_{SP,+} \leq z_+$ holds directly from weak conic duality and Proposition 6. □

1.4 Quadratic Reformulation for POPs and its Relaxations

Section 1.3 showed that CP and CPSD tensor relaxations are tighter than Lagrangian relaxations for general POPs. In this section, we will compare CP and CPSD tensor relaxations with quadratic approach for POPs. For general POPs, a classic approach to obtain relaxations is to reformulate them as quadratic programs by introducing additional variables and constraints to address the higher degree terms in polynomials. And a well-studied SDP relaxation or CP relaxation on quadratic constrained quadratic program (QCQP) can then be applied to the reformulated QCQP. Also as discussed in this paper, general POPs can be relaxed directly by conic programs over the CP or the CPSD tensor cones. In general, it is difficult to compare these two relaxations. In this section, we will focus on POPs with degree 4 and apply these two relaxations and show some specific cases in which tensor cone relaxations of POPs give tighter bounds than convex relaxations of QCQP reformulation of POPs.

1.4.1 QCQP Reformulation of POP

A general QCQP reformulation technique of POPs, including how to add additional variables, is discussed in [92]. In this section, the main focus is on some classes of 4th degree POPs, so we use a specific reformulation approach here. We will introduce additional variables to represent the quadratic terms (i.e. the square of single variable and the multiplication of two variables) of the original variables. Consider the following POPs:

$$\begin{aligned} \sup \quad & p_0(x) \\ \text{s.t.} \quad & p_i(x) \le d_i, \ i = 1, \dots, m_0, \\ & q_j(x) \le 0, \ j = 1, \dots, m_1, \\ & x \in \mathbb{R}^n_+, \end{aligned} \tag{1.14}$$

where $p_0(x) \in \mathbb{R}_4[x], q_j \in \mathbb{R}_2[x]$ (Recall $\mathbb{R}_d[x] := \{p \in \mathbb{R}[x] : \deg(p) \le d\}$) and $p_i(x)$ are homogeneous polynomials of degree 4. Problem (1.14) can encompass a large class of 4th degree optimization problems, including problem with 4th degree objective function and linear/quadratic(binary) constraints and so on. This type of optimization problems also appears in many real life problems, such as biquadratic assignment problem [81, 96], Alternating Current Optimal Power Flow (ACOPF) problem [22, 44, 65, 71], etc.

Define an index set

$$S = \{(a, b, c) \in \mathbb{N}^3 : a = 1, \dots, n, b = a, \dots, n, c = (n + 1 - \frac{a}{2})(a - 1) + b - a + 1\} \tag{1.15}$$

as the index for the additional variables, so that it is from 1 to $|S| = \binom{n+1}{2}$, which is the maximum number of additional variables needed to reformulate 4th degree polynomials using 2nd degree polynomials. Specifically, introducing additional variables $y_c =$

$x_a x_b, \forall (a,b,c) \in S$, the QCQP reformulation of problem (1.14) can be represented as

$$\begin{aligned} \sup \quad & q_0(x,y) \\ \text{s.t.} \quad & h_i(y) \le d_i, \ i = 1, \ldots, m_0, \\ & q_j(x) \le 0, \ j = 1, \ldots, m_1, \\ & y_c - x_a x_b = 0, \forall (a,b,c) \in S, \\ & x \in \mathbb{R}_+^n, y \in \mathbb{R}_+^{|S|}, \end{aligned} \tag{1.16}$$

where $q_0(x,y)$ and $h_i(y)$ are the reformulated quadratic polynomials with original variables x and additional variables y by replacing $x_a x_b$ with y_c, $\forall (a,b,c) \in S$, note that $h_i(y)$ are homogeneous polynomials of degree 2. It is clear that $p_0(x) = q_0(x,y), p_i(x) = h_i(y), i = 1, \ldots, m_0$, therefore problem (1.14) and (1.16) are equivalent. As $p_i(x)$ and $h_i(y)$ are homogeneous polynomials, then it follows that

$$\tilde{p}_i(x) = p_i(x) = h_i(y) = \tilde{h}_i(y), i = 1, \ldots, m_0. \tag{1.17}$$

To make the formula clear and easy to represent in a conic program, let $z = (x,y) \in \mathbb{R}_+^{n+|S|}$, then (1.16) is equivalent to

$$\begin{aligned} \sup \quad & q_0(z) \\ \text{s.t.} \quad & h_i(z) \le d_i, \ i = 1, \ldots, m_0, \\ & q_j(z) \le 0, \ j = 1, \ldots, m_1, \\ & z_{n+c} - z_a z_b = 0, \forall (a,b,c) \in S, \\ & z \in \mathbb{R}_+^{n+|S|}. \end{aligned} \tag{1.18}$$

Example 1. *QP Reformulation*

Consider the following univariate program,

$$\begin{aligned} \sup \quad & x^4 + x^3 + x^2 + x + 1 \\ \text{s.t.} \quad & x^4 \le 1, \\ & x^2 - x - 2 \le 0, \\ & -x + 1 \le 0, \\ & x \in \mathbb{R}_+. \end{aligned} \tag{1.19}$$

Let $y = x^2$ and $z = (x, y) \in \mathbb{R}^2_+$, then problem (1.19) is equivalent to

$$\begin{aligned} \sup \quad & y^2 + xy + y + x + 1 \\ \text{s.t.} \quad & y^2 \le 1, \\ & y - x - 2 \le 0, \\ & -x + 1 \le 0, \\ & x \in \mathbb{R}_+, y \in \mathbb{R}_+. \end{aligned} \quad \equiv \quad \begin{aligned} \sup \quad & z_2^2 + z_1 z_2 + z_2 + z_1 + 1 \\ \text{s.t.} \quad & z_2 \le 1, \\ & z_2 - z_1 - 2 \le 0, \\ & -z_1 + 1 \le 0, \\ & z \in \mathbb{R}^2_+. \end{aligned}$$

1.4.2 CP matrix relaxations for QCQP

Consider the following CP matrix relaxations for problem (1.18),

$$\begin{aligned} \sup \quad & \langle T_2(q_0(z)), Z \rangle \\ \text{s.t.} \quad & \langle T_2(h_i(z)), Z \rangle \le d_i, \ i = 1, \ldots, m_0, \\ & \langle T_2(q_j(z)), Z \rangle \le 0, \ j = 1, \ldots, m_1, \\ & \langle T_2(1), Z \rangle = 1, \\ & Z_{1,n+c+1} - Z_{a+1,b+1} = 0, \forall (a, b, c) \in S, \\ & Z \in \mathcal{C}^*_{n+r+1,2}(\mathbb{R}^{n+r+1}_+). \end{aligned} \tag{1.20}$$

where $r = |S|$ is the number of additional variables in problem (1.18). Problem (1.20) is a natural CP tensor relaxation of problem (1.18) and by relaxing the equality constraints

$Z_{1,c+n+1} - Z_{a+1,b+1} = 0, \forall (a,b,c) \in S$ into inequality constraints, we have the following CP tensor relaxation,

$$
\begin{aligned}
\sup \quad & \langle T_d(q_0(z)), Z\rangle \\
\text{s.t.} \quad & \langle T_2(h_i(z)), Z\rangle \leq d_i,\ i = 1,\ldots,m_0, \\
& \langle T_2(q_j(z)), Z\rangle \leq 0,\ j = 1,\ldots,m_1, \\
& \langle T_2(1), Z\rangle = 1, \\
& Z_{1,c+n+1} - Z_{a+1,b+1} \leq 0, \forall (a,b,c) \in S, \\
& Z \in \mathcal{C}^*_{n+r+1,2}(\mathbb{R}^{n+r+1}_+),
\end{aligned}
\tag{1.21}
$$

Proposition 7. *If problem* (1.20) *is feasible and the coefficients of* $q_0(z)$ *in problem* (1.20) *are nonnegative, then problems* (1.20) *and* (1.21) *are equivalent.*

Proof. It is clear that if the coefficients of objective function $q_0(z)$ are nonnegative, at optimality of problem (1.21), $Z_{1,k+n+1} = Z_{i+1,j+1}$ holds. And the same objective function values are obtained for problems (1.20) and (1.21). □

Recall the CP tensor relaxation (1.10) for general POPs and apply it directly to problem (1.18), then we have the following conic program,

$$
\begin{aligned}
\sup \quad & \langle T_d(p_0(x)), X\rangle \\
\text{s.t.} \quad & \langle T_d(p_i(x)), X\rangle \leq d_i,\ i = 1,\ldots,m_0, \\
& \langle T_d(q_j(x)), X\rangle \leq 0,\ j = 1,\ldots,m_1, \\
& \langle T_d(1), X\rangle = 1, \\
& X \in \mathcal{C}^*_{n+1,d}(\mathbb{R}^{n+1}_+),
\end{aligned}
\tag{1.22}
$$

Problem (1.20) and (1.22) can be seen as two different relaxations for some classes POPs with a form of problem (1.14). Problem (1.20) characterizes the polynomials with higher degree than 2 by reformulating them as quadratic polynomials. SDP and CP matrix relaxations for the reformulated QCQP are well studied in literature [cf., 5, 18, 19, 20, 27, 29, 58, 113]. However, the introduce of additional constraints $Z_{1,c+n+1} - Z_{a+1,b+1} = 0, \forall (a,b,c) \in S$ in problem (1.20) may ruin some exact relaxation conditions for QCQP. Problem (1.22) characterizes the polynomials with degree higher than 2 by using higher order tensors which avoids introducing additional variables and constraints. Next we will show that under some conditions, the latter relaxations will provide tighter bounds for problem (1.14).

Lemma 1 ([92, Lemma 2]). *For any $d > 0$ and $n > 0$, $\mathcal{C}^*_{n+1,d}(\mathbb{R}^{n+1}_+) = conic(M_d(\{0,1\} \times \mathbb{R}^n_+))$.*

Theorem 4. *Consider a feasible problem* (1.14) *where the coefficients of $p_0(x)$ are nonnegative, then problem* (1.20) *is a relaxation of problem* (1.22)*.*

Proof. By Proposition 7, problems (1.20) and (1.21) are equivalent. For any feasible solution $X \in \mathcal{C}^*_{n+1,4}(\mathbb{R}^{n+1}_+)$ to problem (1.22), by Lemma 1,

$$X = \sum_{s=1}^{n_1} \lambda_s M_4(1, u_s) + \sum_{t=1}^{n_0} \gamma_t M_4(0, v_t),$$

for some $n_0, n_1 \geq 0, \lambda_s, \gamma_t > 0$ and $u_s, v_t \in \mathbb{R}^n_+$. Then by using (1.1),

$$\begin{aligned} 1 = \langle T_4(1), X \rangle &= \sum_{s=1}^{n_1} \lambda_s, \\ d_i \geq \langle T_4(p_i), X \rangle &= \sum_{s=1}^{n_1} \lambda_s p_i(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_i(v_t), \ i = 1, ..., m_0, \\ 0 \geq \langle T_4(q_j), X \rangle &= \sum_{s=1}^{n_1} \lambda_s q_j(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_j(v_t), \ j = 1, ..., m_1, \end{aligned} \tag{1.23}$$

with an objective function value of $\sum_{s=1}^{n_1} \lambda_s p_0(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_0(v_t)$. Recall the index set S in (1.15), and construct a vector of w_s, w'_t for $s = 1, ..., n_1, t = 1, ..., n_0$ as follows:

$$\begin{aligned} (w_s)_c &= (u_s)_a (u_s)_b, \quad (a,b,c) \in S, \\ (w'_t)_c &= (v_t)_a (v_t)_b, \quad (a,b,c) \in S. \end{aligned} \tag{1.24}$$

Next we show

$$Z = \sum_{s=1}^{n_1} \lambda_s M_2(1, (u_s, w_s)) + \sum_{t=1}^{n_0} \gamma_t M_2(0, (v_t, w_t')), \tag{1.25}$$

is a feasible solution to problem (1.20). Clearly, $Z \in \mathcal{C}^*_{n+r+1,2}(\mathbb{R}_+^{n+r+1})$, and from equation (1.24) and (1.25), we have

$$\begin{aligned} Z_{1,c+n+1} &= \sum_{s=1}^{n_1} \lambda_s (w_s)_c = \sum_{s=1}^{n_1} \lambda_s (u_s)_a (u_s)_b, \forall (a,b,c) \in S, \\ Z_{a+1,b+1} &= \sum_{s=1}^{n_1} \lambda_s (u_s)_a (u_s)_b + \sum_{t=1}^{n_0} \gamma_t (v_t)_a (v_t)_b, \forall (a,b,c) \in S, \end{aligned}$$

which indicates that $Z_{1,c+n+1} \leq Z_{a+1,b+1}, \ \forall (a,b,c) \in S$. From equations (1.17) and (1.23),

$$\begin{aligned} \langle T_2(1), Z\rangle &= \sum_{s=1}^{n_1} \lambda_s = 1, \\ \langle T_2(h_i), Z\rangle &= \sum_{s=1}^{n_1} \lambda_s h_i(w_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{h}_i(w_t') \\ &= \sum_{s=1}^{n_1} \lambda_s p_i(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{p}_i(v_t) \leq d_i, \ i = 1, ..., m_0, \\ \langle T_2(q_j), Z\rangle &= \sum_{s=1}^{n_1} \lambda_s q_j(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_j(v_t) \leq 0, \ j = 1, ..., m_1, \end{aligned}$$

with an objective value of

$$\sum_{s=1}^{n_1} \lambda_s q_0(u_s, w_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w_t') = \sum_{s=1}^{n_1} \lambda_s p_0(u_s) + \sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w_t'),$$

under the condition that $p_0(x)$ has nonnegative coefficients and $x \in \mathbb{R}^n_+$,

$$\sum_{t=1}^{n_0} \gamma_t \tilde{q}_0(v_t, w_t') \geq \sum_{t=1}^{n_0} \gamma_t \tilde{p}_0(v_t).$$

Therefore, from any feasible solution to problem (1.22), we can construct a feasible solution to problem (1.21) with a larger objective function value, which indicates that problem (1.20) is a relaxation for problem (1.22). □

Theorem 4 illustrates that theoretically the CP tensor relaxation can provide tighter bounds than QCQP approches for a class of POPs. In next section, we will introduce some approximation strategies for CP and CPSD tensor programs to address the intractability.

1.5 Numerical Comparison of Two Relaxations of PO

Unlike the tractability of the PSD matrix cone, the CPSD tensor cone is not tractable in general to our knowledge. Also similar to the intractability of completely positive matrices of dimension greater than 5, the CP tensor cone is also not tractable in general. In this section, we will discuss and develop tractable approximations for CP and CPSD tensor cones, and then use these approximations to address some POPs to show it provides better bounds than approximations for QCQP reformulation.

1.5.1 Approximation of CP and CPSD Tensor Cones

Before presenting results, let us introduce some more notations. For $T = M_d(x), x \in \mathbb{R}^n$, denote $T_{(i_1,...,i_d)}$ as the element in $(i_1, ..., i_d)$ position, where $(i_1, ..., i_d) \in \{1, \dots, n\}^d$. To be more specific, i_j with $j = 1, ..., d$ means the choice of $\{x_1, ..., x_n\}$ in the j^{th} position in the tensor product, i.e. $i_1 = 2$ means choosing x_2 as the first position in the tensor product. To illustrate, let $x \in \mathbb{R}^3$ and let

$$T^1 = M_2(x) = \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 \\ x_1x_2 & x_2^2 & x_2x_3 \\ x_1x_3 & x_2x_3 & x_3^2 \end{pmatrix},$$

then $T^1_{(1,2)} = x_1x_2$ and it is in the (1,2) position in T^1. Also for $T = M_d(x), x \in \mathbb{R}^n$, when $d > 2$, let $T_{(i_1,...,i_{d-2},\cdot,\cdot)}$ denote the matrix in $(i_1, ..., i_{d-2}, \cdot, \cdot)$ position, where $(i_1, ..., i_{d-2}, \cdot, \cdot)$ represents the matrix

$$(T_{(i_1,...,i_{d-2},\cdot,\cdot)})_{jk} = T_{i_1,...,i_{d-2},j,k}, \ j, k = 1, \dots, n,$$

for example, let $T^2 = M_3(x), x \in \mathbb{R}^3$, then

$$T^2_{(1,\cdot,\cdot)} = \begin{pmatrix} x_1^3 & x_1^2x_2 & x_1^2x_3 \\ x_1^2x_2 & x_1x_2^2 & x_1x_2x_3 \\ x_1^2x_3 & x_1x_2x_3 & x_1x_3^2 \end{pmatrix}, T^2_{(2,\cdot,\cdot)} = \begin{pmatrix} x_1^2x_2 & x_1x_2^2 & x_1x_2x_3 \\ x_1x_2^2 & x_2^3 & x_2^2x_3 \\ x_1x_2x_3 & x_2^2x_3 & x_2x_3^2 \end{pmatrix}.$$

Definition 5. *Let $T = M_d(x), x \in \mathbb{R}^n$. For any $(i_1, \ldots, i_{d-2}) \in \{1, \ldots, n\}^{d-2}$, $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)}$ is a* principal matrix *if $I_k \subseteq \{0, \ldots, d-2\}$ is even for all $k = 1, \ldots, n$, where I_k is an ordered set of the number of appearance $i_j = k, \forall j = 1, \ldots, d-2$ where $k = 1, ..., n$.*

For example, let $T^3 = M_8(x), x \in \mathbb{R}^3$, then

$T^3_{(1,1,2,2,3,3,\cdot,\cdot)}, T^3_{(1,2,2,2,1,2,\cdot,\cdot)}, T^3_{(2,3,2,1,3,1,\cdot,\cdot)}$ are principal matrices;

$T^3_{(1,1,1,2,3,3,\cdot,\cdot)}, T^3_{(1,2,2,2,2,2,\cdot,\cdot)}, T^3_{(2,3,2,2,3,1,\cdot,\cdot)}$ are not principal matrices.

Notice the symmetry of symmetric tensors, $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)}$ with the same $I_k, k = 1, ..., n$ are equal. Next we will discuss the approximation strategies for the CP and the CPSD tensor cones based on PSD and DNN matrices.

Definition 6. *A symmetric matrix X is called* doubly nonnegative (DNN) *if and only if $X \succeq 0$ and $X \geq 0$, where $X \geq 0$ indicates every element of X is nonnegative.*

Proposition 8. *For any symmetric tensor T,*

(a) If $T \in \mathcal{C}^*_{n,d}(\mathbb{R}^n_+)$, then $T_{(i_1,\ldots,i_d)} \geq 0, T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \succeq 0, \forall i = 1, ..., n$.

(b) If $T \in \mathcal{C}^*_{n,d}(\mathbb{R}^n)$, for all principal matrices $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)}$, $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \succeq 0, \forall i = 1, ..., n$.

Proof. For part (a), by Proposition 3 (a), $T = \sum_i \lambda_i M_d(x^i)$, where $x^i \in \mathbb{R}^n_+, \lambda_i \geq 0, \sum_i \lambda_i = 1$, then it is clear that $T_{(i_1,\ldots,i_d)} \geq 0$, and

$$T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} = \sum_i \lambda_i \prod_{k=1}^n (x^i_k)^{I_k} (x^i (x^i)^T), \tag{1.26}$$

as $x^i(x^i)^T \succeq 0, \forall i$ and $\prod_{k=1}^n (x_k^i)^{I_k} \geq 0$, then $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \succeq 0$. For part (b), noticing that the number of appearance $I_k, k = 1, ..., n$ is even if $T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)}$ is a principal matrix, then it follows proof of (a) with $\prod_{k=1}^n (x_k^i)^{I_k} \geq 0$ in (1.26). □

Take $T \in \mathcal{C}^*_{2,4}(\mathbb{R}^2_+)$ as an example to illustrate Proposition 8, by Proposition 3 (a), $T = \sum_i \lambda_i M_4(x^i)$, where $\lambda_i \geq 0, \sum_i \lambda_i = 1$ and $x^i \in \mathbb{R}^2_+$, then for any $y \in \mathbb{R}^2$,

$$y^T T_{(1,2,\cdot,\cdot)} y = y^T \sum_i \lambda_i M_4(x^i)_{(1,2,\cdot,\cdot)} y = x_1^i x_2^i \sum_i (y^T x^i)^2 \geq 0,$$

which indicates that $T_{(1,2,\cdot,\cdot)}$ is a 2×2 positive semidefinite matrix.

Next we discuss the approximation of the CPSD and the CP tensor cones. Based on Proposition 8, we define the following tensor cones,

$$\begin{aligned}
\mathcal{K}^{SDP}_{n,d} &= \{T \in \mathcal{S}_{n,d} : T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \succeq 0,\ \forall (i_1,\ldots,i_{d-2}) \in \{1,\ldots,n\}^{d-2}\},\\
\mathcal{K}^{L}_{n,d} &= \{T \in \mathcal{S}_{n,d} : T_{(i_1,\ldots,i_d)} \geq 0,\ \forall (i_1,\ldots,i_d) \in \{1,\ldots,n\}^{d}\},\\
\mathcal{K}^{DNN}_{n,d} &= \{T \in \mathcal{S}_{n,d} : T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \succeq 0, T_{(i_1,\ldots,i_{d-2},\cdot,\cdot)} \geq 0,\ \forall (i_1,\ldots,i_{d-2}) \in \{1,\ldots,n\}^{d-2}\}.
\end{aligned}$$

It is easy to see these cones are *convex closed cones* with the following relationship,

$$\begin{aligned}
\mathcal{C}^*_{n,d}(\mathbb{R}^n) &\subseteq \mathcal{K}^{SDP}_{n,d}\\
\mathcal{C}^*_{n,d}(\mathbb{R}^n_+) &\subseteq \mathcal{K}^{DNN}_{n,d} \subseteq \mathcal{K}^{L}_{n,d}.
\end{aligned} \tag{1.27}$$

Consider the following conic program,

$$\begin{array}{lll}
\textbf{[TP-}\mathcal{K}\textbf{]} & \inf & \langle T_d(p_0), X\rangle\\
& \text{s.t.} & \langle T_d(p_i), X\rangle \leq 0,\ i = 1,\ldots,m\\
& & \langle T_d(1), X\rangle = 1\\
& & X \in \mathcal{K}_{n+1,d}.
\end{array}$$

From (1.27), problem **[TP-$\mathcal{K}$]** is a tractable relaxation for problem (1.8) and (1.10) by choosing appropriate tractable cones, and thus provides relaxations to globally approximate general POPs. It follows that

$$z_{[\textbf{TP-}\mathcal{K}^{SDP}]} \leq z_{SP} \leq z,$$

$$z_{[\textbf{TP-}\mathcal{K}^{L}]} \leq z_{[\textbf{TP-}\mathcal{K}^{DNN}]} \leq z_{CP} \leq z_{+}.$$

1.5.2 Numerical results on general cases

In Section 1.5.1, several tractable approximations for the CP and the CPSD tensor cones have been developed to provide relaxations for CP and CPSD tensor programs. In this section, we will provide numerical results on more general POP cases in order to compare the bounds of two relaxation approaches discussed in Section 1.4.2. Denote $[\mathbf{QP}_L]$ and $[\mathbf{QP}_{DNN}]$ as the linear relaxation and DNN relaxation for problem (1.20) similar to $[\textbf{TP-}\mathcal{K}^{L}]$ and $[\textbf{TP-}\mathcal{K}^{DNN}]$, and denote $[\mathbf{QP}_{SDP}]$ for the SDP relaxation for the quadratic reformulation problem (1.7). Recall the number of additional variables $r = \binom{n+1}{2}$. In Table 1.5.1, we compare the two approaches in terms of number and size of PSD matrices.

	PSD matrix size	PSD matrix number	Number of variables
$[\mathbf{QP}_{SDP}]$	$(1+n+r)\times(1+n+r)$	1	$O(n^4)$
$[\textbf{TP-}\mathcal{K}^{SDP}]$	$(1+n)\times(1+n)$	n	$O(n^3)$
$[\mathbf{QP}_{DNN}]$	$(1+n+r)\times(1+n+r)$	1	$O(n^4)$
$[\textbf{TP-}\mathcal{K}^{DNN}]$	$(1+n)\times(1+n)$	$O(n^2)$	$O(n^4)$

Table 1.5.1: Program size comparison

Followings are some test problems for the comparison. Note that there preliminary results are on small scale problems, only bounds are compared as the time difference is negligible. All the numerical experiments are conducted on a 2.4 GHz CPU laptop with 8 GB memory. We implement all the models with `YALMIP` in `Matlab` . We use `SeDuMi` as the SDP solver and `CPLEX` as the LP solver. For Example 5 and 6, we use `Couenne` as the global solver.

Example 2.

Consider the following problem,

$$\begin{aligned} \min \quad & \left(\sum_{i=1}^{n} x_i\right)^4 \\ \text{s.t.} \quad & x_1^4 = 1, \\ & x_i \geq 0, i = 1, ..., n. \end{aligned} \tag{1.28}$$

By observation, the optimal value is 1, with an optimal solution $x_1^* = 1, x_k^* = 0, k = 2, ..., n$. The QCQP reformulation of (1.28) with least number of additional variables is

$$\begin{aligned} \min \quad & y_1^2 \\ \text{s.t.} \quad & y_1 = \left(\sum_{i=1}^{n} x_i\right)^2, \\ & y_2 = x_1^2, \\ & y_2^2 = 1, \\ & x_i \geq 0, i = 1, ..., n, \\ & y_1, y_2 \geq 0. \end{aligned} \tag{1.29}$$

Relaxation $[\mathbf{TP}\text{-}\mathcal{K}^L]$ can be directly applied to (1.28) and gives an optimal value of 1 while $[\mathbf{QP}_L]$ for (1.29) gives an optimal value of 0, which means the approximation by using tensor relaxation is tight.

