Journal of Convex Analysis
Volume 22 (2015), No. 4, 999–1023

Generalized SOS-Convexity and Strong Duality with SDP Dual Programs in Polynomial Optimization*

V. Jeyakumar
Department of Applied Mathematics,
University of New South Wales, Sydney 2052, Australia
v.jeyakumar@unsw.edu.au