Example 3. *Bi-quadratic POPs*

Bi-quadratic problem and its difficulty have been studied in [75]. Consider the following specific bi-quadratic POPs,

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n, y\in\mathbb{R}^m} \quad & p_0 = \sum_{\substack{1\le i<j\le n\\ 1\le a<b\le m}} x_i x_j y_a y_b \\
\text{s.t.} \quad & \|x\|^2 = 1, \|y\|^2 = 1,
\end{aligned}
\tag{1.30}
$$

where $\|\cdot\|$ is the standard 2-norm in Euclidean spaces. It is clear that problem (1.30) is equivalent to

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n, y\in\mathbb{R}^m} \quad & p_0 = \frac{1}{4}[x^T(e_n e_n^T - I_n)x][y^T(e_m e_m^T - I_m)y] \\
\text{s.t.} \quad & \|x\|^2 = 1, \|y\|^2 = 1,
\end{aligned}
$$

where e_n, e_m are all-one vectors of appropriate dimension and I_n, I_m are diagonal matrices of dimension $n\times n$ and $m\times m$. It is then easy to see the optimal value is $-\frac{1}{4}(\max\{n,m\}-1)$. By defining an index set

$$
S(n) = \{(i,j,k)\in\mathbb{N}^3 : i = 1,\dots,n-1, j = i+1,\dots,n, k = (n-\frac{i}{2})(i-1)+j-i\}
$$

for additional variables, we can reformulate problem (1.30) as a quadratic problem by introducing additional variables,

$$
\begin{aligned}
\min \quad & \sum_{\substack{1\le k\le |S(n)|\\ 1\le c\le |S(m)|}} w_k z_c \\
\text{s.t.} \quad & w_k = x_i x_j, \forall (i,j,k)\in S(n), \\
& z_c = y_a y_b, \forall (a,b,c)\in S(m), \\
& \|x\|^2 = 1, \|y\|^2 = 1,
\end{aligned}
\tag{1.31}
$$

where $w, z \in \mathbb{R}^m$ with $|S(n)| = n(n-1)/2, |S(m)| = m(m-1)/2$. Let $u = \Big[x; y; w; z\Big]$, and a positive semedefinite relaxation can be applied to problem (1.31),

$$
\begin{aligned}
\min \quad & \sum_{\substack{n+m+1 \leq p \leq n+m+|S(n)| \\ n+m+|S(n)|+1 \leq q \leq n+m+|S(n)|+|S(m)|}} Q_{pq} \\
\text{s.t.} \quad & u_{n+m+k} = Q_{ij}, \forall (i,j,k) \in S(n), \\
& u_{n+m+|S(n)|+c} = Q_{n+a,n+b}, \forall (a,b,c) \in S(m), \\
& \sum_{i=1}^{n} Q_{ii} = 1, \\
& \sum_{i=n+1}^{n+m} Q_{ii} = 1, \\
& \begin{pmatrix} 1 & u^T \\ u & Q \end{pmatrix} \in \mathcal{C}^*_{n+m+|S(n)|+|S(m)|+1,2}(\mathbb{R}^{n+m+|S(n)|+|S(m)|+1}).
\end{aligned}
\tag{1.32}
$$

Note that problem (1.32) is a simple SDP relaxation for problem (1.30). More elaborated SDP relaxations that provide bounds with guaranteed performance are discussed for this type of problem in [75].

Proposition 9. *Problem* (1.32) *is unbounded.*

Proof. Let $\bar{u}$ be a $(n+m+|S(n)|+|S(m)|) \times 1$ all-zero vector and let $\bar{Q}$ be a $(n+m+|S(n)|+|S(m)|) \times (n+m+|S(n)|+|S(m)|)$ matrix such that

$$\bar{Q}_{11} = \bar{Q}_{n+1,n+1} = 1, \bar{Q}_{n+m+1,n+m+1} = \bar{Q}_{n+m+|S(n)|+1,n+m+|S(n)|+1} = M^2,$$

$$\bar{Q}_{n+m+1,n+m+|S(n)|+1} = \bar{Q}_{n+m+|S(n)|+1,n+m+1} = -M,$$

where M is a positive number and let all other entries for $\bar{Q}$ be 0. It is clear that $(\bar{u}, \bar{Q})$ is a feasible solution to problem (1.32). However, as $M \to \infty$, the objective function goes to $-\infty$, thus the problem is unbounded. □

Proposition 9 tells that relaxation $[\mathbf{QP}_{SDP}]$ for problem (1.30) will fail to provide a bound. However, a CPSD tensor cone can be directly applied to problem (1.30),

$$\begin{aligned} \min \quad & \langle T_4(p_0), X\rangle \\ \text{s.t.} \quad & \langle T_4(\|x\|^2), X\rangle = 1, \\ & \langle T_4(\|y\|^2), X\rangle = 1, \\ & \langle T_4(1), X\rangle = 1, \\ & X \in \mathcal{C}^*_{n+m+1,4}(\mathbb{R}^{n+m+1}). \end{aligned} \tag{1.33}$$

Problem [**TP-**$\mathcal{K}^{SDP}$] can be used to approximate problem (1.33) and the results are listed in Table 1.5.2. In Table 1.5.2, we can see that relaxation [**TP-**$\mathcal{K}^{SDP}$] can provide the optimal value for problem (1.32) while relaxation [$\mathbf{QP}_{SDP}$] for the QCQP reformulation of problem (1.32) fails to give a bound.

(n,m)	Optimal	[**TP-**$\mathcal{K}^{SDP}$]	(n,m)	Optimal	[**TP-**$\mathcal{K}^{SDP}$]
(2,2)	-0.25	-0.25	(2,10)	-2.25	-2.25
(3,3)	-0.50	-0.50	(3,9)	-2.00	-2.00
(4,4)	-0.75	-0.75	(4,8)	-1.75	-1.75
(5,5)	-1.00	-1.00	(5,7)	-1.50	-1.50
(6,6)	-1.25	-1.25			
(7,7)	-1.50	-1.50			
(8,8)	-1.75	-1.75			
(9,9)	-2.00	-2.00			
(10,10)	-2.25	-2.25			

Table 1.5.2: Relaxation comparisons for Example 3

Example 4. *Non-convex QCQP*

Consider the following nonconvex QCQP,

$$\begin{aligned}
\min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
\text{s.t.} \quad & f_1(x) = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
& f_2(x) = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
& f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
& x_1, x_2 \geq 0
\end{aligned} \tag{1.34}$$

The optimal solution of the example is $x^* = (0, 0.6667)^T$ with $f_0(x^*) = -6.4444$ (see [119]). A semidefinite relaxation and a copositive relaxation has been studied in [119], which gives a bound of -103.43 and -26.67 respectively for problem (1.34) (refer to Table 2 in [119], (SDP+RLT) is actually a DNN relaxation for copositive programming).

For tensor relaxations, we manually add valid inequalities to make the problem a 4-th degree POP,

$$\begin{aligned}
\min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
\text{s.t.} \quad & f_1(x) = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
& f_2(x) = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
& f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
& x_2 f_2(x) \leq 0, \\
& x_1^2 f_1(x) \leq 0, \\
& x_1, x_2 \geq 0
\end{aligned} \tag{1.35}$$

then [**TP-**$\mathcal{K}^{DNN}$] can be used to approximate Problem (1.35), we obtain a bound of -12.83, which provides better bounds than SDP relaxation and completely positive relaxation on problem (1.34). We also add the valid inequalities $x_2 f_2(x) \leq 0, x_1^2 f_1(x) \leq 0$ directly to problem (1.34) by reformulating problem (1.35) as a quadratic program by adding additional variables and constraints as in (1.16):

$$
\begin{aligned}
\min \quad & f_0(x) = -8x_1^2 - x_1x_2 - 13x_2^2 - 6x_1 - x_2 \\
\text{s.t.} \quad & -y_1 = x_1^2 + x_1x_2 + 2x_2^2 - 3x_1 - 3x_2 - 7 \leq 0, \\
& -y_2 = 2x_1x_2 + 33x_1 + 15x_2 - 10 \leq 0, \\
& f_3(x) = x_1 + 2x_2 - 6 \leq 0, \\
& y_3 = x_1^2, \\
& -x_2y_2 \leq 0, \\
& -y_1y_3 \leq 0, \\
& x_1, x_2, y_1, y_2, y_3 \geq 0.
\end{aligned}
$$

In addition to adding valid inequalities to strengthen the tensor relaxation, adding positive semidefinite (PSD) *reformulation linearization technique* (RLT) constraints can further strengthen the relaxations. Similar to second order RLT introduced in [28], with the constraint $\langle T_4(1), X\rangle = 1$ and conic constraint $X \in \mathcal{C}^*_{3,4}(\mathbb{R}^3)$, for a quadratic constraint $c_0 + c_{10}x_1 + c_{01}x_2 + c_{11}x_1^2 + c12x_1x_2 + c_{22}x_2^2 \geq 0$, following PSD-RLT constraints for CP tensor relaxation of problem (1.35) can be constructed,

$$
\begin{aligned}
&(c_0 + c_{10}x_1 + c_{01}x_2 + c_{11}x_1^2 + c_{12}x_1x_2 + c_{22}x_2^2)X_{(0,0,\cdot,\cdot)} \\
=&c_0X_{(0,0,\cdot,\cdot)} + c_{10}X_{(1,0,\cdot,\cdot)} + c_{01}X_{(2,0,\cdot,\cdot)} + c_{11}X_{(1,1,\cdot,\cdot)} + c_{12}X_{(1,2,\cdot,\cdot)} + c_{22}X_{(2,2,\cdot,\cdot)} \succeq 0,
\end{aligned}
\tag{1.36}
$$

where $X_{(0,0,\cdot,\cdot)} \succeq 0$ as discussed in Section 1.5.1. Note that the PSD-RLT can't be used in the CP relaxations on the QCQP reformulation. With the PSD-RLT constraints based on constraints, the optimal value is obtained at -6.4444. A comparison of bounds is listed in Table 1.5.3.

	Without Valid Inequalities		With Valid Inequalities	
	SDP	COP	$[\mathbf{QP}_{DNN}]$	$[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$
Bound	-103.43	-26.67	-26.67	-12.83

Table 1.5.3: Relaxation comparisons for Example 4

Example 5. *Random Objective Function on a Feasible Region*

In this example, we will present our preliminary numerical results on randomly generated 4th degree POPs with feasible regions. The test problem is

$$\begin{array}{ll} \min & \text{Randomly generated 4th degree homogenous polynomial of 3 variables} \\ \text{s.t.} & (x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \geq 0.2^2, \\ & (x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \leq 0.6^2, \\ & 0 \leq x_1, x_2, x_3 \leq 1, \end{array} \tag{1.37}$$

The coefficients in the objective function are integers in the range $[-5, 5]$. The first and second constraints make the problem nonconvex and it is easy to see the problem is feasible. We use $[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$ to directly approximate problem (1.37) and $[\mathbf{QP}_{DNN}]$ to approximate the QCQP reformulation of problem (1.37). We denote *ratio* as the improve ratio similar to that in [119] and

$$ratio = \frac{[\mathbf{TP}\text{-}\mathcal{K}^{DNN}] \ - [\mathbf{QP}_{DNN}]}{f_{opt} - [\mathbf{QP}_{DNN}]}.$$

We also add PSD-RLT constraints for $(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \geq 0.2^2$ and $(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + (x_3 - 0.5)^2 \leq 0.6^2$ to problem (1.37). In Table 1.5.4, the relaxation $[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$ with PSD-RLT constraints provides the tightest bounds where for instances optimal values are obtained. Relaxation $[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$ provides tighter bounds than $[\mathbf{QP}_{DNN}]$ for most test instances. For instances 8,9,18 and 20, relaxation **[TP-**

$\mathcal{K}^{DNN}$] gives the optimal objective value, while [$\mathbf{QP}_{DNN}$] is not tight. For instances 15 and 17, [**TP-**$\mathcal{K}^{DNN}$] and [$\mathbf{QP}_{DNN}$] give the same bound. An average of 50% improve ratio implies that [**TP-**$\mathcal{K}^{DNN}$] provides tighter bounds than [$\mathbf{QP}_{DNN}$] for Problem (1.37). In Table 1.5.4, relaxation [**TP-**$\mathcal{K}^{DNN}$] provides tighter bounds than [$\mathbf{QP}_{DNN}$] for most test instances. For instances 8,9,18 and 20, relaxation [**TP-**$\mathcal{K}^{DNN}$] gives the optimal objective value, while [$\mathbf{QP}_{DNN}$] is not tight. For instances 15 and 17, [**TP-**$\mathcal{K}^{DNN}$] and [$\mathbf{QP}_{DNN}$] give the same bound. An average of 50% improve ratio implies that [**TP-**$\mathcal{K}^{DNN}$] provides better relaxations than [$\mathbf{QP}_{DNN}$] for Example 5.

Test No.	[$\mathbf{QP}_{DNN}$]	[**TP**-$\mathcal{K}^{DNN}$]	*ratio*	[**TP**-$\mathcal{K}^{DNN}$] $^{+}$	`Couenne`
1	-4.6732	-3.2860	36.82%	-0.9055*	-0.9055*
2	-8.8748	-5.2725	78.15%	-4.2654*	-4.2654*
3	-5.6429	-4.1135	76.76%	-3.6477*	-3.6477*
4	-3.5507	-2.1173	53.59%	-0.8761*	-0.8761*
5	-11.0434	-9.5248	37.81%	-7.0268*	-7.0268*
6	-12.6822	-10.5600	24.46%	-4.0055*	-4.0055*
7	-3.0709	-2.4427	45.84%	-1.7005*	-1.7005*
8	0	0.0122	100%	0.0122*	0.0122*
9	-1	0.0091	100%	0.0091*	0.0091*
10	-5.2621	-1.9963	83.15%	-1.3345*	-1.3345*
11	-0.8450	-0.8438	0.30%	-0.4922*	-0.4922*
12	-3.5894	-2.9945	100%	-1.4597*	-1.4597*
13	-0.8554	-0.7762	11.70%	-0.1787*	-0.1787*
14	-6.1631	-2.6502	81.87%	-1.8723*	-1.8723*
15	-0.2666	-0.2666	0	-0.1487*	-0.1487*
16	-6.0238	-5.6216	10.16%	-2.0645*	-2.0645*
17	-4.9579	-4.9579	0	-4.0253*	-4.0253*
18	0	0.0080	100%	0.0080*	0.0080*
19	-12.1584	-10.7368	16.94%	-3.7659*	-3.7659*
20	-0.6545	0.0112	100%	0.0112*	0.0112*

$*$: Optimal value is obtained.

[**TP**-$\mathcal{K}^{DNN}$] $^{+}$: [**TP**-$\mathcal{K}^{DNN}$] with PSD-RLT constraints.

Table 1.5.4: Relaxation comparisons for Example 5

Example 6. *Numerical Results on Random Generated Polynomial Problems*

In this example, we present our preliminary numerical results on randomly generated polynomial optimization problems. The objective function is a 4th degree homogenous polynomial of 3 variables, with two 4th degree polynomial inequality constraints, a linear inequality constraint and nonnegative variables. The coefficients in the objective function are integers in the range $[-5, 5]$ and the coefficients of the two polynomial constraints are integers in the range $[-10, 10]$ and the coefficients of linear constraint are integers in the range $[0, 5]$, with a right hand side coefficient in the range $[5, 15]$. We generated problems and send them to `Couenne`, for those problems which are feasible in `Couenne`, we use $[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$ to directly approximate Example 6 and $[\mathbf{QP}_{DNN}]$ to approximate the QCQP reformulation of Example 6. Note that the convexity of these problems is not tested. Results are shown in Table 1.5.5, and we can clearly see that relaxation $[\mathbf{QP}_{DNN}]$ fail to give a valid bound for instances 1,3,6,7,8 and 10, while tensor relaxation $[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$ can provide a valid lower bound for all tested instances.

Test No.	`Couenne`	$[\mathbf{TP}\text{-}\mathcal{K}^{DNN}]$	$[\mathbf{QP}_{DNN}]$
1	-0.1790	-0.1852	Unbounded
2	10.9275	7.8888	0
3	-158.751	-245.7888	Unbounded
4	1.3041	1.1044	0
5	2.5418	1.9276	0
6	0.7107	-2.0031	Unbounded
7	1.0663	-6.6609	Unbounded
8	-8.0284	-56.0924	Unbounded
9	0.0275	0.0272	0
10	8.0032	2.4765	Unbounded

Table 1.5.5: Relaxation comparisons for Example 6

1.6 Conclusion

This chapter presents convex relaxations for general POPs over CP and CPSD tensor cones discussed in [64]. Bomze showed that completely positive matrix relaxation beats Lagrangian relaxations for quadratic programs with both linear and quadratic constraints in [18]. A natural question is whether similar results hold for general POPs that are not necessarily quadratic. Introducing CP and CPSD tensors to reformulate or relax general POPs, we generalize Bomze's results in QPs to general POPs, that is, the CP tensor relaxation beats Lagrangian relaxation bounds for general POPs with degree higher than 2. These results provide another way of using symmetric tensor cones to globally approximate nonconvex POPs. Burer in [27] showed that every quadratic programs with linear constraints and binary variables can be reformulated as CP programs and programs with quadratic constraints can be relaxed by CP programs, with approximation approaches for CP matrix programs. Note that one can reformulate general POPs as QPs by introducing additional variables and constraints and then apply Burer's results to obtain global bounds on general POPs. Pena et al. generalize Burer's results in [92] that under certain conditions a general POP can be reformulated as a conic program over CP tensors. A natural question is which reformulations or relaxations will provides better bounds for general POPs. In this paper, we showed that the bound of CP tensor relaxations is better than the bound of CP matrix relaxations for the quadratic reformulation of some classes of general POP. This validates the advantages of using tensor cones for convexification of nonconvex POPs . We also provide some tractable approximations of the CP tensor cone as well as CPSD tensor cone, which allows the possibility to compute the bounds of these tensor relaxations. Some preliminary numerical results on small scale POPs showed that these tensor cone approximations can provide bounds for global optimum of the original POPs. More importantly, in the experiments, the bounds obtained by CP or CPSD tensor cone programs yield better bounds than the CP or SDP matrix relaxations for quadratic reformulation of general POPs with similar computational efforts.

As future work we plan to characterize the classes of POPs in which the CP and CPSD tensor cone relaxations provide better bounds than the CP and PSD matrix relaxations for quadratic reformulations of POPs. Also, more POP instances with larger sizes can be tested and numerical comparisons on these more complicated POP cases can be made by developing appropriate code to address these problems.

Chapter 2

Alternative LP and SOCP Approaches for PO

2.1 Introduction

Many real-world problems can be modeled as a polynomial optimization problem (POP), thus devising new approaches to globally solve POPs is an active area of research [see, e.g., 5, 17, for recent surveys in this area]. In his seminal work, Lasserre [66] showed that *semidefinite programming* (SDP) [110] relaxations based on *sum of square* (SOS) polynomials [see, e.g., 17] can provide global bounds for POPs. However, the SDP constraints are computationally expensive and thus even using low-orders of the hierarchy to approximate large-scale POPs becomes computationally intractable in practice [69]. To improve the computational performance of the SDP based hierarchies to approximate the solution of POPs, prior work has focused on exploiting the problem's sparsity [60] and symmetry [34, 42], improving the relaxation by generating and adding appropriate valid inequalities [46], using bounded SOS polynomials [70] and more recently by devising more computationally efficient hierarchies such as linear programming (LP) and second-order cone programming (SOCP) hierarchies [2, 36, 37, 45, 91].

Here, we consider alternative ways to use SOCP restrictions of the SOS condition introduced by Ahmadi and Majumdar [2]. In particular, we show that SOCP hierarchies can be effectively used to strengthen hierarchies of LP relaxations for general POPs. Such hierarchies of LP relaxations have received little attention in the POP literature (a few noteworthy exceptions are [32, 36, 37, 68, 121]). However, in [61] we show that this solution approach is substantially more effective in finding solutions of certain POPs for which the more common hierarchies of SDP relaxations are known to perform poorly [see, e.g., 45]. Furthermore, when the feasible set of the POP is compact, these SOCP hierarchies converge to the POP's optimal value. Note that for the well-known SDP based hierarchies introduced by Lasserre [66], the *quadratic module* (QM) [5] associated with the feasible set of the POP is required to be *Archimedean* [17], which implies the compactness of the POP's feasible set.

The remainder of the chapter is organized as follows. We briefly review several convex approximations of POPs in Section 2.2. The proposed approximation strategies and hierarchies are presented in Section 2.3. Numerical results based on the proposed hierarchies are presented in Section 2.4. Section 2.5 provides some concluding remarks.

2.2 Preliminaries

The following notation is used throughout the chapter: let $\mathbb{R}_d[x] := \mathbb{R}_d[x_1, \ldots, x_n]$ be the set of polynomials in n variables with real coefficients of degree at most d. We define

$$SOS_{2d} := \left\{ \sum_{i=1}^{k} p_i(x)^2 : p_i(x) \in \mathbb{R}_d[x], k \in \mathbb{Z}_+ \right\},$$

as the cone of SOS polynomials in $\mathbb{R}_{2d}[x]$. For any $S \subseteq \mathbb{R}^n$, let $\mathcal{P}_d(S)$ be the cone of polynomials in $\mathbb{R}_d[x]$ of degree at most d that are non-negative over the set S [see, e.g., 17]. We consider the following general POP,

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \geq 0,\ i = 1, \ldots, m, \\ & x \in \mathbb{R}^n, \end{aligned} \tag{PP-P}$$

where the degree of the program is $d = \max\{\deg(f), \deg(g_1), \ldots, \deg(g_m)\}$. Given $S = \{x \in \mathbb{R}^n : g_i(x) \geq 0, i = 1, \ldots, m\}$, problem (PP-P) can be equivalently rewritten as the following *conic program* [see, e.g., 23],

$$\begin{aligned} \max_{x,\lambda} \quad & \lambda \\ \text{s.t.} \quad & f(x) - \lambda \in \mathcal{P}_d(S), \\ & x \in \mathbb{R}^n, \lambda \in \mathbb{R}. \end{aligned} \tag{PP-D}$$

In general, solving (PP-P) is NP-hard [85]. Problem (PP-D) is a (linear) conic program whose complexity is captured in the cone $\mathcal{P}_d(S)$, which is not *tractable* in general. Considering a sequence of *tractable* cones $\mathcal{K}^r \subseteq \mathcal{K}^{r+1} \subseteq \cdots \subseteq \mathcal{P}_d(S)$, then the following convex program

$$\begin{aligned} \max_{x,\lambda} \quad & \lambda \\ \text{s.t.} \quad & f(x) - \lambda \in \mathcal{K}^r, \\ & x \in \mathbb{R}^n, \lambda \in \mathbb{R} \end{aligned} \tag{2.1}$$

provides a lower bound for (PP-D), and hence a lower bound for (PP-P). Above, by *tractable* we mean that inclusion on the set can be expressed as a *linear matrix inequalities* (LMI) [5]. As r increases in (2.1), a tighter bound is achieved. The choice of the tractable cone $\mathcal{K}^r$ is a key factor in obtaining good approximation bounds for (PP-D), and in turn for (PP-P).

For this purpose, in his seminal work, Lasserre [66] proposed a hierarchy of LMI relaxations to approximate $\mathcal{P}_d(S)$, where

$$\mathcal{K}^r = \left\{ p(x) \in \mathbb{R}_{2r}[x] : p(x) = s_0(x) + \sum_{i=1}^{m} s_i(x) g_i(x), s_0(x) \in SOS_{2r}, s_i(x) \in SOS_{2\lfloor r - \deg(g_i)/2 \rfloor} \right\}, \tag{2.2}$$

and $r \geq \lceil d/2 \rceil$ is the level of the hierarchy. In this case, problem (2.1) is equivalent to

$$\begin{aligned} \max_{x, s_i(x)} \quad & \lambda \\ \text{s.t.} \quad & f(x) - \lambda = s_0(x) + \sum_{i=1}^{m} s_i(x) g_i(x), \\ & s_0(x) \in SOS_{2r}, s_i(x) \in SOS_{2\lfloor r - \deg(g_i)/2 \rfloor}, \quad i = 1, \dots, m, \\ & \lambda \in \mathbb{R}. \end{aligned} \tag{QM-SOS$_r$}$$

Problem (QM-SOS$_r$) can be reformulated as a SDP [see, e.g., 17]. Under some conditions related to the compactness of the set S (more precisely, when the *quadratic module* generated by the set of polynomials $\{g_1(x), \dots, g_m(x)\}$ is *Archimedean*), the hierarchy of problems (QM-SOS$_r$) converges to the global solution of (PP-P) as $r \to \infty$ [66]. However, as r increases, the size of the positive semidefinite matrices required to reformulate (QM-SOS$_r$) as a SDP increases exponentially. As a result, this approach is computationally expensive for large-scale problems [see, e.g., 44] or even for small-scale problems that require the solution of high levels of the hierarchy to obtain tight approximations of the POP of interest [see, e.g., 2, 46, 70].

Ahmadi and Majumdar [2] recently proposed a restriction of the SOS condition to address this shortcoming of the SDP-based hierarchies. The restriction of the SOS condition is done by introducing the use of *diagonally dominant sum of square* (DSOS) polynomials and *scaled diagonally dominant sum of square* (SDSOS) polynomials instead of SOS polynomials in (QM-SOS$_r$).

Definition 7 (DSOS polynomials [2])**.** *Let J be an index set, $m_i(x) \in \mathbb{R}_d[x]$ be a monomial* *for all $i \in J$, and $\alpha_i, \beta_{ij} \in \mathbb{R}_+$ for all $i, j \in J$. Then*

$$p(x) = \sum_i \alpha_i m_i(x)^2 + \sum_{i,j} \beta_{ij}(m_i(x) \pm m_j(x))^2, \tag{2.3}$$

is a DSOS polynomial in $\mathbb{R}_{2d}[x]$. Equivalently, DSOS polynomials can be defined as those that can be constructed from a diagonally dominant matrix *(DD). Namely, let $z(x)$ be a vector with the monomials $m_i(x)$ for all $i \in J$, and $Q \in \mathbb{R}^{|J|\times|J|}$ be a (symmetric) diagonally dominant matrix. Then $p(x) = z^T(x)Qz(x)$ is a DSOS.*

Let $DSOS_{2d}$ be the set of all DSOS polynomials in $\mathbb{R}_{2d}[x]$. Then it is clear from (2.3) that $DSOS_{2d} \subseteq SOS_{2d}$. Thus, using DSOS polynomials instead of SOS polynomials in (QM-SOS$_r$) provides a hierarchy of lower bounds for the SOS hierarchy. Moreover the resulting DSOS hierarchy is computationally easier to solve. Namely, recall that a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is DD if $A_{ii} \geq \sum_{j\neq i} |A_{ij}|, \forall i = 1, \ldots, n$. Thus the associated DSOS hierarchy

$$\begin{aligned} \max_{\lambda, d_i(x)} \quad & \lambda \\ \text{s.t.} \quad & f(x) - \lambda = d_0(x) + \sum_{i=1}^m d_i(x) g_i(x), \\ & d_0(x) \in DSOS_{2r}, d_i(x) \in DSOS_{2\lfloor r-\deg(g_i)/2\rfloor}, \\ & \lambda \in \mathbb{R}, \end{aligned} \tag{QM-DSOS$_r$}$$

can be reformulated as a LP. As proposed by Ahmadi and Majumdar [2], the DSOS hierarchy (QM-DSOS$_r$) can be strengthened by considering *scaled diagonally dominant sum of square* (SDSOS) polynomials.

Definition 8 (SDSOS polynomials [2])**.** *Let J be an index set, $m_i(x) \in \mathbb{R}_d[x]$ be a monomial for all $i \in J$, and $\alpha_i, \beta_i, \beta_j \in \mathbb{R}_+$ for all $i, j \in J$. Then*

$$p(x) = \sum_i \alpha_i m_i(x)^2 + \sum_{i,j} (\beta_i m_i(x) \pm \beta_j m_j(x))^2, \tag{2.4}$$

is a SDSOS polynomial in $\mathbb{R}_{2d}[x]$*. Equivalently, SDSOS polynomials can be defined as those that can be constructed from a* scaled diagonally dominant matrix *(SDD). Namely, let* $z(x)$ *be a vector with the monomials* $m_i(x)$ *for all* $i \in J$*, and* $Q \in \mathbb{R}^{|J|\times|J|}$ *be a (symmetric) scaled diagonally dominant matrix. Then* $p(x) = z^T(x)Qz(x)$ *is a SDSOS.*

Let $SDSOS_{2d}$ be the set of all SDSOS polynomial in $\mathbb{R}_{2d}[x]$. Then it is clear from (2.4) that $DSOS_{2d} \subseteq SDSOS_{2d} \subseteq SOS_{2d}$. Thus, using SDSOS polynomials instead of SOS polynomials in (QM-SOS$_r$) provides a hierarchy of lower bounds for the SOS hierarchy that is tighter than the (QM-DSOS$_r$) hierarchy. Moreover the resulting SDSOS hierarchy is computationally easier to solve than the associated SDP-based hierarchy. Namely, notice that a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is SDD if

$$A = \sum_{i,j\in\{1,\ldots,n\}} A^{ij}, \text{ for some } A^{ij} \succeq 0, \text{ with } A^{ij}_{kl} = 0 \text{ for any } k,l \in \{1,\ldots,n\}\setminus\{i,j\}. \quad (2.5)$$

Above, we use the common notation $A \succeq 0$ to indicate that the matrix is positive semidefinite. Notice that because the A^{ij} matrices in (2.5) have only nonzero elements at positions $k,l \in \{i,j\}$, then it follows that

$$A^{ij} \succeq 0 \iff A^{ij}_{ii} + A^{ij}_{jj} \geq \left\| \begin{pmatrix} 2A^{ij}_{ij} \\ \\ A^{ij}_{ii} - A^{ij}_{jj} \end{pmatrix} \right\|_2 \iff \begin{pmatrix} A^{ij}_{ii} + A^{ij}_{jj} \\ 2A^{ij}_{ij} \\ A^{ij}_{ii} - A^{ij}_{jj} \end{pmatrix} \in \mathcal{L}^3, \quad (2.6)$$

where $\mathcal{L}^n$ denotes the *second-order cone* or *Lorentz cone* of dimension n [see, e.g., 23]. Thus the associated SDSOS hierarchy

$$\begin{aligned} \max_{\lambda, d_i(x)} \quad & \lambda \\ \text{s.t.} \quad & f(x) - \lambda = d_0(x) + \sum_{i=1}^m d_i(x) g_i(x), \\ & d_0(x) \in SDSOS_{2r}, d_i(x) \in SDSOS_{2\lfloor r-\deg(g_i)/2\rfloor}, \\ & \lambda \in \mathbb{R}, \end{aligned} \qquad \text{(QM-SDSOS}_r\text{)}$$

can be reformulated as a second-order cone program. Ahmadi and Majumdar [2] have shown that the approximation hierarchies (QM-DSOS$_r$) and (QM-SDSOS$_r$) can be successfully used to approximate POPs arising in control, combinatorics, and general non-linear non-convex optimization [2]. Hierarchies (QM-DSOS$_r$) and (QM-SDSOS$_r$) are computationally easier to solve than (QM-SOS$_r$), however, their bounds might not be as good as the one obtained with the (QM-SOS$_r$) hierarchy of the same order [see, e.g., 62].

2.3 Alternative Hierarchies for Polynomial Optimization

Lasserre's hierarchy [66] has been shown to provide very tight bounds for multiple classes of POPs. However, this approach becomes computationally intractable for large-scale problems or even for small-scale problems that require the solution of high levels of the hierarchy to obtain good approximations for the solution of the problem of interest. Loosely speaking, this intractability stems from the fact that the size of the SDP reformulation of the SOS conditions in (QM-SOS$_r$) grows exponentially with the dimension of the decision variables of the problem n, as well as the level of the hierarchy r.

A key building block behind the convergence properties of the hierarchy defined by (QM-SOS$_r$) is a representation theorem for polynomials in $\mathcal{P}_d(S)$ by Putinar [95] that makes use of SOS polynomials [see, e.g., 17, 66]. Other convergent SDP hierarchies can be constructed similarly using the representation theorem by Schmüdgen [101], when the set S is compact. Besides these SOS representation theorems, there are however well-known representations theorems for non-negative polynomials that use polynomials with non-negative coefficients (instead of SOS polynomials) in the representation. Examples of these are the representation theorem of Hardy et al. [49], when the set S is a polytope, and Pólya's Theorem [49], when the set $S = \mathbb{R}^n_+$.

Theorem 5 (Pólya [49])**.** *Let $p(x) \in \mathbb{R}^n[x]$ be a multivariate polynomial. Then $p(x) > 0$ for all $x \geq 0 \Rightarrow \left(1 + \sum_{i=1}^n x_i\right)^r p(x) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha x^\alpha$ for some $r \geq 0$, $c_\alpha \geq 0$ for all $\alpha \in \mathbb{N}^n$.*

In stating Theorem 5, we make use of the common notation $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for any $x \in \mathbb{R}^n$, and $\alpha \in \mathbb{N}^n$. Note that in Theorem 5, the non-negativity of the polynomial is certified using polynomials with non-negative coefficients. As a result, this type of representation theorems can be used to construct hierarchies of LP problems that converge to the optimal solution of (PP-P) (when the required conditions on the set S are satisfied). Such approach has been used in [32, 68, 121]. It is worthy to mention that Pólya's approach is also used in [36, 37], to address the solution of POPs.

Here, we take advantage of this type of computationally easier LP hierarchy approach to address the solution of certain classes of POPs for which the more common SDP hierarchy is known to perform poorly [see, e.g., 45]. In particular, we use a representation theorem for non-negative polynomials in a semi-algebraic set recently introduced in [91] to construct a converging hierarchy of LPs for POPs. Formally, consider the following optimization problem:

$$\begin{aligned} z_{r,LP} := \max_{\lambda, c_{\alpha,\beta}} \quad & \lambda \\ \text{s.t.} \quad & \left(1 + \sum_{i=1}^{n} x_i + \sum_{j=1}^{m} g_j(x)\right)^r (f(x) - \lambda) = \sum_{(\alpha,\beta)\in I} c_{\alpha,\beta} x^\alpha g(x)^\beta \\ & c_{\alpha,\beta} \in \mathbb{R}_+ \text{ for all } (\alpha,\beta) \in I, \\ & \lambda \in \mathbb{R}, \end{aligned} \tag{Po-LP$_r$}$$

where

$$I := \{(\alpha,\beta) \in \mathbb{N}^{n+m} : \|(\alpha,\beta)\|_1 \leq r \max\{\deg(g_i) : i = 1,\dots,m\} + \deg(f)\}.$$

By matching the coefficients of each monomial in the left-hand side and the right-hand side of equation (Po-LP$_r$), the resulting problem is a LP with decision variables $\lambda \in \mathbb{R}, c_{\alpha,\beta} \in \mathbb{R}_+$, for all $(\alpha,\beta) \in I$. Similar to the (QM-SOS$_r$) hierarchy, but under milder conditions, the resulting bound $z_{r,LP}$ of (Po-LP$_r$), obtained at each level of the hierarchy converges as r increases.

Theorem 6 (Peña et al. [91])**.** *Let $S = \{x \in \mathbb{R}^n_+ : g_i(x) \geq 0, i = 1, \dots, m\}$ be a compact set, then as $r \to \infty$, $z_{r,LP}$ converges to the global optimum of* (PP-P)*.*

Thus, Theorem 6 provides the convergence guarantee of the (Po-LP$_r$) hierarchy to the optimal value of (PP-P) with a compact feasible set in $\mathbb{R}_n^+$. This allows us to use LP techniques to globally solve non-convex problems. However, this type of LP approximations for POPs are known to provide very weak approximation bounds for the objective value of the POP of interest [see, e.g., 32, 68]. To address this, we next propose the use of DSOS, SDSOS and SOS polynomials with fixed degree (degree 2) instead of the non-negative constant $c_{\alpha,\beta}$ in the definition of the hierarchy (Po-LP$_r$).

For a general POP (PP-P) with feasible set $S \subseteq \mathbb{R}^n_+$, consider the following hierarchies of optimization problems:

$$\begin{aligned}
\max_{\lambda, p_{\alpha,\beta}} \quad & \lambda \\
\text{s.t.} \quad & \left(1 + \sum_{i=1}^n x_i + \sum_{j=1}^m g_j(x)\right)^r (f(x) - \lambda) = \sum_{(\alpha,\beta)\in I'} p_{\alpha,\beta}(x) x^\alpha g(x)^\beta, \\
& p_{\alpha,\beta}(x) \in SOS_2, \text{ for all } (\alpha,\beta) \in I', \\
& \lambda \in \mathbb{R},
\end{aligned} \tag{Po-SOS$_r$}$$

$$\begin{aligned}
\max_{\lambda, p_{\alpha,\beta}} \quad & \lambda \\
\text{s.t.} \quad & \left(1 + \sum_{i=1}^n x_i + \sum_{j=1}^m g_j(x)\right)^r (f(x) - \lambda) = \sum_{(\alpha,\beta)\in I'} p_{\alpha,\beta}(x) x^\alpha g(x)^\beta, \\
& p_{\alpha,\beta}(x) \in SDSOS_2, \text{ for all } (\alpha,\beta) \in I', \\
& \lambda \in \mathbb{R},
\end{aligned} \tag{Po-SDSOS$_r$}$$

$$\begin{aligned} \max_{\lambda, p_{\alpha,\beta}} \quad & \lambda \\ \text{s.t.} \quad & \left(1 + \sum_{i=1}^{n} x_i + \sum_{j=1}^{m} g_j(x)\right)^r (f(x) - \lambda) = \sum_{(\alpha,\beta)\in I'} p_{\alpha,\beta}(x) x^{\alpha} g(x)^{\beta}, \\ & p_{\alpha,\beta}(x) \in DSOS_2, \text{ for all } (\alpha,\beta) \in I', \\ & \lambda \in \mathbb{R}, \end{aligned} \tag{Po-DSOS$_r$}$$

where $r \geq 0$ and

$$I' := \{(\alpha,\beta) \in \mathbb{N}^{n+m} : \|(\alpha,\beta)\|_1 \leq r \max\{\deg(g_i) : i = 1, \ldots, m\} + \deg(f) - 2\}.$$

Similar to Lasserre's hierarchy (QM-SOS$_r$), problem (Po-SOS$_r$) can be reformulated as a SDP. In turn, similar to the hierarchies (QM-DSOS$_r$) and (QM-SDSOS$_r$) (cf., Section 2.2), the optimization problems (Po-DSOS$_r$) and (Po-SDSOS$_r$) can be reformulated as a LP and as a SOCP respectively. Note that in the hierarchies discussed in Section 2.2, as the level of the hierarchy r increases, the complexity of checking that a fixed number, $m + 1$, of polynomials are SOS, SDSOS, or DSOS increases. Instead in the hierarchy defined in (Po-SOS$_r$), (Po-SDSOS$_r$) and (Po-DSOS$_r$), the complexity of checking that the involved polynomials are SOS, SDSOS, or DSOS does not change as the degree of these polynomials is fixed to 2. Instead, it is the number of these polynomials that increases as the level of the hierarchy increases (a similar approach has been used in [70]). This turns out to be key to obtain the results presented later in next section on the performance of the hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$) and (Po-DSOS$_r$).

Clearly, the hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$), and (Po-DSOS$_r$) provide tighter bounds on the associated POP than the LP based hierarchy (Po-LP$_r$). As a result, under the same conditions of Theorem 6, these hierarchies will converge as $r \to \infty$ to the global optimal solution of (PP-P). Below, we state this formally.

Proposition 10. *Consider problem* (PP-P) *with a compact feasible region and assume that* $S \subseteq \mathbb{R}^n_+$ *whose global optimal objective function is* z^**, and let* $z_{r,DSOS}$*,* $z_{r,SDSOS}$*,* $z_{r,SOS}$ *be the optimal value of hierarchies* (Po-DSOS$_r$)*,* (Po-SDSOS$_r$) *and* (Po-SOS$_r$) *respectively, then it follows that for any* $r = 1, 2, \dots$*:*

$$z_{r,LP} \leq z_{r,DSOS} \leq z_{r,SDSOS} \leq z_{r,SOS} \leq z^*.$$

Moreover,

$$\lim_{r\to\infty} z_{r,DSOS} = \lim_{r\to\infty} z_{r,SDSOS} = \lim_{r\to\infty} z_{r,SOS} = z^*.$$

Proof. The inequalities $z_{r,DSOS} \leq z_{r,SDSOS} \leq z_{r,SOS}$ follow from $DSOS_{2d} \subseteq SDSOS_{2d} \subseteq SOS_{2d}$. It is easy to see $z_{r,LP} \leq z_{r,DSOS}$ since all the nonnegative constants belong to $DSOS_0$. By Theorem 6, $\lim_{r\to\infty} z_{r,LP} = z^*$ when the feasible region of (PP-P) is compact. Thus $\lim_{r\to\infty} z_{r,DSOS} = \lim_{r\to\infty} z_{r,SDSOS} = \lim_{r\to\infty} z_{r,SOS} = z^*$ follows. □

2.4 Numerical Results

To illustrate the performance of the hierarchies discussed in Section 2.3, we test the Lasserre-type hierarchies and the proposed hierarchies in this article on some relevant POP instances. We use APPS [43] together with `Matlab` to implement all the hierarchies. Numerical experiments are conducted on an AMD Opteron 2.0 GHz(x16) Linux machine with 32 GB memory. We use `MOSEK` [6] as the LP and SOCP solver. Also, we use `SeDuMi` [108] as the SDP solver, to exploit its well-known precision for solving SDPs.

Due to the different approach used in the Lasserre-type hierarchies and the hierarchies proposed in Section 2.3, with the same r, the degree of the polynomials involved in the problem might not be equal. Thus, to make it easier to compare the results obtained from each hierarchy, instead of reporting the hierarchy level r, we report the maximum degree $\hat{d}$ of the polynomials involved in the formulation as r increases in each of them.

In the tables of numerical results that follow, the symbol (*) indicates that the reported value is the optimal objective value of the problem. We use "T" as the solution time in seconds for each hierarchy and "Infeas." to indicate that the optimization problem is infeasible. The symbol ($\diamond$) indicates that the solver runs out of memory. Lastly the symbol ($\circ$) indicates that generating the program that matches coefficient in the hierarchy in `Matlab` runs out of memory.

2.4.1 Illustrative Examples

We begin by testing a set of POPs from [46], which are highly non-convex and require a high level of Lasserre's hierarchy to converge to their global optimum.

Example 7. *Consider the following quadratic POP with 5 variables:*

$$\begin{aligned} \min_{x\in\mathbb{R}^5} \quad & 2x_1 - x_2 + x_3 - 2x_4 - x_5 \\ \text{s.t.} \quad & (x_1-2)^2 - x_2^2 - (x_3-1)^2 - (x_5-1)^2 \geq 0, \\ & x_1x_3 - x_4x_5 + x_1^2 \geq 1, \\ & x_3 - x_2^2 - x_4^2 \geq 1, \\ & x_1x_5 - x_2x_3 \geq 2, \\ & x_1 + x_2 + x_3 + x_4 + x_5 \leq 14, \\ & x_i \geq 0, i = 1, \ldots, 5. \end{aligned}$$

As shown in Table 2.4.1, the (QM-SOS$_r$) hierarchy converges to the global optimum when $\hat{d} = 8$ with a computational time of 49.82 seconds, while the hierarchy (Po-SOS$_r$) converges to global optimum when $\hat{d} = 6$ with only 8.21 seconds of computational time. Hierarchies (QM-SDSOS$_r$) and (QM-DSOS$_r$) fail to converge to the global optimum up to $\hat{d} = 8$. However, the hierarchy (Po-SDSOS$_r$) also converges to the global optimum when $\hat{d} = 8$ with 13.28 seconds of computational time. The hierarchy (Po-DSOS$_r$) provides a weaker bound than hierarchy (Po-SDSOS$_r$) and does not converge to the problem's global optimum when $\hat{d} = 8$.

Although the degree $\hat{d}$ provides an approximate measure of the size (variables and constraints) involved in the formulations of the hierarchies' problems, a better comparison of the hierarchies can be done by illustrating the trade-off between the solution time and the quality of the bound obtained from each hierarchy. In Figure 2.4.1 (left), the different line plots show the bound and solution time associated with increasing orders of each of the hierarchies. Clearly, within one second, the (Po-SDSOS$_r$) hierarchy gives the best bound; within ten seconds, the (Po-SOS$_r$) hierarchy gives the optimal value while there is still a gap between the problem's optimal value (illustrated by the dashed horizontal line) and the bounds obtained by other hierarchies. Clearly, the hierarchies proposed in Section 2.3 have better performance over the Lasserre-type hierarchies for this problem.

	(QM-SOS$_r$)		(QM-SDSOS$_r$)		(QM-DSOS$_r$)		(Po-SOS$_r$)		(Po-SDSOS$_r$)		(Po-DSOS$_r$)	
$\hat{d}$	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T
2	-25.00	0.35	-25.00	0.12	-25.00	0.01	-6.63	0.74	-7.40	0.03	-25.00	0.02
4	-6.01	1.22	-6.35	0.15	-25.00	0.09	-2.35	1.53	-2.96	0.19	-6.14	0.05
6	-2.40	6.75	-4.46	1.85	-14.39	1.46	*-1.57	8.21	-1.72	0.71	-2.93	0.74
8	*-1.57	49.82	-2.81	15.00	-7.49	18.62			*-1.57	13.28	-1.86	15.49

$*$: Optimal value is obtained.

Table 2.4.1: Bound and time comparison of different hierarchies for Example 7.

In Table 2.4.1, note that for the same level of hierarchies (Po-DSOS$_r$) and (Po-SDSOS$_r$), the linear representation of $DSOS_2$ introduces more decision variables than the SOCP representation of $SDSOS_2$. This explains why the running time of the LP-based hierarchy can be larger than the running time of the SOCP-based hierarchy.

Example 8. *Consider the following quadratic POP with 10 variables:*

$$
\begin{aligned}
\min_{x\in\mathbb{R}^{10}} \quad & -x_1 - x_2 + x_3 - 2x_4 - x_5 + x_6 + x_7 - x_8 + x_9 - 2x_{10} \\
\text{s.t.} \quad & (x_3-2)^2 - (x_5-1)^2 - 2x_6 + x_8^2 - (x_9-2)^2 \geq -4, \\
& -x_2^2 + x_3x_{10} - x_4^2 + x_6x_7 \geq 1, \\
& x_1x_8 - x_2x_3 + x_4x_7 - x_5x_{10} \geq 2, \\
& \sum_{i=1}^{10} x_i \leq 5, \\
& x_i \geq 0, i = 1, \dots, 10.
\end{aligned}
$$

As shown in Table 2.4.2, Lasserre's hierarchy (QM-SOS$_r$) and hierarchy (QM-SDSOS$_r$) converge to the global optimum at the third level when $\hat{d} = 6$ with a computational time of 2369.50 seconds and 72.43 seconds respectively. In contrast, hierarchy (Po-SOS$_r$) converges to the global optimum when $\hat{d} = 4$ with 8.27 seconds of computational time. The hierarchy (Po-SDSOS$_r$) also converges to the global optimum when $\hat{d} = 4$ with 2.23 seconds of computational time. Similar to Example 1, hierarchies (QM-DSOS$_r$) and (Po-DSOS$_r$) provide the weakest bound and the problem's global optimum is not reached by $\hat{d} = 6$, but the hierarchy (Po-DSOS$_r$) provides tighter bounds with less computational time than the hierarchy (QM-DSOS$_r$) at each level.

As discussed previously, a better comparison among the different hierarchies can be obtained by illustrating the trade-off between the solution time and the quality of the bound obtained from each hierarchy. In Figure 2.4.1 (right), the different line plots show the bound and solution time associated with increasing orders of each of the hierarchies. Notice that within one second, the (Po-DSOS$_r$) gives the best bound. Also, within ten seconds, only the (Po-SOS$_r$) and (Po-SDSOS$_r$) hierarchies obtain the problem's optimal value (illustrated by the dashed horizontal line), and the (Po-SDSOS$_r$) hierarchy takes less computational time than the (Po-SOS$_r$) hierarchy.

	(QM-SOS$_r$)		(QM-SDSOS$_r$)		(QM-DSOS$_r$)		(Po-SOS$_r$)		(Po-SDSOS$_r$)		(Po-DSOS$_r$)	
$\hat{d}$	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T
2	-10.00	0.09	-10.00	0.04	-10.00	0.02	-7.76	0.17	-7.76	0.04	-10.00	0.02
4	-7.76	25.89	-7.76	1.34	-10.00	0.47	*-5.18	8.27	*-5.18	2.23	-5.59	0.16
6	*-5.18	2369.50	*-5.18	72.43	-8.28	63.09					-5.19	35.29

$*$: Optimal value is obtained.

Table 2.4.2: Bound and time comparison of different hierarchies for Example 8.

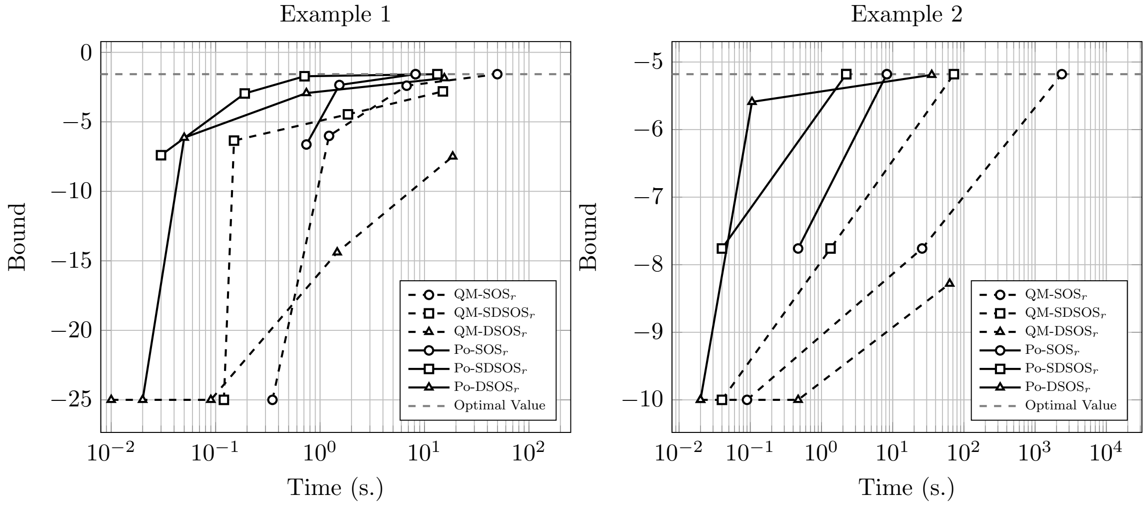

Figure 2.4.1: Bound and time comparison of different hierarchies for Example 7 (left) and Example 8 (right).

Example 9. *Consider the following quadratic POP with 15 variables:*

$$
\begin{aligned}
\min_{x\in\mathbb{R}^{15}} \quad & x_1 - x_2 + x_3 - x_4 - x_5 + x_6 + x_7 - x_8 + x_9 - x_{10} + x_{11} - x_{12} + x_{13} - x_{14} + x_{15} \\
\text{s.t.} \quad & (x_1-2)^2 - x_2^2 + (x_3-2)^2 - (x_4-1)^2 - (x_5-1)^2 + (x_6-2)^2 - (x_7-1)^2 - x_8^2 \\
& - (x_9-2)^2 - (x_{10}-1)^2 + x_{11}^2 - x_{12}^2 + (x_{13}-2)^2 + x_{14}^2 - (x_{15}-1)^2 \geq 0, \\
& - x_1x_7 - x_4x_5 - x_3^2 + x_6x_9 + x_{10}x_{12} \geq 1, \\
& x_2x_3 - x_8x_{11} - x_{14}^2 + x_5x_{15} \geq 2, \\
& \sum_{i=1}^{15} x_i \leq 10, \\
& x_i \geq 0, i = 1, \ldots, 15.
\end{aligned}
$$

	(QM-SOS$_r$)		(QM-SDSOS$_r$)		(QM-DSOS$_r$)		(Po-SOS$_r$)		(Po-SDSOS$_r$)		(Po-DSOS$_r$)	
$\hat{d}$	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T
2	-10.00	0.09	-10.00	0.02	-10.00	0.02	-8.07	0.27	-8.74	0.06	-10.00	0.02
4	-8.06	2754.30	-8.29	10.17	-10.00	2.27	*-7.43	640.60	*-7.43	59.85	-8.22	0.51
6	○	○	○	○	○	○					-7.64	2340.00

$*$: Optimal value is obtained.

○: `Matlab` runs out of memory while formulating LMI.

Table 2.4.3: Bound and time comparison of different hierarchies for Example 9.

The results for Example 3 are shown in Table 2.4.3. Lasserre's hierarchy (QM-SOS$_r$) and the hierarchy (QM-SDSOS$_r$) fail to provide the problem's global optimal value when $\hat{d} = 4$. `Matlab` runs out of memory when generating the LMI for Lasserre-type hierarchies when $\hat{d} = 6$. In contrast, hierarchies (Po-SOS$_r$) and (Po-SDSOS$_r$) converge to the global optimum when $\hat{d} = 4$ with 640.60 and 59.85 seconds of computational time respectively. Similar to Example 1 and 2, hierarchies (QM-DSOS$_r$) and (Po-DSOS$_r$) provide the weakest bound and the problem's global optimum is not reached when $\hat{d} = 6$. However, the (Po-DSOS$_r$) hierarchy provides tighter bounds with less computational time than (QM-DSOS$_r$) when $\hat{d} = 2$ and $\hat{d} = 4$.

2.4.2 Numerical Results on Global Optimization Library

Next, we compare Lasserre-type hierarchies with the proposed hierarchies on some problems from the GLOBAL Library available at `http://www.gamsworld.org/global/globallib.htm`. These problems have been used as benchmark in [59, 114, 115].

In Figure 2.4.2, we show the performance of different hierarchies for problem `ex2_1_1` and problem `ex3_1_4`. Similar to Figure 2.4.1, the different line plots show the bound and and solution time associated with increasing orders of each of the hierarchies. Clearly, for problem `ex2_1_1`, within one second, the (Po-SOS$_r$) and (Po-SDSOS$_r$) hierarchies give the optimal value while the bounds obtained by other hierarchies is not tight. Overall, the (Po-SDSOS$_r$) has the best performance in terms of bound and computational time for

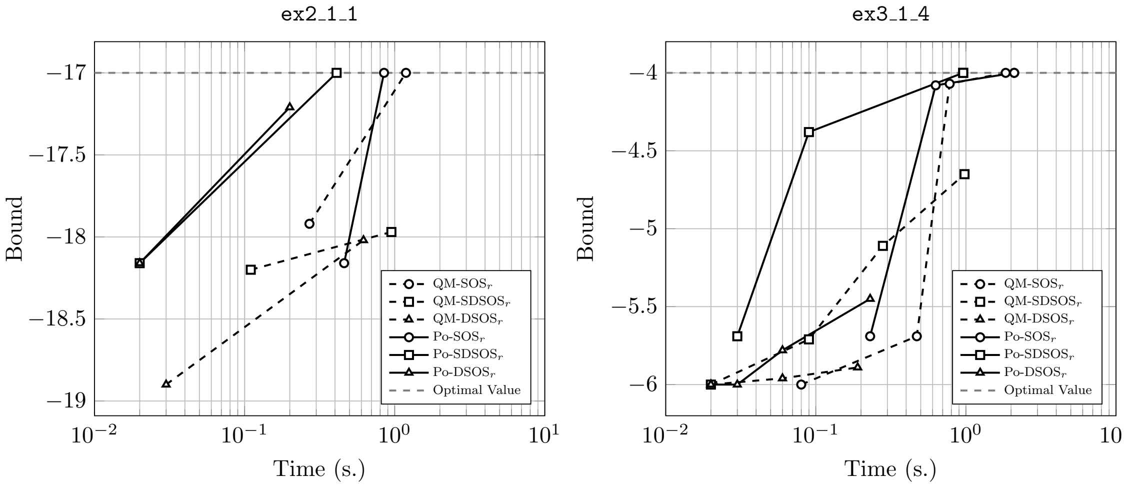

Figure 2.4.2: Bound and time comparison of different hierarchies for ex2_1_1 (left) and ex3_1_4 (right).

problem `ex2_1_1`. For problem `ex3_1_4`, within one second, only the (Po-SDSOS$_r$) reaches the optimal value, again, the (Po-SDSOS$_r$) has the best performance in terms of bound and computational time for problem `ex3_1_4`.

Table 2.4.4 shows the bound and time comparison of all hierarchies applied to different test problems. Column 1 shows the name of the problem and column 2 states the number of variables in the problem and its degree, while the degree of each hierarchy $\hat{d}$ is listed in column 3. The results for the Lasserre-type hierarchies are given in columns 4-9 while the remaining columns show the results for the proposed hierarchies in Section 2.3. We can see that for problems `ex_2_1_2`, `ex_2_1_3`, `ex2_1_4`, and `ex2_1_5`, the Lasserre-type hierarchies (QM-SOS$_r$), (QM-SDSOS$_r$), and (QM-DSOS$_r$) are infeasible when $\hat{d} = 2$ and provide the optimal solution when $\hat{d} = 4$. In contrast, the proposed hierarchies give the optimal value when $\hat{d} = 2$.

			(QM-SOS$_r$)		(QM-SDSOS$_r$)		(QM-DSOS$_r$)		(Po-SOS$_r$)		(Po-SDSOS$_r$)		(Po-DSOS$_r$)	
ID	(n,d)	$\hat{d}$	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T	Bound	T
`ex2_1_1`	(5,2)	2	Infeas.	0.14	Infeas.	0.08	Infeas.	0.02	-18.16	0.46	-18.16	0.02	-18.16	0.02
		4	-17.92	0.27	-18.20	0.11	-18.90	0.03	*-17.00	0.85	*-17.00	0.41	-17.21	0.20
		6	*-17.00	1.19	-17.97	0.95	-18.02	0.62					-17.21	0.46
`ex2_1_2`	(6,2)	2	Infeas.	0.11	Infeas.	0.73	Infeas.	0.00	*-213.00	0.64	*-213.00	0.03	*-213.00	0.02
		4	*-213.00	0.46	*-213.00	0.25	*-213.00	0.06						
`ex2_1_3`	(13,2)	2	Infeas.	0.17	Infeas.	0.39	Infeas.	0.01	*-15.00	0.27	*-15.00	0.06	*-15.00	0.03
		4	*-15.00	108.88	*-15.00	2.35	-15.00	0.91						
`ex2_1_4`	(6,2)	2	Infeas.	0.08	Infeas.	0.37	Infeas.	0.37	*-11.00	0.76	*-11.00	0.03	*-11.00	0.02
		4	*-11.00	0.45	*-11.00	0.20	*-11.00	0.06						
`ex2_1_5`	(10,2)	2	Infeas.	0.13	Infeas.	0.75	Infeas.	0.01	*-268.01	0.26	*-268.01	0.08	*-268.01	0.03
		4	*-268.01	5.05	*-268.01	0.86	-268.06	0.84						
`ex2_1_6`	(10,2)	2	Infeas.	0.05	Infeas.	0.45	Infeas.	0.00	-39.83	0.39	-39.83	0.14	-39.83	0.06
		4	*-39.00	8.90	-39.86	0.86	-40.88	1.41	*-39.00	16.88	*-39.00	1.61	-39.52	0.85
`ex2_1_7`	(20,2)	2	Infeas.	0.28	Infeas.	0.36	Infeas.	0.02	-4331.16	1.10	-4332.53	0.61	-4332.53	0.16
		4	$\diamond$	$\diamond$	-4686.70	23.08	$-\infty$	8.79	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$	$\circ$
`ex2_1_10`	(20,2)	2	Infeas.	0.22	Infeas.	0.38	Infeas.	0.02	*49318.02	0.69	*49318.02	0.25	*49318.02	0.06
		4	$\diamond$	$\diamond$	*49318.02	22.92	$-\infty$	7.70						
`ex3_1_3`	(6,2)	2	Infeas.	0.11	Infeas.	0.74	Infeas.	0.00	*-310.00	0.67	*-310.00	0.02	*-310.00	0.02
		4	*-310.00	0.71	*-310.00	0.38	$-\infty$	0.08						
`ex3_1_4`	(3,2)	2	-6.00	0.08	-6.00	0.02	-6.00	0.02	-5.69	0.23	-5.69	0.03	-6.00	0.02
		4	-5.69	0.47	-5.71	0.09	-6.00	0.02	-4.08	0.63	-4.38	0.09	-6.00	0.03
		6	-4.07	0.78	-5.11	0.28	-5.96	0.06	*-4.00	2.10	*-4.00	0.96	-5.78	0.06
		8	*-4.00	1.84	-4.65	0.98	-5.89	0.19					-5.45	0.23
`ex4_1_9`	(2,4)	4	-7.00	0.15	-7.00	0.04	-7.00	0.02	*-5.51	0.20	*-5.51	0.04	-7.00	0.02
		6	*-5.51	0.23	*-5.51	0.07	-7.00	0.02					$\circ$	$\circ$

$*$: Optimal value is obtained.

$\diamond$: Solver runs out of memory.

$\circ$: `Matlab` runs out of memory while formulating LMI.

Table 2.4.4: Bound and time comparison of different hierarchies for examples in Global Optimization Library.

The time to obtain the optimal value is greatly reduced by using (Po-SOS$_r$) over (QM-SOS$_r$) for problem ex_2_1_3. For problems ex_2_1_7 and ex2_1_10 with relatively large numbers of variables (20 variables), the Lasserre-type hierarchies are infeasible when $\hat{d} = 2$, which means that the Lasserre-type hierarchies fail to give a bound, however, by using the hierarchies proposed here, the optimal value for problem ex2_1_10 and a global lower bound for problem ex_2_1_7 are obtained when $\hat{d} = 2$. The hierarchy (QM-SOS$_r$) runs out of memory when $\hat{d} = 4$ for problems ex_2_1_7 and ex2_1_10. For ex_2_1_7, the hierarchy (QM-SDSOS$_r$) gives a bound when $\hat{d} = 4$; however, it is weaker than the ones obtained from the hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$) and (Po-DSOS$_r$) when $\hat{d} = 2$. Matlab runs out of memory when formulating the LMI for the hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$) and (Po-DSOS$_r$) when $\hat{d} = 4$ for ex_2_1_7, due to a large number of constraints in ex_2_1_7. For other cases, our proposed hierarchies mostly converge to global optimum with a smaller $\hat{d}$ than that of Lasserre-type hierarchies.

From Table 2.4.4, one can notice that in some instances of the problems, there is no improvement in the bound obtained by using the sequentially tighter (Po-SOS$_r$), (Po-SDSOS$_r$) and (Po-DSOS$_r$) hierarchies. For some problems (like ex2_1_2), this is due to the hierarchy (Po-DSOS$_r$) providing the problem's optimal solution. For other problems (like ex2_1_1), this is a result of the structure of the problem, which results in hierarchies (Po-SOS$_r$) and (Po-SDSOS$_r$) not helping to improve the bounds. For example, for problem ex2_1_1, the objective function is given by $f(x_1, \ldots, x_5) = 42x_1 + 44x_2 + 45x_3 + 47x_4 + 47.5x_5 - 50(x_1^2 + x_2^2 + x_2^2 + x_4^2 + x_5^2)$. Since there are no cross-variable terms, the bound obtained from the hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$) does not take advantage of providing a tighter formulation of the POP for cross-variable monomials. On the other hand, it is clear that in problems like ex3_1_4, the tighter (Po-SOS$_r$), (Po-SDSOS$_r$) hierarchies provide better bounds than the (Po-DSOS$_r$) hierarchy.

2.4.3 Numerical Results on Problems with More Variables

Consider the following non-convex problem,

$$\begin{aligned} \min \quad & \sum_{|\alpha|\leq 2} c_\alpha x^\alpha \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i^2 \leq 1, \\ & \sum_{i=1}^{n} x_i^2 \geq 0.6^2, \\ & x_i \geq 0, i = 1, \ldots, n, \end{aligned} \tag{2.7}$$

where c_α are randomly generated in $[-1, 1]$. We construct this nonconvex problem inspired by [116] to test the performance of the proposed hierarchies. The problem is to find the minimal value of a polynomial on the Euclidean unit ball intersected with the positive orthant while excluding the Euclidean ball with radius 0.6. We use instances with relatively large number of variables $n \in [20, 30, 50, 100, 150]$ and compare the Lasserre-type hierarchies with the hierarchies proposed in Section 2.3. Note that it will be computationally expensive to run higher levels of the hierarchies for large-scale problems. The purpose of this comparison mainly focuses on the bound obtained for quadratic programs when $\hat{d} = 2$.

Similar to the figures in previous sections, in Figure 2.4.3, we plot the performance of all hierarchies for problem (2.7) with $n = 20$. Clearly, only the (Po-SOS$_r$) obtains the optimal value with $n = 20$. Unlike the instances in Section 2.4.1 and Section 2.4.2, the increasing order of the hierarchies doesn't improve the bound significantly, thus the lines in Figure 2.4.3 are flat.

Table 2.4.5 lists the bound and time of all hierarchies for problem (2.7) with different n. The results for $n \geq 30$ and $\hat{d} = 4$ are not listed since `Matlab` runs out of memory when formulating the LMI for these cases. Column 2 is the upper bound we obtain from a global optimization solver, BARON [98]; columns 4-15 list the results by running Lasserre-type hierarchies and the proposed hierarchies. The optimal value (indicated by $*$) is obtained when the upper bound obtained by BARON is equal to any of the lower bound from the six hierarchies. It is clear that our proposed hierarchies (Po-SOS$_r$), (Po-SDSOS$_r$), and (Po-DSOS$_r$) yield tighter bounds than corresponding the Lasserre-type hierarchies. For cases with $n = 20, 30, 50, 100$, the hierarchy (Po-SOS$_r$) converges to the global optimum when $\hat{d} = 2$. For the case with $n = 150$, SOS-based hierarchies (QM-SOS$_r$) and (Po-SOS$_r$) fail to give a bound due to the computationally difficult SDP constraints. SOCP-based hierarchies can be used to obtain global bounds, in which case our proposed hierarchy (Po-SDSOS$_r$) improves the bounds obtained from (QM-SDSOS$_r$) by approximately 100%. LP-based hierarchies provide the worst bounds among the same type of hierarchies, however, the bound obtained by the LP-based hierarchy (Po-DSOS$_r$) is even tighter than the bound obtained by SOCP-based hierarchy (QM-SDSOS$_r$).

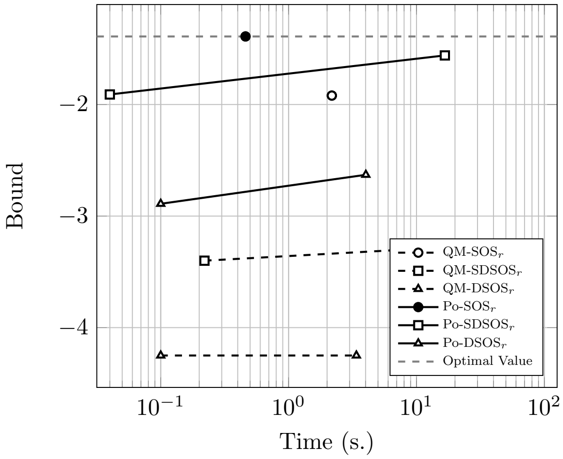

Figure 2.4.3: Bound and time comparison of different hierarchies for problem (2.7) with $n = 20$.

2.5 Concluding Remarks

In this paper, we propose alternative LP, SOCP and SDP approximation hierarchies to obtain global bounds for general POPs, by using SOS, SDSOS and DSOS polynomials to strengthen existing LP hierarchy for POPs. Comparing with the classic Lasserre's hierarchy, the LP and SOCP approximation hierarchies are shown to be computationally more

	BARON		(QM-SOS$_r$)		(QM-SDSOS$_r$)		(QM-DSOS$_r$)		(Po-SOS$_r$)		(Po-SDSOS$_r$)		(Po-DSOS$_r$)	
(n, d)	UB	$\hat{d}$	LB	T	LB	T	LB	T	LB	T	LB	T	LB	T
(20,2)	*-1.39	2	-1.92	2.18	-3.40	0.22	-4.25	0.01	*-1.39	0.46	-1.91	0.04	-2.89	0.01
		4	◇	◇	-3.29	11.99	-4.25	3.39			-1.56	16.67	-2.63	4.02
(30,2)	*-1.96	2	-2.21	2.61	-4.82	0.16	- 6.40	0.01	*-1.96	2.08	-2.91	0.06	-4.93	0.01
(50,2)	*-2.06	2	-2.90	43.99	-8.30	1.82	-11.38	0.38	*-2.06	50.26	-4.32	0.15	-6.38	0.01
(100,2)	*-3.02	2	-3.89	1606.00	-16.56	1.12	-20.47	0.01	*-3.02	1824.30	-8.49	0.45	-11.92	0.04
(150,2)	-3.58	2	◇	◇	-25.02	0.90	-29.12	0.04	◇	◇	-12.66	1.11	-16.42	0.09

LB,UB: lower bound, upper bound.

$*$: Optimal value is obtained.

◇: Solver runs out of memory.

Table 2.4.5: Bound and time comparison of different hierarchies for problem (2.7).

efficient to find the global optimum of POPs for which Lasserre's hierarchy is known to perform poorly. In particular, this shows that the relaxation approach introduced by Ahmadi and Majumdar [2] produces better results as a way to strengthen LP-based hierarchies for POPs. Furthermore, these hierarchies are shown to converge as the level of the hierarchy increases to the global optimum of the corresponding POP. Unlike other hierarchies proposed in the literature, this property is obtained whenever the feasible set of the POP is compact but the *quadratic module* of the polynomials defining the problem's feasible set is not necessarily*Archimedean*.

The fact that the hierarchies considered here are based on using LP and SOCP allows for the future use of column generation approaches in order to be able to address the solution of larger-scale POPs.

Chapter 3

Alternative SOCP Hierarchies for ACOPF Problems

3.1 Introduction

The Alternating Current Optimal Power Flow (ACOPF) problem is a challenging optimization problem in power systems. Several optimization methods have been devised to efficiently solve different variations of this problem. However, the ACOPF problem continues to be challenging. This is mainly due to the large-scale size of real systems and the nonlinearities and the nonconvexities in the underlying formulation.

One solution approach that has been actively investigated to address the global solution of the ACOPF relies on reformulating the problem as a polynomial program (PP); that is an optimization problem whose objective and constraints can be written in terms of polynomials on the decision variables. Then, a hierarchy of *semidefinite programming* (SDP) [cf., 5] relaxations can be used to obtain globally optimal solutions under mild conditions [44, 57, 82, 83]. This hierarchy of SDP relaxations is based on the seminal work of Lasserre [66], who showed that SDP relaxations based on *sum of square* (SOS) polynomials can provide global bounds for a general class of PPs. However, the associated SDP relaxations are computationally expensive and thus, even using the hierarchy's low-order relaxations to approximate large-scale PPs like the ACOPF becomes computationally intractable in practice [69].

To improve the computational performance of the SDP based hierarchies, prior work has focused on exploiting the problem structure through sparsity [60] and symmetry [34], improving the relaxation through the generation of valid inequalities [46], and more recently through devising more computationally efficient hierarchies based on the use of linear programming (LP) and second-order cone programming (SOCP) relaxations of the problem of interest [2, 45, 91].

Rather than using the computationally expensive SDP hierarchy, in [62] and [65], we explore the use of LP and SOCP hierarchies for solving the ACOPF problem and show that the SOCP approach that is proposed in this paper can be used to obtain solutions for ACOPF within limited computational time. Furthermore, we show that the first-order SOCP hierarchy obtained by weakening the more common hierarchy of SDP relaxations for this problem is equivalent to solving the conic dual of the SOCP approximations of ACOPF that have been recently proposed in [22, 54, 105], which provide the optimal solution of the ACOPF problem for special network topologies. In turn, the SOCP hierarchy approach provides a natural hierarchy of increasingly stronger SOCP approximations for the ACOPF problem.

Following the notations and concepts introduced in Section 2.2 in Chapter 2, we briefly introduce polynomial programming approaches for ACOPF as well as applying the proposed LP and SOCP based hierarchies in Chapter 2 to the ACOPF problem in next section.

3.2 ACOPF Formulation

In this section, we apply the LP and SOCP hierarchies to the alternating current optimal flow problem and we exploit the network structure of electricity transmission grids to specialize the proposed SOCP approximations for the ACOPF problem and obtain improving results. The results are then compared to the SDP based hierarchy [66] and to the SOCP relaxation that is proposed in [22].

3.2.1 ACOPF problem as a Polynomial Program

To formulate the ACOPF problem as a polynomial problem, we follow the same notation as in [71, 83], that is, we consider an undirected graph $P(N, E)$ where each vertex $n \in N$ is called a "bus" and each edge $e \in E$ is called a "branch". The subset $G \subseteq N$ denotes the set of generators. Additionally, we define the following parameters:

- P_k^-, P_k^+, Q_k^-, Q_k^+, V_k^- and V_k^+ are respectively the limits on active and reactive generation capacity and the absolute value of the voltage at bus k.
- S_{lm}^+ is the limit on the absolute value of the apparent power of a branch $(l, m) \in E$.
- P_k^d and Q_k^d are the active and reactive power demand respectively.
- c_k^2, c_k^1, c_k^0 are nonnegative coefficients for the power generation cost function.

We also define the following decision variables:

- P_k^g and Q_k^g: active and reactive power generated at bus k.
- P_{lm} and Q_{lm}: active and reactive power flow on branch (l, m).

Given a complex voltage V_i at bus i, let $\Re V_i$ denote the real part of V_i and $\Im V_i$ denote the imaginary part. The power flow equations are

$$P_k^g = P_k^d + \Re V_k \sum_{i=1}^{n} (\Re y_{ik} \Re V_i - \Im y_{ik} \Im V_i) + \Im V_k \sum_{i=1}^{n} (\Im y_{ik} \Re V_i - \Re y_{ik} \Im V_i),$$

$$Q_k^g = Q_k^d + \Re V_k \sum_{i=1}^{n} (-\Im y_{ik} \Re V_i - \Re y_{ik} \Im V_i) + \Im V_k \sum_{i=1}^{n} (\Re y_{ik} \Re V_i - \Im y_{ik} \Im V_i),$$

$$P_{lm} = b_{lm}(\Re V_l \Im V_m - \Re V_m \Im V_l) + g_{lm}(\Re V_l{}^2 + \Im V_m{}^2 - \Im V_l \Im V_m - \Re V_l \Re V_m),$$

$$\begin{aligned} Q_{lm} = {} & b_{lm}(\Re V_l \Im V_m - \Im V_l \Im V_m - \Re V_l{}^2 - \Im V_l{}^2) + g_{lm}(\Re V_l \Im V_m - \Re V_m \Im V_l - \Re V_m \Im V_l) \\ & - \frac{\bar{b}_{lm}}{2}(\Re V_l{}^2 + \Im V_l{}^2). \end{aligned}$$

Additionally, the network admittance matrix and other related matrices are defined as follows:

$$y_k = e_k e_k^T y,$$

$$y_{lm} = (j\frac{\bar{b}_{lm}}{2} + g_{lm} + jb_{lm})e_l e_l^T - (g_{lm} + jb_{lm})e_l e_m^T,$$

$$Y_k = \frac{1}{2}\begin{bmatrix} \Re(y_k + y_k^T) & \Im(y_k^T - y_k) \\ \Im(y_k - y_k^T)) & \Re(y_k + y_k^T) \end{bmatrix},$$

$$\bar{Y}_k = -\frac{1}{2}\begin{bmatrix} \Im(y_k + y_k^T) & \Re(y_k - y_k^T) \\ \Re(y_k^T - y_k) & \Im(y_k + y_k^T) \end{bmatrix},$$

$$M_k = \begin{bmatrix} e_k e_k^T & 0 \\ 0 & e_k e_k^T \end{bmatrix},$$

$$Y_{lm} = \frac{1}{2}\begin{bmatrix} \Re(y_{lm} + y_{lm}^T) & \Im(y_{lm}^T - y_{lm}) \\ \Im(y_{lm} - y_{lm}^T) & \Re(y_{lm} + y_{lm}^T) \end{bmatrix},$$

$$\bar{Y}_{lm} = -\frac{1}{2}\begin{bmatrix} \Im(y_{lm} + y_{lm}^T) & \Re(y_{lm}^T - y_{lm}) \\ \Re(y_{lm}^T - y_{lm}) & \Im(y_{lm} + y_{lm}^T) \end{bmatrix}.$$

where e_k is the k^{th} standard basis vector in $\mathbb{R}^{|N|}$. Let $x := [\Re V_k \quad \Im V_k]^T$ be the vector of variables in addition to the variables P_k^g, P_{lm} and Q_{lm}. The ACOPF can be formulated as the following degree-2 PP:

$$
\begin{aligned}
\min \quad & \sum_{k\in G} \left(c_k^2 (P_k^g)^2 + c_k^1 (P_k^d + \mathrm{tr}(Y_k x x^T)) + c_k^0 \right) && (3.1)\\
\text{s.t.} \quad & P_k^- \leq \mathrm{tr}(Y_k x x^T) + P_k^d \leq P_k^+,\ \forall k \in G, && (3.2)\\
& Q_k^- \leq \mathrm{tr}(\bar{Y}_k x x^T) + Q_k^d \leq Q_k^+,\ \forall k \in G, && (3.3)\\
& (V_k^-)^2 \leq \mathrm{tr}(M_k x x^T) \leq (V_k^+)^2,\ \forall k \in N, && (3.4)\\
& P_{lm}^2 + Q_{lm}^2 \leq (S_{lm}^+)^2,\ \forall (l,m) \in E, && (3.5)\\
& P_k^g = \mathrm{tr}(Y_k x x^T) + P_k^d,\ \forall k \in G, && (3.6)\\
& P_{lm} = \mathrm{tr}(Y_{lm} x x^T),\ \forall (l,m) \in E, && (3.7)\\
& Q_{lm} = \mathrm{tr}(\bar{Y}_{lm} x x^T),\ \forall (l,m) \in E. && (3.8)
\end{aligned}
\tag{PP-OPF}
$$

The objective function (3.1) minimizes the cost of power generation. Constraints (3.2), (3.3), and (3.4) set limits on the active power, reactive power, and voltage on each bus. Constraints (3.5) set a limit on the apparent power flow at every branch. Constraints (3.6) define the power generated while constraints (3.7), (3.8) define the active and reactive power flow. Problem (PP-OPF) is a PP of degree two and the proposed hierarchies are applicable to this problem. Additionally, the ACOPF problem exhibits a structure that can be exploited to improve the computational performance of the solution approach discussed here.

3.2.2 ACOPF Problem as a Second Order Cone Program

By exploring the structure of the quadratic formulation of the ACOPF, we have the following observations:

Proposition 11. *Let $\mathcal{X}_{ori}$ denote the set of x variables in* (PP-OPF) *and let $\mathcal{X}_{ext}$ denote the set of P_k^g, P_{lm}, and Q_{lm} variables. Then for $v_i \in \mathcal{X}_{ori}$ and $v_j \in \mathcal{X}_{ext}$, $v_i v_j$ does not* *appear in the quadratic model* (PP-OPF)*.*

Proposition 11 allows us to develop specific hierarchies of tractable conic programming relaxations that take advantage of the structure of ACOPF problem. Note that such structure is also exploited in the related conic programming relaxations for the ACOPF problem studied in [22].

Additionally, some product terms of variables in $\mathcal{X}_{ori}$ do not appear based on the network structure. For instance, buses i and j are connected if $(i, j) \in E$, and thus the network structure can be used to delete products of variables that are not needed in the formulation.

Proposition 12. *For any $k \in G$, the term $tr(Y_k xx^T)$ in the objective function* (3.1) *is equivalent to*

$$
\begin{aligned}
tr(Y_k xx^T) = \alpha_k \Re V_k^2 + \beta_k \Im V_k^2 + \sum_{j:(k,j)\in E} (\gamma_1^{kj} \Re V_k \Re V_k + \gamma_2^{kj} \Re V_k \Im V_j \\
+ \gamma_3^{kj} \Im V_k \Re V_j + \gamma_4^{kj} \Im V_k \Im V_j), \forall k \in N,
\end{aligned}
$$

where $\alpha_k, \beta_k, \gamma_1^{kj}, \gamma_2^{kj}, \gamma_3^{kj}, \gamma_4^{kj}$ are parameters.

Proof. Directly follows by exploring the structure of Y_k, $k \in G$. □

Proposition 12 uses the branch information in the network to simplify the hierarchies of tractable conic programming relaxations considered here by removing unnecessary monomial terms in the definitions of (QM-DSOS$_r$) and (QM-SDSOS$_r$). Only the cross product of voltages of two connected buses appears in tr$(Y_k xx^T)$ and the same goes for tr$(\bar{Y}_k xx^T)$, which means that there is no need to generate all the monomials in the right hand side of hierarchies (QM-DSOS$_r$) and (QM-SDSOS$_r$). Since electricity transmission grids are normally sparse graphs then the computational effort of solving the (QM-SDSOS$_r$) (or (QM-DSOS$_r$)) hierarchy is reduced significantly thanks to the reduction in the number of SOCP (or LP) constraints required in the hierarchy formulation. Additionally, the structure of the SOCP relaxation of ACOPF can be easily exploited since in each

SOCP constraint of (QM-SDSOS$_r$), only a combination of two monomials is involved. For example, if the cross product $m_i m_j$ does not appear in the model, then for $d_0(x)$ in hierarchy (QM-SDSOS$_r$) the SOCP constraint related to $(\beta_i m_i + \beta_j m_j)^2$ is not needed.

3.2.3 Duality in ACOPF Formulations

It has been shown in [44, Theorem 1] that the first level of the SDP hierarchy obtained by using the (QM-SOS$_r$) on (PP-OPF) is equivalent to the conic dual of the SDP relaxation for the ACOPF problem that is considered in [71, Optimization 3]. In this section, we show that the first level of the (QM-SDSOS$_r$) hierarchy of the ACOPF problem is the conic dual problem of the SOCP relaxation of the ACOPF problem that is presented in [22]. Intuitively, this result follows from the fact that the conic dual of the SDD matrices is the set of symmetric matrices $A \in \mathbb{R}^{n\times n}$ such that every 2×2 *principal submatrix* of A is positive semidefinite. Similar to (2.6), the latter condition can be formulated using the second-order cone, which corresponds to the relaxation approach used in [22]. Below, we state this result formally.

Theorem 7. *The first level of the SOCP hierarchy obtained by using the* (QM-SDSOS$_r$) *hierarchy on* (PP-OPF) *is equivalent to the conic dual of the SOCP relaxation for the ACOPF problem considered in [22].*

Proof. The basic idea is to derive the Lagrangian dual problem of the SOCP relaxation for the ACOPF problem considered in [22] and compare it with the first level of SOCP hierarchy discussed in this paper. First, problem (PP-OPF) is reformulated as a PP of the form given in (QM-SDSOS$_r$). We notice that since not all the monomials appear in the polynomial formulation then using the first level of the SOCP hierarchy ($r = 2$), one can approximate (PP-OPF) as

$$
\begin{aligned}
\max \quad & \varphi \\
\text{s.t.} \quad & \sum_{k \in N} \left(c_k^2 (P_k^g)^2 + c_k^1 (P_k^d + \operatorname{tr}(Y_k x x^T)) + c_k^0 \right) - \varphi = S(x) + S(P^g) + S(P, Q) \\
& \quad + \sum_{k \in N} \overline{\lambda}_k \left(P_k^+ - P_k^d - \operatorname{tr}(Y_k x x^T) \right) + \sum_{k \in N} \underline{\lambda}_k \left(-P_k^- + P_k^d + \operatorname{tr}(Y_k x x^T) \right) \\
& \quad + \sum_{k \in N} \overline{\gamma}_k \left(Q_k^+] - Q_k^d - \operatorname{tr}(\bar{Y}_k x x^T) \right) + \sum_{k \in N} \underline{\gamma}_k \left(-Q_k^- + Q_k^d + \operatorname{tr}(\bar{Y}_k x x^T) \right) \\
& \quad + \sum_{k \in N} \overline{\mu}_k \left((V_k^+)^2 - \operatorname{tr}(M_k x x^T) \right) + \sum_{k \in N} \underline{\mu}_k \left((-V_k^-)^2 + \operatorname{tr}(M_k x x^T) \right) \\
& \quad + \sum_{(l,m) \in L} a_{lm} \left((S_{lm}^+)^2 - P_{lm}^2 - Q_{lm}^2 \right) + \sum_{k \in G} b_k \left(P_k^g - \operatorname{tr}(Y_k x x^T) - P_k^d \right) \\
& \quad \sum_{(l,m) \in L} c_{lm} \left(P_{lm} - \operatorname{tr}(Y_{lm} x x^T) \right) + \sum_{(l,m) \in L} d_{lm} \left(Q_{lm} - \operatorname{tr}(\bar{Y}_{lm} x x^T) \right),
\end{aligned}
\tag{3.9}
$$

where $S(x)$ is an SDSOS polynomial based on the branch information, and $S(P^g), S(P, Q)$ are SDSOS polynomials as a function of P_k^g, P_{lm} and Q_{lm} respectively. Furthermore, by observation of the monomials in the objective function and constraints of (PP-OPF) , we remove some unnecessary monomials in $S(x), S(P^g)$ and $S(P, Q)$ for simplicity. To be more specific, the following terms are not needed:

$$
\begin{aligned}
& x_k, \forall k \in N \text{ in } S(x), \\
& P_i^g P_j^g, \forall i, j \in N, i \neq j \text{ in } S(P^g), \\
& P_{l_1, m_1} P_{l_2, m_2}, \forall (l_1, m_1), (l_2, m_2) \in E, (l_1, m_1) \neq (l_2, m_2), \\
& Q_{l_1, m_1} Q_{l_2, m_2}, \forall (l_1, m_1), (l_2, m_2) \in E, (l_1, m_1) \neq (l_2, m_2), \\
& P_{l_1, m_1} Q_{l_2, m_2}, \forall (l_1, m_1), (l_2, m_2) \in E \text{ in } S(P, Q).
\end{aligned}
$$

The variables $\alpha, \beta, \theta, \alpha_k, \beta_k, \theta_{lm}, \delta_{im}, \overline{\lambda}_k, \underline{\lambda}_k, \overline{\gamma}_k, \underline{\gamma}_k, \overline{\mu}_k, \underline{\mu}_k$ and a_{lm} are non-negative variables and α_{lm}^1, α_{lm}^2, β_k^1, β_k^2, θ_{lm}^1, θ_{lm}^2, δ_{lm}^1, δ_{lm}^2, b_k, c_{lm} and d_{lm} are free variables. As discussed in Section 2.2, an SDSOS polynomial has a natural second order cone representation, so we can write $S(x), S(P^g)$ and $S(P,Q)$ as

$$
\begin{aligned}
S(x) &= \alpha + \sum_{k \in N} \alpha_k x_k^2 + \sum_{(l,m)\in E} \begin{pmatrix} x_l \\ x_m \end{pmatrix}^T A_{lm} \begin{pmatrix} x_l \\ x_m \end{pmatrix}, \\
S(P^g) &= \beta + \sum_{k \in N} \beta_k (P_k^g)^2 + \sum_{k \in N} \begin{pmatrix} 1 \\ P_k^g \end{pmatrix}^T B_k \begin{pmatrix} 1 \\ P_k^g \end{pmatrix} \\
S(P,Q) &= \theta + \sum_{(l,m)\in E} (\theta_{lm} P_{lm}^2 + \delta_{lm} Q_{lm}^2) \\
&+ \sum_{(l,m)\in E} \Bigg(\begin{pmatrix} 1 \\ P_{lm} \end{pmatrix}^T C_{lm}^1 \begin{pmatrix} 1 \\ P_{lm} \end{pmatrix} + \begin{pmatrix} 1 \\ Q_{lm} \end{pmatrix}^T C_{lm}^2 \begin{pmatrix} 1 \\ Q_{lm} \end{pmatrix} \Bigg),
\end{aligned}
$$

Define η as $\eta = \alpha + \beta + \theta$, for ease of notation. By equating the coefficients of the monomials of Problem (3.9) and substituting some of the variables, we have the following second-order cone program:

$$
\begin{aligned}
\max\ & \sum_{k\in G} c_k^1 P_k^d - \sum_{k\in N} B_k^{(11)} - \eta - \sum_{(l,m)\in E} (C_{lm}^{1(11)} - C_{lm}^{2(11)}) + \sum_{k\in G} c_k^0 - \sum_{k\in N} \overline{\lambda}_k (P_k^+ - P_k^d) \\
& - \sum_{k\in N} \underline{\lambda}_k (-P_k^- + P_k^d) - \sum_{k\in N} \overline{\gamma}_k (Q_k^+ - Q_k^d) - \sum_{k\in N} \underline{\gamma}_k (-Q_k^- + Q_k^d) - \sum_{k\in N} \overline{\mu}_k (V_k^+)^2 \\
& + \sum_{k\in N} \underline{\mu}_k (V_k^-)^2 - \sum_{k\in G} 2B_k^{1(12)} P_k^d - \sum_{(l,m)\in L} (\theta_{lm} + C_{lm}^{1(22)})(S_{lm}^+)^2 \\
\text{s.t.}\ & A_{ij}^{(12)} = \sum_{(l,m)\in L} 2C_{lm}^{1(12)} Y_{lm}^{(ij)} + \sum_{(l,m)\in L} 2C_{lm}^{2(12)} \bar{Y}_{lm}^{(ij)} + \sum_{k\in N} c_k^1 Y_k^{(ij)} + \sum_{k\in N} \Big(\overline{\lambda}_k Y_k^{(ij)} \\
& -\underline{\lambda}_k Y_k^{(ij)} + \overline{\gamma}_k \bar{Y}_k^{(ij)} - \underline{\gamma}_k \bar{Y}_k^{(ij)} + 2B_k^{1(12)} Y_k^{(ij)} \Big),\ \forall (i,j) \in E, \\
& \alpha_i + \sum_{i:(i,j)\in E} A_{ij}^{(11)} + \sum_{j:(i,j)\in E} A_{ij}^{(22)} = \sum_{k\in N} c_k^1 Y_k^{(ii)} + \sum_{(l,m)\in L} \Big(2C_{lm}^{1(12)} Y_{lm}^{(ii)} + 2C_{lm}^{2(12)} \bar{Y}_{lm}^{(ii)} \Big) \\
& + \sum_{k\in N} \Big(\overline{\lambda}_k Y_k^{(ii)} - \underline{\lambda}_k Y_k^{(ii)} + \overline{\gamma}_k \bar{Y}_k^{(ii)} - \underline{\gamma}_k \bar{Y}_k^{(ii)} + 2B_k^{1(12)} Y_k^{(ii)} + \overline{\mu}_k - \underline{\mu}_k \Big),\ \forall i \in N, \\
& c_k^2 = \beta_k + B_k^{(22)}, \quad \forall k \in N, \\
& \theta_{lm} + C_{lm}^{1(22)} - \delta_{lm} - C_{lm}^{2(22)} = 0, \quad \forall (l,m) \in E, \\
& \eta \geq 0, \\
& \alpha_k, \beta_k, \overline{\lambda}_k, \underline{\lambda}_k, \overline{\gamma}_k, \underline{\gamma}_k, \overline{\mu}_k, \underline{\mu}_k \geq 0, \quad \forall k \in N, \\
& \theta_{lm}, \delta_{lm} \geq 0, \quad \forall (l,m) \in E, \\
& B_k \succeq 0, \quad \forall k \in N, \\
& A_{lm}, C_{lm}^1, C_{lm}^2 \succeq 0, \quad \forall (l,m) \in E,
\end{aligned}
\tag{3.10}
$$

where $B_k, A_{lm}, C_{lm}^1, C_{lm}^2$ are 2×2 positive semidefinite matrices. From Equation (2.6), $B_k, A_{lm},\ C_{lm}^1, C_{lm}^2 \succeq 0$ can be represented as second order cone constraints, so Problem (3.10) is a second order cone program.

Next we derive the dual of Problem $\mathcal{R}_2$ in [22]. In [22], the authors also use the notations in [71]. Also note that for ease of presentation, the authors omit some refinements (such as the shunt element) in their model.

Let $Z_k^{12} = Z_k^{21} = \sqrt{c_{k2}} \sum_{e\in E} \operatorname{tr}(Y_k^e W_e) + b_k$ be the non-diagonal element of matrix Z_k, and we present the model of Problem $\mathcal{R}_2$ in [22] derived from Optimization 3 in [71] as follows:

$$
\begin{aligned}
\min\ \ & \sum_{k\in N} \alpha_k \\
\text{s.t.}\ \ & P_k^- \leq \sum_{e\in E} \operatorname{tr}(Y_k^e W_e) + P_k^d \leq P_k^+,\ \forall k \in N, \\
& Q_k^- \leq \sum_{e\in E} \operatorname{tr}(\bar{Y}_k^e W_e) + Q_k^d \leq Q_k^+,\ \forall k \in N, \\
& (V_k^-)^2 \leq \sum_{e\in E} \operatorname{tr}(M_k^e W_e) \leq (V_k^+)^2,\ \forall k \in N, \\
& Z_k = \begin{bmatrix} \alpha_k - c_{k1} \sum_{e\in E} \operatorname{tr}(Y_k^e W_e) - a_k & Z_k^{12} \\ Z_k^{21} & 1 \end{bmatrix} \succeq 0,\ k \in N \\
& Z_{lm}^1 = \begin{bmatrix} S_{lm,\max}^2 & \sum_{e\in E} \operatorname{tr}(Y_k^e W_e) \\ \sum_{e\in E} \operatorname{tr}(Y_k^e W_e) & 1 \end{bmatrix} \succeq 0,\ \forall (l,m) \in E \\
& Z_{lm}^2 = \begin{bmatrix} S_{lm,\max}^2 & \sum_{e\in E} \operatorname{tr}(\bar{Y}_k^e W_e) \\ \sum_{e\in E} \operatorname{tr}(\bar{Y}_k^e W_e) & 1 \end{bmatrix} \succeq 0,\ \forall (l,m) \in E \\
& W_e \succeq 0,\ \forall e \in E,
\end{aligned}
\tag{3.11}
$$

where $\operatorname{tr}(Y_k W)$ in Optimization 3 [71] is replaced by $\sum_{e\in E} \operatorname{tr}(Y_k^e W_e)$ to take the branch structure into account and we relax constraint (5) in Optimization 3 [71] to two second-order cone constraints Z_{lm}^1, Z_{lm}^2. So basically, optimization problem (3.11) is a relaxation of Optimization 3 in [71] in terms of the variable W_e. The Lagrangian function of optimization problem (3.11) is

$$\begin{aligned}
\mathcal{L}(W,Z,\alpha,\lambda,\gamma,\mu,\eta,A,B,C) = & \sum_{k\in G}\alpha_k + \sum_{e\in E}\operatorname{tr}(W_e,A_e) + \sum_{k\in N}\operatorname{tr}(Z_k,B_k) \\
& + \sum_{(l,m)\in E}\left(\operatorname{tr}(Z^1_{lm},C^1_{lm}) + \operatorname{tr}(Z^2_{lm},C^2_{lm})\right) + \sum_{k\in N}\overline{\lambda}_k\left(P^+_k - P^d_k - \sum_{e\in E}\operatorname{tr}(Y^e_k W_e)\right) \\
& + \sum_{k\in N}\underline{\lambda}_k\left(-P^-_k + P^d_k + \sum_{e\in E}\operatorname{tr}(Y^e_k W_e)\right) + \sum_{k\in N}\overline{\gamma}_k\left(Q^+_k - Q^d_k - \sum_{e\in E}\operatorname{tr}(\bar{Y}^e_k W_e)\right) \\
& + \sum_{k\in N}\underline{\gamma}_k\left(-Q^-_k + Q^d_k + \sum_{e\in E}\operatorname{tr}(\bar{Y}^e_k W_e)\right) + \sum_{k\in N}\overline{\mu}_k\left((V^+_k)^2 - \sum_{e\in E}\operatorname{tr}(M^e_k W_e)\right) \\
& + \sum_{k\in N}\underline{\mu}_k\left((-V^-_k)^2 + \sum_{e\in E}\operatorname{tr}(M^e_k W_e)\right),
\end{aligned}$$

where $A_e, B_k, C^1_{lm}, C^2_{lm}$ are 2×2 positive semidefinite matrices and λ,γ,μ,η are nonnegative Lagrangian multipliers. Therefore the Lagrangian dual problem of optimization problem (3.11) is

$$\begin{aligned}
\max_{\lambda,\gamma,\mu,\eta,A,B,C}\quad & \min_{W,Z\succeq 0,\alpha}\mathcal{L}(W,Z,\alpha,\lambda,\gamma,\mu,\eta,A,B,C) \\
\text{s.t.}\quad & \lambda,\gamma,\mu,\eta \geq 0 \\
& A_e \succeq 0,\ \forall e\in E, \\
& B_k \succeq 0,\ \forall k\in N, \\
& C^1_{lm}, C^2_{lm} \succeq 0,\ \forall (l,m)\in E.
\end{aligned} \tag{3.12}$$

Once expanded, problem (3.12) has the same structure to the first level SOCP hierarchy for (PP-OPF) given in (3.9), with exactly the same conic constraints. □

Furthermore, for this pair of dual SOCP problems, strong duality holds.

Theorem 8. *Strong duality holds between the first level of the SOCP hierarchy obtained by using* (QM-SDSOS$_r$) *on* (PP-OPF) *and the SOCP relaxation for the ACOPF problem considered in [22].*

Proof. The basic idea is to try to find a strictly feasible solution for the dual problem. In order to prove strong conic duality, we need to find a strict feasible solution for either the primal or dual problem. For the dual SOCP problem (3.9), consider the following point,

$$\begin{aligned}
&\eta > 0, \\
&\alpha_k = |E|_{max} + 1 - |E_k| > 0, \beta_k = c_k^2 - \epsilon > 0, \quad \forall k \in N, \\
&\theta_{lm} = \delta_{lm} = 1, \quad \forall (l,m) \in E, \\
&\underline{\lambda}_k = c_k^1 + \Lambda > 0, \overline{\lambda}_k = \Lambda, \underline{\gamma}_k = \overline{\gamma}_k = 1, \quad \forall k \in N, \\
&\overline{\mu}_k = 1 + \frac{1 + |E|_{max}}{|N|} > 0, \underline{\mu}_k = 1, \quad \forall k \in N, \\
&B_k = \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon \end{bmatrix} \succ 0, \quad \forall k \in N, A_{lm} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \succ 0, \\
&C_{lm}^1 = C_{lm}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succ 0, \quad \forall (l,m) \in E.
\end{aligned} \tag{3.13}$$

In (3.13), $|E_k|$ is the number of edges whose endpoints contain node k, $|E|_{max} = \max\{|E_k|, k \in N\}$, for a fixed graph, $|E_k|, |E|_{max}$ are parameters; ϵ is a sufficient small positive number such that $c_k^2 - \epsilon > 0, \forall k \in N$, as c_k^2 is a positive parameter; Λ is a sufficient large positive number such that $c_k^1 + \Lambda > 0$. We substitute (3.13) into the constraints in Problem (3.9) and it is easy to verify that it satisfies all the constraints. Thus it is straightforward to see that the point is in the interior of the feasible set, which means point (3.13) is a strict feasible point for problem (3.9). Therefore, strong duality holds for Theorem 8. □

3.2.4 Numerical Results on ACOPF Instances

In this section we apply the LP and the SOCP hierarchies to solve the polynomial programming formulation of the ACOPF problem (PP-OPF), and then we compare the results to the ones that are obtained using the SDP hierarchy based on Lasserre's approach [66]. The SDP, SOCP, and LP based hierarchies are implemented in `Matlab`, and `SeDuMi` is used to solve the SDP programs while `MOSEK` is used to solve the SOCP and LP programs. The ACOPF instances are taken from [25, 111] and the results are summarized in Table 3.2.1. In computing the gaps, the best known bounds for the ACOPF instances are taken from [44]. Table 3.2.2 presents computational results of the structured SOCP approach where the structure of the ACOPF problem is exploited and a computational comparison with optimization problem R_2 of [22]. In Tables 3.2.1, and 3.2.2, - indicates that the problem was not solved within one hour of computational time and $\star$ implies that `Matlab` ran out of memory while generating the hierarchy. From Tables 3.2.1-3.2.2, the SDP hierarchy (3rd, 4th, and 5th columns in Table 3.2.1) provides solutions for instances up to 14 buses within 1 hour using the first level of the hierarchy. By using the SOCP hierarchy (6th, 7th, and 8th columns in Table 3.2.1), we obtained lower bounds for instances of up to 300 buses using the first level of the hierarchy with an average gap of 2% from the best known bounds.

	Best	SDP hierarchy			SOCP hierarchy			LP hierarchy		
Instance	Bound	Bound	Time (s.)	Gap	Bound	Time (s.)	Gap	Bound	Time (s.)	Gap
case9Q	5297.4	5296.71	49.59	0.01%	5230.49	0.09	1.25%	4448.00	0.05	16.02%
case14	8081.7	8081.52	1.59	0.002%	7677.48	0.17	4.99%	6859.31	0.07	15.12%
case30	574.5	-	-	-	567.96	0.48	1.55%	373.93	0.36	35.18%
case39	41889.1	-	-	-	41276.26	0.95	1.40%	3495.12	0.40	91.65%
case57	41712.0	-	-	-	41144.28	1.69	1.42%	38822.46	0.28	6.98%
case118	129372.4	-	-	-	126001.18	8.17	2.82%	96774.44	0.49	25.36%
case300	720031.0	-	-	-	706616.52	81.69	1.82%	$\star$	$\star$	$\star$

Table 3.2.1: Computational time comparison of first level SDP, SOCP and LP approximation.

	R_2		SOCP hierarchy		Structured SOCP		Structured SOCP Level 2	
instance	Bound	Time(s)	Bound	Time(s)	Bound	Time(s)	Bound	Time(s)
case9Q	5220.01	0.14	5230.49	0.09	5220.01	0.01	5298.89	27.15
case14	7661.05	0.15	7677.48	0.17	7659.96	0.08	7953.84	48.75
case30	567.79	0.43	567.96	0.48	567.79	0.31	$\star$	$\star$
case39	41283.67	0.55	41276.26	0.95	41278.11	0.51	$\star$	$\star$
case57	41157.72	0.38	41144.28	1.69	41157.72	0.26	$\star$	$\star$
case118	126072.18	0.95	126001.18	8.17	126050.17	0.71	$\star$	$\star$
case300	706779.64	2.62	706616.52	81.69	706779.64	1.57	$\star$	$\star$

Table 3.2.2: Computational time comparison of R_2 with SOCP and structured SOCP approximation.

Furthermore, the computational times for the SOCP approach are substantially reduced when the structure of the ACOPF problem is exploited as described in Table 3.2.2, while obtaining a comparable lower bound to R_2. Note that there are slight numerical differences in the bounds due to numerical errors in the solvers.

3.3 Concluding Remarks

In this chapter, we applied the recently proposed LP and SOCP approximation hierarchies to the ACOPF problem. Numerical results on ACOPF instances from the literature show that the use of these hierarchies, together with the sparsity structure of the problem, allow the computation of global bounds for large-scale ACOPF problems where the SDP hierarchy fails to provide such bounds. Moreover, the first level of the SOCP hierarchy is shown to be equivalent to the dual of the SOCP relaxations proposed in [22] for these problems.

The fact that the hierarchies considered here are based on using LP and SOCP allows for the future use the column generation approaches recently proposed in [1] to address the solution of larger-scale ACOPF problems.

Chapter 4

Pricing in Non-Convex Markets with Quadratic Costs

The presence of non-convexities is inherent to markets with economies of scale, start-up and/or shut-down costs, avoidable costs, indivisibilities, and minimum supply requirements [e.g., 21]. These non-convexities make the problem of finding appropriate prices that result in a market equilibrium challenging. This issue has been addressed in classical work by Gomory and Baumol [48], Scarf [99, 100], Starr [107]. Continued work in this area has been recently reviewed by Liberopoulos and Andrianesis [74] and Van Vyve et al. [112]. Here, we focus on considering the effect of non-convexities, and also potential convex quadratic costs that affect market prices. As discussed below, these market characteristics are particularly important in deregulated electricity markets with high penetration of renewable energy sources. We exemplify on the rest of the article using the electricity market. However note that this features are common to other network market economies.

Consider a deregulated electricity market [see, e.g., 88, 97], where a market maker or independent system operator (ISO) receives offers with information about generation constraints, marginal generation costs, and fixed commitment costs, from generators participating in the clearing process. Based on this information, the ISO decides the generators that should be available (committed) in the market, as well as their appropriate compensation.

Due to the potential existence of non-convexities, related to start-up/shut-down costs and minimum output requirements (amongst others), it is difficult for market operators, e.g., ISOs, to obtain the appropriate compensation values for committed generators. Furthermore, there can be long-term substantial differences when using different market pricing techniques [50]. One way in which ISOs can obtain an estimate of the correct compensation values is by finding the shadow (dual) prices associated to the *linear programming relaxation* [cf., 31] of the *mixed-integer linear program* (MILP) [cf., 31] associated to the market's *unit commitment problem* (UC) [cf., 51]. The generators' compensation values are then computed by adding *uplift payments* [cf., 55]. However, these uplift payments may be significant enough to modify the suppliers' incentives [74]. To address this issue, a number of alternative pricing schemes have been recently developed by Bjørndal and Jörnsten [16], Garcı et al. [41], Hogan and Ring [52], Liberopoulos and Andrianesis [74], O'Neill et al. [88], Ruiz et al. [97]; Araoz and Jörnsten [7]; among others [see, 74, for a recent review]. However, to the best of our knowledge, none of these approaches directly consider potential ramping costs in the electricity market.

The penetration of renewable energy sources such as wind and solar into the electricity market has been steadily increasing. The U.S. Energy Information Administration estimates that in the U.S., the percentage of energy generated from renewable energy sources has increased from 9.5% in 2006 to 13.3% in 2015, reaching close to 47% in the ERCOT system (Texas) and surpassing hydroelectric generation [40]. As a result, ISO's commitment decisions require conventional generators to ramp up or down regularly due to the variability and uncertainty in the power that renewable sources can provide throughout a given day [see, e.g., 104]. At the same time, ramping results in substantial wear and tear costs [see, e.g., 73]. Therefore, considering these ramping costs in the computation of compensation values is becoming necessary to internalize externalities nowadays absorbed by conventional (dispatchable) units, and provide adequate settlement prices for the all participants in the electricity market.

In order to take into account ramping costs in the electricity market, we extend the seminal results introduced by O'Neill et al. [88] to obtain appropriate prices in markets with non-convexities. O'Neill et al. [88] show that after solving the MILP associated with the ISO's UC, the desired clearing prices can be obtained from the shadow prices of the demand and capacity constraints of the linear program resulting from fixing the commitment (binary) variables of the ISO's UC to their optimal value.

Here, we extend this approach by considering ramping costs (or more generally, quadratic convex costs) into the formulation of the unit commitment problem in [63]. Then, using *convex optimization* techniques [cf., 23], we obtain the market-clearing prices in the presence of ramping costs from the *dual variables* values (shadow prices) associated with the optimal solution of an appropriate convex optimization problem. That is, we obtain a set of market clearing prices that satisfies incentive compatibility, i.e., the Walrasian market equilibrium conditions assure that suppliers would not want to change their energy dispatch at these prices. With these results in hand, we perform numerical experiments to show the impact of ramping costs on the clearing prices of a market with non-convexities.

The pricing results are mainly motivated by an ISO seeking to obtain clearing prices in an electricity market in which ramping becomes more prevalent, e.g., due to the presence of renewable energy sources. However, following O'Neill et al. [88], we use the general formulation of a market in which the auctioneer is buying and/or selling a commodity, and has an objective of maximizing the value to participants, when potential quadratic commodity costs (e.g., in a labor market [13]) or quadratic transaction costs (e.g., in a finance market [24, 87] are present in the market).

It is worth mentioning that related advances on obtaining market-clearing prices in markets with non-convexities have been done recently. For example, consider the work of Zoltowska [120], who considers demand shifting bids and transmission constraints in the market; Ye et al. [118], who consider non-convexities arising when considering flexible demand options from the point of view of the customer (e.g., the option to forgo demand); Sioshansi and Nicholson [103], who consider market non-convexities in both centrally- and self-committed and markets; Van Vyve et al. [112], who propose a new model to obtain prices in a market with non-convexities by combining the approaches used in both the US and European electricity markets; and Muatore [84], who proposes an algorithm to find market-clearing prices in markets with non-convexities that provide incentives to both maintain and increase generating capacity.

The rest of the chapter is organized as follows. In Section 4.2, we introduce the market's assignment problem in the presence of convex quadratic costs. Furthermore, we generalize the results of O'Neill et al. [88] to obtain clearing prices for the market. In Section 4.3, we illustrate our results by computing and analyzing the clearing prices associated to Scarf's classical market problem [100] when potential ramping costs are taken into account. In Section 4.4, we offer concluding remarks.

4.1 Introduction

In today's advanced economies, it has been widely believed that in the presence of nonconvexities (binary decision variables) in the cost function, it is not possible to guarantee the existence of prices that will allow the market to clear, unless the solution to the relaxed convex problem just happens to produce an integral solution (e.g. assignment problems). The economics and management science literature has occasionally addressed the problem of finding dual price interpretations to integer programs and MIPs. Motivated by the electric power markets in which nonconvexities arise from the operating characteristics of generators, O'Neill et al. in [88] discussed the existence of market clearing prices and the economic interpretation of strong duality for integer programs in the economic analysis of markets with nonconvexities (indivisibilities). They showed that the optimal solution to a linear program that solves the mixed integer program has dual variables that: (1) have the traditional economic interpretation as prices; (2) explicitly price integral activities; and (3) clear the market in the presence of nonconvexities.

Consider an auction market (like electric power markets) that can be represented by a

Primal Mixed Integer Program (PIP). The formulation below assumes that the auctioneer is buying and/or selling a set of goods, and has an objective of maximizing the value to bidders. The auctioneer is simply a computer code that finds a solution to the problem:

4.2 Market problem with convex quadratic costs

Consider a market, where the auctioneer is buying and/or selling a commodity, and has an objective of maximizing the value to participants, when potential transaction convex quadratic costs are present in the market. Following O'Neill et al. [88],[1] the market's assignment problem can be formulated as the following *mixed-integer quadratic program* (MIQP) [cf., 72]:

$$
\begin{aligned}
p^*_{\text{MIQP}} = \quad \min \quad & \sum_{k=1}^{n} c_k x_k + d_k z_k + r_k (x_k - x_k^0)^2 \\
\text{s.t.} \quad & \sum_{k=1}^{n} a_k x_k = b_0, \\
& g_k x_k + h_k z_k \geq b_k, && k = 1, \ldots, n, \\
& z_k \in \{0, 1\}, && k = 1, \ldots, n, \\
& x_k \geq 0, && k = 1, \ldots, n,
\end{aligned}
\tag{4.1}
$$

where for any market bidder $k = 1, \ldots, n$: $x_k \in \mathbb{R}_+$ represents the units of commodity provided by the bidder; $z_k \in \{0, 1\}$ indicates whether the bidder is committed or not to provide units of the commodity; $c_k, d_k \in \mathbb{R}$, are the variable and fixed costs (i.e., commodity and start-up) associated with the bidder's activities respectively; $a_k \in \mathbb{R}$reflects the production or demand characteristics of the bidder in the market-clearing constraint $\sum_{k=1}^{n} a_k x_k = b_0$, where $b_0 \in \mathbb{R}$ is the amount of commodity to be auctioned, with $b_0 \neq 0$ in a one-sided auction and $b_0 = 0$ in a two-sided auction; $g_k, h_k \in \mathbb{R}$ reflect restrictions on the bidder's operations (e.g., production of a particular plant is limited to the capacity of that plant); $b_k \in \mathbb{R}$ represents the right hand sides of the internal constraints of the

[1] We also follow the terminology used in terms of sellers as "bidders". Note however that in an auction, buyers submit *bids* to purchase a product and suppliers submit *offers* to sell a product.

bidder. Also, in an extension of the market assignment problem in O'Neill et al. [88, Sec. 5], $r_k \in \mathbb{R}_+$ denotes the quadratic costs associated with the deviation of the commodity provided by bidder $k = 1, \ldots, n$ from a target or previous commodity production level $x_k^0 \in \mathbb{R}$. The parameters $r_k \in \mathbb{R}_+, x_k^0 \in \mathbb{R}$ can also be used to model quadratic commodity costs, as well as quadratic transaction costs.

Notice that by assuming that $r_k \geq 0$ for $k = 1, \ldots, n$, one ensures that (4.1) has a convex quadratic objective. Also, note that if $r_k = 0$ for $k = 1, \ldots, n$, problem (4.1) is equivalent to the PIP problem in O'Neill et al. [88, Sec. 5] when bidders are assumed to provide a single commodity. This single-commodity assumption is made here for ease of presentation and to identify the commodity with power in electricity markets. However, all the results presented thereof extend in straightforward fashion to the more general multi-commodity market considered in O'Neill et al. [88].

In what follows, we assume that problem (4.1) is feasible; that is, there is an assignment of the bidders that satisfies both the operating constraints of the bidders as well as the market-clearing constraint. Also, without loss of generality, we assume that $a_k \neq 0$, $k = 1, \ldots, n$ (i.e., bidders that do not contribute to the market-clearing are not considered), and $r_k > 0$, $k = 1, \ldots, n$ (i.e., bidders that do not incur ramping costs do not need a corresponding quadratic constraint in formulation (4.2) below).

By introducing the auxiliary variables $y_k \in \mathbb{R}_+, k = 1, \ldots, n$, the market assignment problem (4.1) can be reformulated as:

$$\begin{array}{rlll} p^*_{\text{MIQP}} = & \min & \displaystyle\sum_{k=1}^n c_k x_k + d_k z_k + r_k y_k & \\ & \text{s.t.} & \displaystyle\sum_{k=1}^n a_k x_k = b_0, & \\ & & g_k x_k + h_k z_k \geq b_k, & k = 1, \ldots, n, \\ & & y_k \geq (x_k - x_k^0)^2, & k = 1, \ldots, n, \\ & & z_k \in \{0, 1\}, & k = 1, \ldots, n, \\ & & x_k, y_k \geq 0, & k = 1, \ldots, n. \end{array} \tag{4.2}$$

This follows from the fact that for any optimal solution of (4.2), the constraints $y_k \geq (x_k - x_k^0)^2$, $k = 1, \ldots, n$, are tight. Now let $\mathcal{K}^{n+1} \subseteq \mathbb{R}^{n+1}$ denote the *second-order* (or Lorentz) cone [cf., 4] in dimension $n+1$; that is,

$$\mathcal{K}^{n+1} = \{(w_0, w) \in \mathbb{R}^{n+1} : w_0 \geq \|(w_1, w_2, \ldots, w_n)\|_2\}, \tag{4.3}$$

where $\|\cdot\|_2$ represents the Euclidean norm. Note that like $\mathbb{R}^n$ or $\mathbb{R}^n_+$, the cone $\mathcal{K}^{n+1}$ is a *closed convex cone* [cf., 12]. Moreover, $\mathbb{R}^n$, $\mathbb{R}^n_+$, $\mathcal{K}^{n+1}$ are *self-dual* cones; that is, $(\mathbb{R}^n)^* = \mathbb{R}^n$, $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, $(\mathcal{K}^{n+1})^* = \mathcal{K}^{n+1}$, where for any convex set, the dual cone of $S \subseteq \mathbb{R}^n$ is equal to $S^* = \{u \in \mathbb{R}^n : u^\intercal w \geq 0 \text{ for all } w \in S\}$. Here, for any $u, w \in \mathbb{R}^n$, $u^\intercal w = \sum_{i=1}^n u_i w_i$ denotes the usual inner product in $\mathbb{R}^n$. The self-duality of $\mathbb{R}^n$, $\mathbb{R}^n_+$, and $\mathcal{K}^{n+1}$ is key in deriving an appropriate set of shadow prices, or the dual problem associated with the continuous relaxation (i.e., when $z_i \in \{0,1\}$ is replaced by $z_i \in [0,1]$ for $k = 1, \ldots, n$) of problem (4.2).

From the definition (4.3), it is not difficult to see [cf., 76, eq. (7)] that for any $k = 1, \ldots, n$,

$$y_k \geq (x_k - x_k^0)^2 \Leftrightarrow \big(y_k + 1, y_k - 1, 2(x_k - x_k^0)\big) \in \mathcal{K}^3.$$

Thus, problem (4.2) can be reformulated as the following *mixed-integer second-order cone program* (MISOCP) [109]:

$$\begin{aligned}
p^*_{\text{MISOCP}} = \quad \min \quad & \sum_{k=1}^n c_k x_k + d_k z_k + r_k y_k \\
\text{s.t.} \quad & \sum_{k=1}^n a_k x_k = b_0, \\
& g_k x_k + h_k z_k \geq b_k, & k = 1, \ldots, n, \\
& \big(y_k + 1, y_k - 1, 2(x_k - x_k^0)\big) \in \mathcal{K}^3, & k = 1, \ldots, n, \\
& z_k \in \{0,1\}, & k = 1, \ldots, n, \\
& x_k, y_k, z_k \geq 0, & k = 1, \ldots, n.
\end{aligned} \tag{4.4}$$

Similar to problem (4.4) above, a MISOCP is an optimization problem that besides a linear objective and linear constraints, has second-order cone constraints, as well as integer or binary variables. Problems with second-order cone constraints are widely used in applications in engineering and science. In particular, this type of constraints appear in structural design problems [76, Sec. 3.5], electrical engineering [15, 77], healthcare [79], and supply chain management [10]. In many instances, this is a result of the need to take into account the uncertainty of problem parameters and obtain solutions that are robust; that is, perform well in different scenarios [see, e.g., 79]. Moreover the solution of MISOCP problems can be obtained using commercial MISOCP solvers like `MOSEK`, `CPLEX`, and `Gurobi`.

Following O'Neill et al. [88], we next use the optimal solution of (4.4) to obtain the shadow (dual) prices associated with the market-clearing and bidder operational constraints in (4.4). Namely, let

$$z^* = \operatorname{argmin}_{z \in \{0,1\}^n} \{(4.4)\}; \tag{4.5}$$

that is, $z^* \in \{0,1\}^n$ is the vector of optimal values of the binary variables in (4.4). After replacing $z_k = z_k^*$, $k = 1, \ldots, n$ in (4.4), we obtain the following *second-order conic program* (SOCP) [cf., 76]:

$$\begin{array}{rlll}
p^*_{\text{SOCP}} = \min & \displaystyle\sum_{k=1}^n c_k x_k + d_k z_k + r_k y_k & & \\
\text{s.t.} & \displaystyle\sum_{k=1}^n a_k x_k = b_0, & (p_0) & \\
& g_k x_k + h_k z_k \geq b_k, & (q_k) & k = 1, \ldots, n, \\
& \big(y_k + 1, y_k - 1, 2(x_k - x_k^0)\big) \in \mathcal{K}^3, & (\gamma_k, \alpha_k, \beta_k) & k = 1, \ldots, n, \\
& z_k = z_k^*, & (p_k) & k = 1, \ldots, n, \\
& x_k, y_k \geq 0, & & k = 1, \ldots, n.
\end{array} \tag{4.6}$$

Similar to problem (4.6) above, a SOCP is an optimization problem that besides a linear objective and linear constraints, has second-order cone constraints. The key characteristic of SOCPs that are used here, is the fact that SOCPs are *convex optimization* problems. This follows from the fact that the second-order cone constraint $(w_0, w) \in \mathcal{K}^{n+1}$ is a convex constraint. As a result, there is a rich *duality* theory (which generalizes linear programming duality) for these problems, as well as polynomial-time solution algorithms [cf., 4]. In turn, these algorithms can be used together with *branch & bound* techniques [cf., 31] to solve MISOCP problems like (4.4).

Notice that in (4.6), we have associated the dual variables: p_0 with the market-clearing constraint, q_k with the k-th bidder's operation constraint, p_k with the k-th bidder's commitment constraint, and $(\gamma_k, \alpha_k, \beta_k)$ with the k-th second-order constraint, for $k = 1, \ldots, n$. Also, for ease of notation, let $u \in \mathbb{R}^n$ represent the vector of variables u_k, $k = 1, \ldots, n$. With this notation [cf., 4], the *dual* SOCP corresponding to the *primal* SOCP problem (4.6) can be obtained by constructing the *Lagrangean dual* of (4.6). Namely, let

$$L(x, y, z, p_0, p, q, \gamma, \alpha, \beta) = \sum_{k=1}^{n} L_k(x_k, y_k, z_k, p_0, p_k, q_k, \gamma_k, \alpha_k, \beta_k), \tag{4.7}$$

where

$$\begin{aligned} L_k(x_k, \cdots, \beta_k) &= c_k x_k + d_k z_k + r_k y_k + p_0(\tfrac{1}{k} b_0 - a_k x_k) - q_k(g_k x_k + h_k z_k - b_k) \\ &\quad + p_k(z_k^* - z_k) - (y_k + 1, y_k - 1, 2(x_k - x_k^0))^\intercal (\gamma_k, \alpha_k, \beta_k) \\ &= (c_k - a_k p_0 - g_k q_k - 2\beta_k) x_k + (d_k - h_k q_k - p_k) z_k \\ &\quad + (r_k - \gamma_k - \alpha_k) y_k + \tfrac{1}{k} b_0 p_0 + b_k q_k + z_k^* p_k - \gamma_k + \alpha_k + 2\beta_k x_k^0, \end{aligned} \tag{4.8}$$

for $k = 1, \ldots, n$, where $p_0 \in (\mathbb{R})^* = \mathbb{R}$, $q \in (\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, $p \in (\mathbb{R}^n)^* = \mathbb{R}^n$, and $(\gamma_k, \alpha_k, \beta_k) \in (\mathcal{K}^3)^* = \mathcal{K}^3$, for all $k = 1, \ldots, n$, are the dual variables or *lagrangian multipliers* associated to each of the constraints in problem (4.6).

From (4.7) and (4.8), it follows that the dual problem of (4.6):

$$\max_{\substack{p_0 \in \mathbb{R}, p \in \mathbb{R}^n, q \in \mathbb{R}^n_+, \\ (\gamma_k, \alpha_k, \beta_k) \in \mathcal{K}^3, k=1,\dots,n}} \quad \min_{x \geq 0, y \geq 0, z \geq 0} L(x, y, z, p, q, u, v, w, \gamma, \alpha, \beta),$$

is equivalent to

$$\begin{array}{rlll} d^*_{\text{SOCP}} = & \max & b_0 p_0 + \displaystyle\sum_{k=1}^{n} \left(b_k q_k + z^*_k p_k - \gamma_k + \alpha_k + 2\beta_k x^0_k \right) & \\ & \text{s.t.} & c_k - a_k p_0 - g_k q_k - 2\beta_k \geq 0, & k = 1, \dots, n, \\ & & d_k - h_k q_k - p_k \geq 0, & k = 1, \dots, n, \\ & & r_k - \gamma_k - \alpha_k \geq 0, & k = 1, \dots, n, \\ & & (\gamma_k, \alpha_k, \beta_k) \in \mathcal{K}^3, & k = 1, \dots, n, \\ & & q_k \geq 0, & k = 1, \dots, n. \end{array} \tag{4.9}$$

In Proposition 13 below, we show that strong duality holds between (4.6) and (4.9), and that their optimal objectives are attained. For this purpose, we introduce, for any set $S \subseteq \mathbb{R}^n$ the notion of its *interior*; that is, $\text{int}(S) = \{w \in S: \text{ for any } u \in \mathbb{R}^n, \text{ there exists } \epsilon > 0, \text{ such that } w + \epsilon u \in S\}$.

Proposition 13. *Assume that* (4.1) *is feasible, and* $a_k \neq 0,\ c_k \geq 0, r_k > 0$*, for all* $k = 1, \dots, n$*. Then*

$$p^*_{\text{MIQP}} = p^*_{\text{MISOCP}} = p^*_{\text{SOCP}} = d^*_{\text{SOCP}}.$$

Proof. From the discussion above, it is clear that problems (4.1), (4.4), and (4.6) are equivalent. Therefore, $p^*_{\text{MIQP}} = p^*_{\text{MISOCP}} = p^*_{\text{SOCP}}$. It then remains to show both p^*_{SOCP} and d^*_{SOCP} are attained and strong duality holds between (4.6) and (4.9); that is, $p^*_{\text{SOCP}} = d^*_{\text{SOCP}}$. For this purpose, we first show that both (4.6) and (4.9) are *strictly feasible* [cf., 4].

Notice that from the feasibility of (4.1), the fact that the feasible set of (4.1) is closed, and that the objective of (4.1) is bounded below by $\min\{0, k \min_{k=1,\dots,n}\{d_k\}\}$, it follows from Weierstrass' Theorem that (4.1) has an optimal solution. Let $x^* \in \mathbb{R}^n_+$, $z^* \in \{0,1\}^n$ be the optimal solution of (4.1), and consider any vector $y \in \mathbb{R}^n_+$ such that $y_k > (x^*_k - x^0_k)^2$, for $k = 1, \dots, n$. It is easy to see that $(x^*, y, z^*) \in \mathbb{R}^{2n}_+ \times \{0,1\}^n$ is feasible for (4.6). Furthermore, we have that

$$y_k + 1 > \|(y_k - 1, 2(x^*_k - x^0_k))\|_2,$$

for $k = 1, \dots, n$. Thus, $(x^*, y, z^*) \in \mathbb{R}^{2n}_+ \times \{0,1\}^n$ is a *strictly feasible* solution for (4.6); that is, $(x^*, y, z^*) \in \mathbb{R}^{2n}_+ \times \{0,1\}^n$ is feasible for (4.6), and $(y_k + 1, y_k - 1, 2(x^*_k - x^0_k)) \in \mathrm{int}(\mathcal{K}^3)$, $k = 1, \dots, n$. Thus, problem (4.6) is strictly feasible. Now consider the assignment

$$\begin{aligned} & p_0 = \min_{k=1,\dots,n} \left\{ \tfrac{c_k - g_k}{a_k} \right\}, \\ & (p_k, q_k, \gamma_k, \alpha_k, \beta_k) = \left(d_k - h_k - 1, 1, \tfrac{1}{2} r_k, \tfrac{1}{4} r_k, -\tfrac{1}{4} r_k \right), \end{aligned} \tag{4.10}$$

for $k = 1, \dots, n$. Clearly, (4.10) is feasible for (4.9), with

$$\gamma_k = \frac{1}{2} r_k > \frac{1}{2\sqrt{2}} r_k = \|(\alpha_k, \beta_k)\|_2,$$

for $k = 1, \dots, n$. That is, (4.10) is a feasible solution for (4.9), and $(\gamma_k, \alpha_k, \beta_k) \in \mathrm{int}(\mathcal{K}^3)$, $k = 1, \dots, n$. Thus, problem (4.9) is strictly feasible. The result then follows from SOCP duality [see, e.g., 4, Thm. 13]. □

Now let us consider the individual problems associated to the bidders in the market. For each bidder $k = 1, \dots, n$, let t_0 be the unit commodity price and t_k be the price reflecting the commitment action offered to individual k by the auctioneer. Define the following problem,

$$
\begin{aligned}
\min \quad & c_k x_k + d_k z_k + r_k(x_k - x_k^0)^2 - t_0(a_k x_k) - t_k z_k \\
\text{s.t.} \quad & g_k x_k + h_k z_k \geq b_k, \\
& x_k \geq 0, \\
& z_k \in \{0,1\};
\end{aligned}
$$

which similar to problem (4.1) is equivalent to the following MISOCP:

$$
\begin{aligned}
p^*_{\text{MISOCP}_k}(t_0, t_k) = \min \quad & c_k x_k + d_k z_k + r_k y_k - t_0(a_k x_k) - t_k z_k \\
\text{s.t.} \quad & g_k x_k + h_k z_k \geq b_k, \\
& (y_k + 1, y_k - 1, 2(x_k - x_k^0)) \in \mathcal{K}^3, \\
& x_k, y_k \geq 0, \\
& z_k \in \{0,1\}.
\end{aligned}
\tag{4.11}
$$

Following O'Neill et al. [88], below we define both the market-clearing prices and associated market-clearing contracts between the auctioneer and the bidders.

Definition 9. *A competitive equilibrium for the market is a set of prices $\{t_0^*, t_k^*\}$ and allocations $\{x_k^*, z_k^*\}$, such that*

(a) At the prices $\{t_0^*, t_k^*\}$, the allocations $\{x_k^*, z_k^*\}$ solve (4.11) *for all $k = 1, \dots, n$;*

(b) The market clears: $\sum_{k=1}^n a_k x_k^* = b_0$.

Definition 10. *Let T_k be a contract between the auctioneer and bidder $k \in \{1, \dots, n\}$ with the following terms:*

(a) Bidder k operates following $z_k = z_k^*, x_k = x_k^*$.

(b) Bidder k receives an amount from the auctioneer that is equal to the following payment: $p_0^* a_k x_k + p_k^* z_k$.

In what follows, we refer to $T = \{T_k \text{ for } k = 1, \ldots, n\}$, to describe the market-clearing contracts between the auctioneer and the bidders. Below, we provide the main result of the article; namely, a characterization of the market-clearing prices for a market with non-convexities and convex quadratic costs.

Theorem 9. *Assume that* (4.1) *is feasible, and* $a_k \neq 0$, $c_k \geq 0, r_k > 0$, *for all* $k = 1, \ldots, n$. *Let* $\{x_k^*, y_k^*, z_k^*\}$ *for all* $k = 1, \ldots, n$ *be an optimal solution to* (4.1). *Also, let* p_0^*, *and* $\{p_k^*, q_k^*, \gamma_k^*, \alpha_k^*, \beta_k^*\}$ *for all* $k = 1, \ldots, n$ *be an optimal solution to* (4.9). *If in* (4.11) *we define* $t_0 = p_0^*$ *and* $t_k = p_k^*$ *for all* $k = 1, \ldots, n$, *then the prices* $\{p_0^*, p_k^*\}$ *and allocations* $\{x_k^*, z_k^*\}$ *for all* $k = 1, \ldots, n$ *represent a competitive equilibrium.*

Proof. Note that $\{x_k^*, y_k^*, z_k^*\}$ for all $k = 1, \ldots, n$ is also an optimal solution of (4.6). From the Karush-Kuhn-Tucker (KKT) conditions associated to the optimal solutions of both (4.6) and (4.9), it follows that:

$$0 \leq (c_k - a_k p_0^* - g_k q_k^* - 2\beta_k^*) \perp x_k^* \geq 0, \qquad k = 1, \ldots, n, \tag{4.12}$$

$$0 \leq (d_k - h_k q_k^* - p_k^*) \perp z_k^* \geq 0, \qquad k = 1, \ldots, n,$$

$$0 \leq (r_k - \gamma_k^* - \alpha_k^*) \perp y_k^* \geq 0, \qquad k = 1, \ldots, n, \tag{4.13}$$

$$0 = p_0^*(b_0 - \sum_{k=1}^{n} a_k x_k^*),$$

$$0 \leq q_k^* \perp (g_k x_k^* + h_k z_k^* - b_k) \geq 0, \qquad k = 1, \ldots, n, \tag{4.14}$$

$$0 = p_k^*(z_k' - z_k^*), \qquad k = 1, \ldots, n,$$

$$(y_k^* + 1, y_k^* - 1, 2(x_k^* - x_k^0)) \perp (\gamma_k^*, \alpha_k^*, \beta_k^*), \qquad k = 1, \ldots, n, \tag{4.15}$$

where the notation $u \perp w$, for $u, w \in \mathbb{R}$ denotes the complementary slackness between u and w. Now consider the following problem under the contract T; that is, each individual bidder k is offered prices $\{p_0^*, p_k^*\}$, then each participant $k = 1, \ldots, n$ solves (4.11) with

$t_0 = p_0^*$, $t_k = p_k^*$, $k = 1, \ldots, n$ to minimize their operation cost. That is, each bidder solves

$$\begin{aligned} p^*_{\mathrm{MISOCP}_k}(p_0^*, p_k^*) = \min \quad & c_k x_k + d_k z_k + r_k y_k - p_0^*(a_k x_k) - p_k^* z_k \\ \text{s.t.} \quad & g_k x_k + h_k z_k \geq b_k, \\ & (y_k + 1, y_k - 1, 2(x_k - x_k^0)) \in \mathcal{K}^3, \\ & x_k, y_k \geq 0, \\ & z_k \in \{0, 1\}. \end{aligned} \tag{4.16}$$

Clearly, $(x_k, y_k, z_k) = (x_k^*, y_k^*, z_k^*)$ is feasible for (4.16), and the objective value of this solution, denoted $\hat{p}_{\mathrm{MISOCP}_k}(p_0^*, p_k^*)$, is

$$\hat{p}_{\mathrm{MISOCP}_k}(p_0^*, p_k^*) = c_k x_k^* + d_k z_k^* + r_k y_k^* - p_0^* a_k x_k^* - p_k^* z_k^*.$$

Using the complementarity equations (4.13), (4.14), and (4.15), it follows that

$$\begin{aligned} \hat{p}_{\mathrm{MISOCP}_k}(p_0^*, p_k^*) = & c_k x_k^* + d_k z_k^* + r_k y_k^* - p_0^* a_k x_k^* - p_k^* z_k^* - q_k^*(g_k x_k^* + h_k z_k^* - b_k) \\ & - (r_k - \gamma_k^* - \alpha_k^*) y_k^* - (y_k^* + 1, y_k^* - 1, 2(x_k^* - x_k^0))^\intercal (\gamma_k^*, \alpha_k^*, \beta_k^*), \\ = & (c_k - a_k p_0^* - g_k q_k^* - 2\beta_k^*) x_k^* + (d_k - p_k^* - h_k q_k^*) z_k^* + y_k^*(\gamma_k^* + \alpha_k^*) \\ & - (y_k^* \gamma_k^* + \gamma_k^* + y_k^* \alpha_k^* - \alpha_k^* - 2\beta_k^* x_k^0), \\ = & q_k^* b_k - \gamma_k^* + \alpha_k^* + 2\beta_k^* x_k^0. \end{aligned}$$

Next we show that $(x_k, y_k, z_k) = (x_k^*, y_k^*, z_k^*)$ is the optimal solution for problem (4.16). Let $(x_k, y_k, z_k) \in \mathbb{R}_+^2 \times \{0, 1\}$ be a feasible solution of (4.16). It follows that $g_k x_k + h_k z_k - b_k \geq 0$, $y_k \geq 0$ and $(y_k + 1, y_k - 1, 2(x_k - x_k^0)) \in \mathcal{K}^3$. Therefore

$$\begin{aligned} & q_k^*(g_k x_k + h_k z_k - b_k) \geq 0 \\ & y_k(r_k - \gamma_k^* - \alpha_k^*) \geq 0, \\ & (y_k + 1, y_k - 1, 2(x_k - x_k^0))^\intercal (\gamma_k^*, \alpha_k^*, \beta_k^*) \geq 0, \end{aligned} \tag{4.17}$$

since the feasibility of p_0^*, and $\{p_k^*, q_k^*, \gamma_k^*, \alpha_k^*, \beta_k^*\}$ for all $k = 1, \dots, n$, for (4.9) ensures that $q_k^* \geq 0$, $r_k - \gamma_k^* - \alpha_k^* \geq 0$, and $(\gamma_k^*, \alpha_k^*, \beta_k^*) \in \mathcal{K}^3 = (\mathcal{K}^3)^*$. Also, from the fact that p_0^*, and $\{p_k^*, q_k^*, \gamma_k^*, \alpha_k^*, \beta_k^*\}$ for all $k = 1, \dots, n$ is feasible for (4.9), $x_k, z_k \in \mathbb{R}_+$, it follows that:

$$\begin{aligned} &(c_k - a_k p_0^* - g_k q_k^* - 2\beta_k^*) x_k \geq 0, \\ &(d_k - h_k q_k^* - p_k^*) z_k \geq 0. \end{aligned} \tag{4.18}$$

Now let p_{MISOCP_k} denote the objective value of (4.16) associated with the feasible solution $(x_k, y_k, z_k) \in \mathbb{R}_+^2 \times \{0, 1\}$; that is

$$p_{\mathrm{MISOCP}_k} = c_k x_k + d_k z_k + r_k y_k - p_0^* a_k x_k - p_k^* z_k.$$

Using (4.17) and then (4.18), we have

$$\begin{aligned} p_{\mathrm{MISOCP}_k} &\geq c_k x_k + d_k z_k + r_k y_k - p_0^* a_k x_k - p_k^* z_k - q_k^*(g_k x_k + h_k z_k - b_k) \\ &\quad -y_k(r_k - \gamma_k^* - \alpha_k^*) - (y_k \gamma_k^* + \gamma_k^* + y_k \alpha_k^* - \alpha_k^* + 2\beta_k^* x_k - 2\beta_k^* x_k^0) \\ &= (c_k - a_k p_0^* - g_k q_k^* - 2\beta_k^*) x_k + (d_k - h_k q_k^* - p_k^*) z_k + q_k^* b_k - \gamma_k^* + \alpha_k^* + 2\beta_k^* x_k^0 \\ &\geq q_k^* b_k - \gamma_k^* + \alpha_k^* + 2\beta_k^* x_k^0 \\ &= \hat{p}_{\mathrm{MISOCP}_k}(p_0^*, p_k^*). \end{aligned}$$

This shows that (x_k^*, y_k^*, z_k^*) is the optimal solution for (4.16). Furthermore, the solution (x_k^*, y_k^*, z_k^*) satisfies the market-clearing condition $\sum_{k=1}^n a_k x_k^* = b_0$, therefore, (x_k^*, y_k^*, z_k^*) provides a market-clearing allocation. □

4.3 Scarf's market instance

As an example of a market with non-convexities, consider a problem proposed by Scarf [100]. The objective is to minimize the total cost subject to satisfying the demand in an electricity market. Two types of plants are available to provide the electricity in the market. The characteristics of each type of plant, including costs and operational constraints, are summarized in Table 4.3.1.

Characteristics	(type 1 plant) Smokestack	(type 2 plant) High Tech
Capacity	16.00	7.00
Construction cost	53.00	30.00
Marginal cost	3.00	2.00
Average cost at capacity	6.31	6.28
Total cost at capacity	101.00	44.00

Table 4.3.1: Characteristics of Smokestack and High Tech plants [88].

Scarf's market problem can be formulated as the following mixed-integer linear program

$$\begin{aligned}
\min \quad & \sum_{i=1}^{5}(3x_{1i} + 53z_{1i}) + \sum_{j=1}^{10}(2x_{2j} + 30z_{2j}) & & \\
\text{s.t.} \quad & \sum_{i=1}^{5} x_{1i} + \sum_{j=1}^{10} x_{2j} = D, & & \\
& x_{1i} - 16z_{1i} \leq 0, & & i = 1, \dots, 5, \\
& x_{2j} - 7z_{2j} \leq 0, & & j = 1, \dots, 10, \\
& x_{1i}, x_{2j} \geq 0, & & i = 1, \dots, 5, j = 1, \dots, 10, \\
& z_{1i}, z_{2j} \in \{0, 1\}, & & i = 1, \dots, 5, j = 1, \dots, 10,
\end{aligned} \tag{4.19}$$

where D is the total demand. The market-clearing price for this problem is studied in [88]. Table 4.3.2 summarizes the optimal solution of (4.19) for different values of the demand. It is clear that as the demand increases, the number of plants of different types used can change dramatically. For example, when the demand is 56, all type 1 plants are closed and eight (8) type 2 plants are open; however, when the demand is 60, two (2) type 1 plants are open whereas now only four (4) type 2 plants are open. The market-clearing prices in Table 4.3.3 are obtained from the dual (shadow) prices of the linear program obtained from (4.19) by fixing its binary variables to their optimal value [cf., 88]. Note that for all the instances with different demand, the market-clearing prices remain the same.

	Units of type		Unit's output		
Demand	1	2	1	2	Total Cost
56	0	8	0	56	352
58	1	6	16	42	365
60	2	4	32	28	378
62	3	2	48	14	391
64	4	0	64	0	404
66	2	5	31	35	419
68	3	3	47	21	432
70	0	10	0	70	440

Table 4.3.2: Optimal solution of Scarf's market problem (4.19) [88].

Commodity Price	Plant 1 Start-up Price	Plant 2 Start-up Price
3	53	23

Table 4.3.3: Market-clearing price of Scarf's problem.

Now we consider a modified Scarf problem with quadratic ramping costs,

$$\begin{aligned}
\min \quad & \sum_{i=1}^{5}(3x_{1i} + 53z_{1i} + r_1(x_{1i} - x_{1i}^0)^2) \\
& + \sum_{j=1}^{10}(2x_{2j} + 30z_{2j} + r_2(x_{2j} - x_{2j}^0)^2) \\
\text{s.t.} \quad & \sum_{i=1}^{5} x_{1i} + \sum_{j=1}^{10} x_{2j} = D, \\
& x_{1i} - 16z_{1i} \leq 0, && i = 1, \ldots, 5, \\
& x_{2j} - 7z_{2j} \leq 0, && j = 1, \ldots, 10, \\
& x_{1i}, x_{2j} \geq 0, && i = 1, \ldots, 5, j = 1, \ldots, 10, \\
& z_{1i}, z_{2j} \in \{0, 1\}, && i = 1, \ldots, 5, j = 1, \ldots, 10,
\end{aligned} \tag{4.20}$$

where x_{1i}^0, x_{2j}^0, for all $i = 1, \ldots, 5$, $j = 1, \ldots, 10$, are set to be the optimal unit's generation outputs, obtained after solving (4.19) with a demand $D = 55$. The quadratic terms in the objective function can be interpreted as the ramping costs of moving from the original output plan set from (4.19) with a demand $D = 55$. It is clear that by setting $D = 55$ in (4.20), regardless of the values assigned to r_1, r_2, problem (4.20) has the same optimal solution.

After setting $r_1 = r_2 = 0.1$, we can see from Table 4.3.4 that in most cases, the number of type 1 plants and type 2 plants with full capacity remain the same as a result of the ramping costs (in contrast with Table 4.3.2). Note that as the demand increases, a type 2 plant with partial capacity is opened for production for $D \geq 64$. Table 4.3.5 summarizes the market-clearing prices obtained using the results in Section 4.2. In order to maintain a competitive equilibrium, the market-clearing prices vary from case to case. In contrast with the start-up prices obtained by O'Neill et al. [88], note that from the results of Table 4.3.5, it follows that the start-up price can differ for type 2 plants producing at full capacity, and type 2 plants producing at partial capacity. Specifically, the start-up price of type 2 plants producing at full capacity changes with the demand, whereas the start-up price of type-2 plants producing at partial capacity remains constant at 30 for demands between 56 to 68. The start-up price for closed type 2 plants is the same as the price for type 2 plants with full capacity. The unit commodity price varies in a small range but it does not remain the same. The start-up price for all type 1 plants remains mostly equal to 53, but this price is different for some demand levels (e.g., compare $D = 60$ and $D = 62$).

	Type 1 Output				Type 2 Output					
	Partial		Full		Partial		Full		Cost	
Demand	No.	Prod.	No.	Prod.	No.	Prod.	No.	Prod.	Ramp	Total
56	3	45.00	0	0	1	4.00	1	7	1.90	377.90
58	3	46.50	0	0	1	4.50	1	7	2.10	383.60
60	0	0	3	48	1	5.00	1	7	2.50	389.50
62	0	0	3	48	1	7.00	1	7	4.90	395.90
64	3	47.40	0	0	2	9.60	1	7	4.62	429.02
66	0	0	3	48	2	11.00	1	7	6.05	435.05
68	0	0	3	48	2	13.00	1	7	8.45	441.45
70	1	15.00	3	48	0	0	1	7	22.50	467.50

Table 4.3.4: Optimal solution of modified Scarf's problem with $r_1 = 0.1, r_2 = 0.1$.

Tables 4.3.6 and Table 4.3.7 show that after setting $r_1 = 0.1, r_2 = 0.3$ in (4.20), similar conclusions as in the case $r_1 = r_2 = 0.1$ can be reached. However, with a higher ramping cost on type 2 plants, we can see that with demand values between 64 to 68, instead of operating one more type 2 plant to satisfy the demand, the optimal solution suggests to operate an additional type 1 plant instead.

		Plant 1 Start-up Price		Plant 2 Start-up Price	
Demand	Unit Price	Partial	Full & Closed	Partial	Full & Closed
56	2.80	53.00	53.00	30	24.40
58	2.90	53.00	53.00	30	23.70
60	3.00	53.00	53.00	30	23.00
62	3.40	46.60	46.60	30	20.20
64	2.96	53.00	53.00	30	23.28
66	3.10	51.40	51.40	30	22.30
68	3.30	48.20	48.20	30	20.90
70	6.00	53.00	5.00	2	2.00

Table 4.3.5: Market-clearing price of modified Scarf's problem with $r_1 = 0.1, r_2 = 0.1$.

	Type 1 Output				Type 2 Output					
	Partial		Full		Partial		Full		Cost	
Demand	No.	Prod.	No.	Prod.	No.	Prod.	No.	Prod.	Ramping	Total
56	3	47.4	0	0	1	1.6	1	7	0.78	379.18
58	0	0	3	48	1	3.0	1	7	2.70	385.70
60	0	0	3	48	1	5.0	1	7	7.50	394.50
62	0	0	3	48	0	0	2	14	14.70	405.70
64	1	9.0	3	48	0	0	1	7	4.62	435.10
66	1	11.0	3	48	0	0	1	7	12.10	445.10
68	1	13.0	3	48	0	0	1	7	16.90	455.90
70	1	15.0	3	48	0	0	1	7	22.50	467.50

Table 4.3.6: Optimal solution of modified Scarf's problem with $r_1 = 0.1, r_2 = 0.3$.

		Start-up Compensation			
		Type 1 Output		Type 2 Output	
Demand	Unit Price	Partial	Full & Closed	Partial	Full & Closed
56	2.96	53.00	53.00	30.00	23.28
58	3.80	40.20	40.20	30.00	17.40
60	5.00	21.00	21.00	30.00	9.00
62	6.20	1.80	1.80	30.00	0.60
64	4.80	53.00	24.20	10.40	10.40
66	5.20	53.00	17.80	7.60	7.60
68	5.60	53.00	11.40	4.80	4.80
70	6.00	53.00	5.00	2.00	2.00

Table 4.3.7: Market-clearing price of modified Scarf's problem with $r_1 = 0.1, r_2 = 0.3$.

Figures 4.3.1 and 4.3.2 compare the solutions obtained from the three cases discussed thus far; that is when $r_1 = r_2 = 0$, when $r_1 = r_2 = 0.1$, and when $r_1 = 0.1, r_2 = 0.3$.

Next we consider an extreme case where the parameter of ramping cost is relatively large. With $r_1 = r_2 = 1$, similar analysis can be applied from the analysis with $r_1 = r_2 = 0.1$. However, with a relative large ramping cost parameter, the unit commodity price

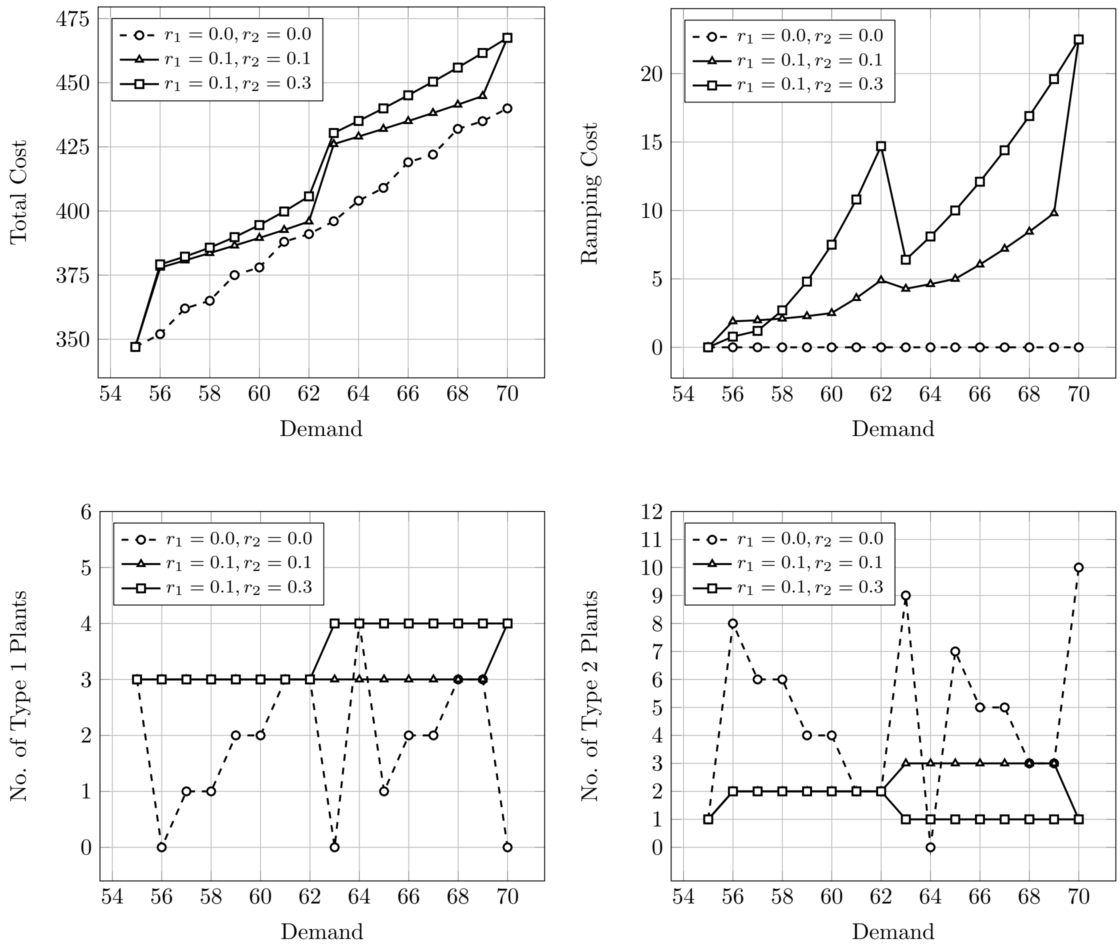

Figure 4.3.1: Comparison of optimal solution of (4.20) for different values of ramping costs r_1, r_2 ($r_1 = 0, r_2 = 0$ indicates no ramping costs).

Table 4.3.8: Optimal solution of modified Scarf's problem with $r_1 = r_2 = 1$.

	Type 1 Output				Type 2 Output					
	Partial		Full		Partial		Full		Cost	
Demand	No.	Prod.	No.	Prod.	No.	Prod.	No.	Prod.	Ramp	Total
56	0	0	3	48	1	1	1	7	1.0	380.0
58	0	0	3	48	1	3	1	7	9.0	392.0
60	0	0	3	48	1	5	1	7	25.0	412.0
62	0	0	3	48	1	7	1	7	49.0	440.0
64	0	0	3	48	2	9	1	7	40.5	465.5
66	0	0	3	48	2	11	1	7	60.5	489.5
68	0	0	3	48	2	13	1	7	84.5	517.5
70	0	0	3	48	3	15	1	7	75.0	542.0

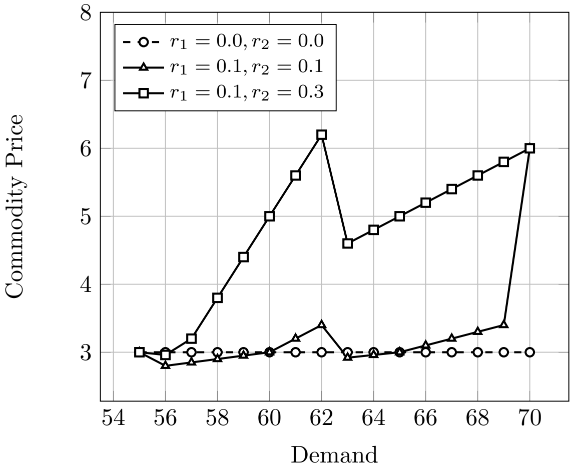

Figure 4.3.2: Unit commodity prices obtained from the solution of (4.20), for different values of ramping costs r_1, r_2 ($r_1 = 0, r_2 = 0$ indicates no ramping costs).

as well as start-up price can change dramatically. It is interesting to see that when the demand becomes 58, the start-up prices for type 1 plant, type 2 plant with full capacity and closed type 2 plant are negative, which means these plants need to pay instead of getting paid to open in exchange for a very high commodity price (compared with Table 4.3.3 and Table 4.3.5).

Now consider the case in which $r_1 = r_2 = 1$ and $D = 60$ in (4.20), as an example to illustrate that the dual prices obtained by using the methodology of Section 4.2, result in a competitive equilibrium. For any $j = 1, \ldots, n$, the individual problem for a type 2 plant j committed by the central operator to produce at partial capacity is

$$\begin{aligned} \min \quad & 2x_{2j} + 30z_{2j} + (x_{2j} - 0)^2 - 12x_{2j} - 30z_{2j} \\ \text{s.t.} \quad & x_{2j} - 7z_{2j} \leq 0, \\ & x_{2j} \geq 0, \\ & z_{2j} \in \{0, 1\}, \end{aligned}$$

and its optimal solution is $(x^*_{2j}, z^*_{2j}) = (5, 1)$, which matches the optimal solution in Table 4.3.8. It is not difficult to check, given the dual prices in Table 4.3.9, that the optimal solution for every individual problem matches its corresponding solution in Table 4.3.8, verifying Theorem 9.

		Plant 1 Start-up Price		Plant 2 Start-up Price	
Demand	Unit Price	Partial	Full & Closed	Partial	Full & Closed
56	4	37	37	30	16
58	8	-27	-27	30	-12
60	12	-91	-91	30	-40
62	16	-155	-155	30	-68
64	11	-75	-75	30	-33
66	13	-107	-107	30	-47
68	15	-139	-139	30	-61
70	12	-91	-91	30	-40

Table 4.3.9: Market-clearing price of modified Scarf's problem with $r_1 = r_2 = 1$.

4.4 Conclusion

We consider the problem of obtaining appropriate market-clearing prices when the market has both non-convexities and convex quadratic costs. Our results show that by using convex optimization techniques, the work of [88] on pricing in markets with non-convexities, can be extended to markets in which both non-convexities and convex quadratic costs arise. For the electricity market these two features arise due to generator fixed costs (or other operational constraints), and ramping or quadratic generation costs. Considering both of these characteristics have become increasingly important due the high penetration of renewable energy sources (RES). This is due to the output volatility of RES, which requires conventional generators to ramp up or down more frequently. Besides electricity markets, non-convexity features appear in other markets such as the financial and labor market [13, 24]. Thus our results have an impact in a wide range of potential markets.

Furthermore, we believe that the techniques outlined here can be used to extend other pricing methodologies for markets with non-convexities [74, see,] that aim to obtain prices with different characteristics. Finally, in the context of electricity markets, it is natural to consider how the consideration of ramping costs affect prices in an electrical network with congested lines. Addressing these questions provides intriguing directions for future research work in this area.

Bibliography

[1] Ahmadi, A. A., Dash, S., and Hall, G. (2017). Optimization over structured subsets of positive semidefinite matrices via column generation. *Discrete Optimization*, 24:129–151.

[2] Ahmadi, A. A. and Majumdar, A. (2014). DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, pages 1–5. IEEE.

[3] Ahmed, F., Dür, M., and Still, G. (2013). Copositive programming via semi-infinite optimization. *Journal of optimization theory and applications*, 159(2):322–340.

[4] Alizadeh, F. and Goldfarb, D. (2003). Second-order cone programming. *Mathematical programming*, 95(1):3–51.

[5] Anjos, M. and Lasserre, J. B., editors (2012). *Handbook on Semidefinite, Conic and Polynomial Optimization Handbook on Semidefinite, Conic and Polynomial Optimization*, volume 166 of *International Series in Operations Research & Management Science*. Springer.

[6] ApS, M. (2015). *The MOSEK optimization toolbox for MATLAB manual. Version 7.1.*

[7] Araoz, V. and Jörnsten, K. (2011). Semi-Lagrangean approach for price discovery in markets with non-convexities. *European Journal of Operational Research*, 214(2):411–417.

[8] Arima, N., Kim, S., and Kojima, M. (2013). A quadratically constrained quadratic

optimization model for completely positive cone programming. *SIAM Journal on Optimization*, 23(4):2320–2340.

[9] Arima, N., Kim, S., and Kojima, M. (2016). Extension of completely positive cone relaxation to moment cone relaxation for polynomial optimization. *Journal of Optimization Theory and Applications*, 168(3):884–900.

[10] Atamtürk, A., Berenguer, G., and Shen, Z.-J. (2012). A conic integer programming approach to stochastic joint location-inventory problems. *Operations Research*, 60(2):366–381.

[11] Bai, L., Mitchell, J. E., and Pang, J.-S. (2015). On conic QPCCs, conic QCQPs and completely positive programs. *Mathematical Programming*, pages 1–28.

[12] Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM.

[13] Bentolila, S. and Saint-Paul, G. (1992). The macroeconomic impact of flexible labor contracts, with an application to Spain. *European Economic Review*, 36(5):1013–1047.

[14] Berman, A. and Plemmons, R. J. (1979). Nonnegative matrices. *The Mathematical Sciences, Classics in Applied Mathematics*, 9.

[15] Bienstock, D., Chertkov, M., and Harnett, S. (2014). Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495.

[16] Bjørndal, M. and Jörnsten, K. (2008). Equilibrium prices supported by dual price functions in markets with non-convexities. *European Journal of Operational Research*, 190(3):768–789.

[17] Blekherman, G., Parrilo, P. A., and Thomas, R. R. (2012). *Semidefinite optimization and convex algebraic geometry*. SIAM.

[18] Bomze, I. M. (2013). Copositive relaxation beats Lagrangian dual bounds in quadratically and linearly constrained QPs. *Preprint NI13064-POP, Isaac Newton Institute, Cambridge UK.*

[19] Bomze, I. M. and De Klerk, E. (2002). Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24(2):163–185.

[20] Bomze, I. M., Dür, M., De Klerk, E., Roos, C., Quist, A. J., and Terlaky, T. (2000). On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18(4):301–320.

[21] Borenstein, S. (2012). The redistributional impact of nonlinear electricity pricing. *American Economic Journal: Economic Policy*, 4(3):56–90.

[22] Bose, S., Low, S. H., Teeraratkul, T., and Hassibi, B. (2015). Equivalent relaxations of optimal power flow. *IEEE Transactions on Automatic Control*, 60(3):729–742.

[23] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization.* Cambridge University Press.

[24] Brown, D. B. and Smith, J. E. (2011). Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. *Management Science*, 57(10):1752–1770.

[25] Bukhsh, W., Grothey, A., McKinnon, K., and Trodden, P. (2013). Local solutions of the optimal power flow problem. *IEEE Transactions on Power Systems*, 28(4):4780–4788.

[26] Bundfuss, S. and Dür, M. (2009). An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization*, 20(1):30–53.

[27] Burer, S. (2009). On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120(2):479–495.

[28] Burer, S. and Anstreicher, K. M. (2013). Second-order-cone constraints for extended trust-region subproblems. *SIAM Journal on Optimization*, 23(1):432–451.

[29] Burer, S. and Dong, H. (2012). Representing quadratically constrained quadratic programs as generalized copositive programs. *Operations Research Letters*, 40(3):203–206.

[30] Chen, J. and Burer, S. (2012). Globally solving nonconvex quadratic programming problems via completely positive programming. *Mathematical Programming Computation*, 4(1):33–52.

[31] Conforti, M., Cornuéjols, G., and Zambelli, G. (2014). *Integer programming*, volume 271. Springer.

[32] de Klerk, E. and Pasechnik, D. (2002). Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimimization*, 12(4):875–892.

[33] de Klerk, E. and Pasechnik, D. V. (2007). A linear programming reformulation of the standard quadratic optimization problem. *Journal of Global Optimization*, 37(1):75–84.

[34] de Klerk, E. and Sotirov, R. (2010). Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Mathematical Programming*, 122(2):225–246.

[35] Dickinson, P. J. and Povh, J. (2013a). Moment approximations for set-semidefinite polynomials. *Journal of Optimization Theory and Applications*, 159(1):57–68.

[36] Dickinson, P. J. and Povh, J. (2013b). New linear and positive semidefinite programming based approximation hierarchies for polynomial optimisation. *Preprint, submitted. Available at http://www.optimization-online.org/DB_HTML/2013/06/3925.html.*

[37] Dickinson, P. J. and Povh, J. (2015). On an extension of Pólyas positivstellensatz. *Journal of global optimization*, 61(4):615–625.

[38] Dong, H. (2013). Symmetric tensor approximation hierarchies for the completely positive cone. *SIAM Journal on Optimization*, 23(3):1850–1866.

[39] Dukanovic, I. and Rendl, F. (2010). Copositive programming motivated bounds on the stability and the chromatic numbers. *Mathematical programming*, 121(2):249–268.

[40] EIA, A. L. (2017). U.S. wind generating capacity surpasses hydro capacity at the end of 2016. Technical report, EIA.

[41] Garcı, R., Conejo, A. J., Gabriel, S., et al. (2006). Electricity market near-equilibrium under locational marginal pricing and minimum profit conditions. *European Journal of Operational Research*, 174(1):457–479.

[42] Gatermann, K. and Parrilo, P. (2004). Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128.

[43] Ghaddar, B. (2011). New conic optimization techniques for solving binary polynomial programming problem. *PhD thesis*.

[44] Ghaddar, B., Marecek, J., and Mevissen, M. (2016). Optimal power flow as a polynomial optimization problem. *IEEE Transactions on Power Systems*, 31(1):539–546.

[45] Ghaddar, B., Vera, J. C., and Anjos, M. F. (2011). Second-order cone relaxations for binary quadratic polynomial programs. *SIAM Journal on Optimization*, 21(1):391–414.

[46] Ghaddar, B., Vera, J. C., and Anjos, M. F. (2015). A dynamic inequality generation scheme for polynomial programming. *to appear in Mathematical Programming (Series A)*.

[47] Goemans, M. X. (1997). Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79(1-3):143–161.

[48] Gomory, R. E. and Baumol, W. J. (1960). Integer programming and pricing. *Econometrica: Journal of the Econometric Society*, pages 521–550.

[49] Hardy, G. H., Littlewood, J. E., and Pólya, G. (1988). *Inequalities.* Cambridge Mathematical Library.

[50] Herrero, I., Rodilla, P., and Batlle, C. (2015). Electricity market-clearing prices and investment incentives: The role of pricing rules. *Energy Economics*, 47:42–51.

[51] Hobbs, B. F., Rothkopf, M. H., O'Neill, R. P., and Chao, H.-p. (2006). *The next generation of electric power unit commitment models*, volume 36. Springer Science & Business Media.

[52] Hogan, W. W. and Ring, B. J. (2003). On minimum-uplift pricing for electricity markets. *Electricity Policy Group.*

[53] Hu, S., Qi, L., and Zhang, G. (2016). Computing the geometric measure of entanglement of multipartite pure states by means of non-negative tensors. *Physical Review A*, 93(1):012304.

[54] Jabr, R. (2006). Radial distribution load flow using conic programming. *IEEE Transactions on Power Systems*, 21(3):1458–1459.

[55] Jayantilal, A., Cheung, K. W., Shamsollahi, P., and Bresler, F. S. (2001). Market based regulation for the PJM electricity market. In *Power Industry Computer Applications, 2001. PICA 2001. Innovative Computing for Power - Electric Energy Meets the Market. 22nd IEEE Power Engineering Society International Conference on*, pages 155–160.

[56] Jiang, B., Yang, F., and Zhang, S. (2016). Tensor and its tucker core: the invariance relationships. *arXiv preprint arXiv:1601.01469.*

[57] Josz, C., Maeght, J., Panciatici, P., and Gilbert, J. (2015). Application of the moment-SOS approach to global optimization of the OPF problem. *IEEE Transactions on Power Systems*, 30(1):463–470.

[58] Kim, S. and Kojima, M. (2003). Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26(2):143–154.

[59] Kleniati, P.-M., Parpas, P., and Rustem, B. (2010). Partitioning procedure for polynomial optimization. *Journal of Global Optimization*, 48(4):549–567.

[60] Kojima, M., Kim, S., and Waki, H. (2003). Sparsity in sums of squares of polynomials. *Mathematical Programming*, 103(1):45–62.

[61] Kuang, X., Ghaddar, B., Naoum-Sawaya, J., and Zuluaga, L. F. (2015a). Alternative SDP and SOCP approximations for polynomial optimization. *arXiv preprint arXiv:1510.06797*.

[62] Kuang, X., Ghaddar, B., Naoum-Sawaya, J., and Zuluaga, L. F. (2017a). Alternative LP and SOCP hierarchies for acopf problems. *IEEE Transactions on Power Systems*, 32(4):2828–2836.

[63] Kuang, X., Lamadrid, A. J., and Zuluaga, L. F. (2017b). Pricing in non-convex markets with quadratic deliverability costs. *arXiv preprint arXiv:1705.01675*.

[64] Kuang, X. and Zuluaga, L. F. (2016). Completely positive and completely positive semidefinite tensor relaxations for polynomial optimization. *arXiv preprint arXiv:1604.06823*.

[65] Kuang, X., Zuluaga, L. F., Ghaddar, B., and Naoum-Sawaya, J. (2015b). Approximating the ACOPF problem with a hierarchy of SOCP problems. In *Power & Energy Society General Meeting, 2015 IEEE*, pages 1–5. IEEE.

[66] Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817.

[67] Lasserre, J. B. (2002a). An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM Journal on Optimization*, 12(3):756–769.

[68] Lasserre, J. B. (2002b). Semidefinite programming vs. LP relaxations for polynomial programming. *Mathematics of Operations Research*, 27(2):347–360.

[69] Lasserre, J. B. (2009). Moments and sums of squares for polynomial optimization and related problems. *Journal of Global Optimization*, 45(1):39–61.

[70] Lasserre, J. B., Toh, K.-C., and Yang, S. (2015). A bounded degree SOS hierarchy for polynomial optimization. *EURO Journal on Computational Optimization*, pages 1–31.

[71] Lavaei, J. and Low, S. H. (2012). Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107.

[72] Lazimy, R. (1982). Mixed-integer quadratic programming. *Mathematical Programming*, 22(1):332–349.

[73] Lew, D., Brinkman, G., Kumar, N., Besuner, P., Agan, D., and Lefton, S. (2012). Impacts of wind and solar on emissions and wear and tear of fossil-fueled generators. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE.

[74] Liberopoulos, G. and Andrianesis, P. (2016). Critical review of pricing schemes in markets with non-convex costs. *Operations Research*, 64(1):17–31.

[75] Ling, C., Nie, J., Qi, L., and Ye, Y. (2009). Biquadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM Journal on Optimization*, 20(3):1286–1310.

[76] Lobo, M. S., Vandenberghe, L., Boyd, S., and Lebret, H. (1997). Second-order cone programming. *Manuscript, Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford, CA, USA*.

[77] Lubin, M., Dvorkin, Y., and Backhaus, S. (2016). A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Transactions on Power Systems*, 31(5):3840–3849.

[78] Luo, Z., Qi, L., and Ye, Y. (2015). Linear operators and positive semidefiniteness of symmetric tensor spaces. *Science China Mathematics*, 58(1):197–212.

[79] Mak, H.-Y., Rong, Y., and Zhang, J. (2014). Sequencing appointments for service systems using inventory approximations. *Manufacturing & Service Operations Management*, 16(2):251–262.

[80] Malek, A. and Hosseinipour-Mahani, N. (2015). Solving a class of non-convex quadratic problems based on generalized KKT conditions and neurodynamic optimization technique. *Kybernetika*, 51(5):890–908.

[81] Mavridou, T., Pardalos, P., Pitsoulis, L., and Resende, M. G. (1998). A grasp for the biquadratic assignment problem. *European Journal of Operational Research*, 105(3):613–621.

[82] Molzahn, D. and Hiskens, I. (2014). Moment-based relaxation of the optimal power flow problem. In *Power Systems Computation Conference (PSCC), 2014*, pages 1–7.

[83] Molzahn, D. K., Holzer, J. T., Lesieutre, B. C., and DeMarco, C. L. (2013). Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Transactions on Power Systems*, 28(4):3987–3998.

[84] Muatore, G. (2008). Equilibria in markets with non-convexities and a solution to the missing money phenomenon in energy markets. *CORE Discussion Paper 2008/6.*

[85] Murty, K. G. and Kabadi, S. N. (1987). Some NP-complete problems in quadratic and nonlinear programming. *Mathematical programming*, 39(2):117–129.

[86] Nesterov, Y. (1997). Structure of non-negative polynomials and optimization problems. Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

[87] Olivares-Nadal, A. V. and DeMiguel, V. (2015). A robust perspective on transaction costs in portfolio optimization. Technical report, working paper.

[88] O'Neill, R. P., Sotkiewicz, P. M., Hobbs, B. F., Rothkopf, M. H., and Stewart, W. R. (2005). Efficient market-clearing prices in markets with nonconvexities. *European journal of operational research*, 164(1):269–285.

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

[90] Pena, J., Vera, J., and Zuluaga, L. F. (2007). Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105.

[91] Peña, J., Vera, J. C., and Zuluaga, L. F. (2014). A certificate of non-negativity for polynomials over unbounded sets. *Mathematics of Operations Research*.

[92] Pena, J., Vera, J. C., and Zuluaga, L. F. (2014). Completely positive reformulations for polynomial optimization. *Mathematical Programming*, 151(2):405–431.

[93] Povh, J. and Rendl, F. (2007). A copositive programming approach to graph partitioning. *SIAM Journal on Optimization*, 18(1):223–241.

[94] Povh, J. and Rendl, F. (2009). Copositive and semidefinite relaxations of the quadratic assignment problem. *Discrete Optimization*, 6(3):231–241.

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

[96] QELA, R. E. B. E. and KLINZ, B. (1994). On the biquadratic assignment problem. In *Quadratic Assignment and Related Problems: DIMACS Workshop, May 20-21, 1993*, volume 16, page 117. American Mathematical Soc.

[97] Ruiz, C., Conejo, A. J., and Gabriel, S. A. (2012). Pricing non-convexities in an electricity pool. *IEEE Transactions on Power Systems*, 27(3):1334–1342.

[98] Sahinidis, N. V. and Tawarmalani, M. (2005). Baron 7.2.5: Global optimization of mixed-integer nonlinear programs. *Users manual*.

[99] Scarf, H. E. (1990). Mathematical programming and economic theory. *Operations Research*, 38(3):377–385.

[100] Scarf, H. E. (1994). The allocation of resources in the presence of indivisibilities. *The Journal of Economic Perspectives*, 8(4):111–128.

[101] Schmüdgen, K. (1991). The K-moment problem for compact semi-algebraic sets. *Math. Ann.*, 289:203–206.

[102] Shor, N. (1987). Class of global minimum bounds of polynomial functions. *Cybernetics and Systems Analysis*, 23(6):731–734.

[103] Sioshansi, R. and Nicholson, E. (2011). Towards equilibrium offers in unit commitment auctions with nonconvex costs. *Journal of Regulatory Economics*, 40(1):41–61.

[104] Sioshansi, R. and Short, W. (2009). Evaluating the impacts of real-time pricing on the usage of wind generation. *IEEE Transactions on Power Systems*, 24(2):516–524.

[105] Sojoudi, S. and Lavaei, J. (2012). Physics of power networks makes hard optimization problems easy to solve. In *IEEE Power and Energy Society General Meeting*, pages 1–8.

[106] Song, Y. and Qi, L. (2015). Tensor complementarity problem and semi-positive tensors. *Journal of Optimization Theory and Applications*, pages 1–10.

[107] Starr, R. M. (1969). Quasi-equilibria in markets with non-convex preferences. *Econometrica: Journal of the Econometric Society*, pages 25–38.

[108] Sturm, J. F. (1999). Using SeDuMi 1.02, a matlab toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653.

[109] Terlaky, T. (2013). Mixed-integer conic linear programming: Challenges and perspectives. Technical report, DTIC Document.

[110] Todd, M. (2001). Semidefinite optimization. *Acta Numerica*, 10:515–560.

[111] University of Washington (2015). Power systems test case archive. *http://www.ee.washington.edu/research/pstca/.*

[112] Van Vyve, M. et al. (2011). Linear prices for non-convex electricity markets: models and algorithms. *CORE Discussion Paper 2011/50.*

[113] Vandenberghe, L. and Boyd, S. (1996). Semidefinite programming. *SIAM review*, 38(1):49–95.

[114] Waki, H., Kim, S., Kojima, M., and Muramatsu, M. (2006). Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242.

[115] Waki, H., Kim, S., Kojima, M., Muramatsu, M., and Sugimoto, H. (2008). Algorithm 883: Sparsepop—a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Transactions on Mathematical Software (TOMS)*, 35(2):15.

[116] Yang, B., Anstreicher, K., and Burer, S. (2016). Quadratic programs with hollows. Technical report, Technical report, January 2016. http://www.optimization-online.org/DB_FILE/2016/01/5277.pdf.

[117] Ye, Y. (1999). Approximating quadratic programming with bound and quadratic constraints. *Mathematical programming*, 84(2):219–226.

[118] Ye, Y., Papadaskalopoulos, D., and Strbac, G. (2015). Factoring flexible demand non-convexities in electricity markets. *IEEE Transactions on Power Systems*, 30(4):2090–2099.

[119] Zheng, X. J., Sun, X. L., and Li, D. (2011). Convex relaxations for nonconvex quadratically constrained quadratic programming: matrix cone decomposition and polyhedral approximation. *Mathematical programming*, 129(2):301–329.

[120] Zoltowska, I. (2016). Demand shifting bids in energy auction with non-convexities and transmission constraints. *Energy Economics*, 53:17–27.

[121] Zuluaga, L., Vera, J., and Peña, J. (2006). LMI approximations for cones of positive semidefinite forms. *SIAM Journal on Optimization*, 16:1076–1091.

Vita

Education

Ph.D.	Industrial Engineering, Lehigh University	2017.09
M.Eng.	Industrial and Systems Engineering, Lehigh University	2014.09
B.S.	Systems Engineering, Huazhong University of Science and Technology	2012.05

Experiences

Amazon.com, Inc.	Research Scientist	2017.06 - Present
Apple Inc.	Operations Research Scientist Intern	2016.05 - 2017.05
United Airlines	Enterprise Optimization Intern	2015.05 - 2015.08
IBM Research - Ireland	Ph.D. Research Intern	2014.06 - 2014.09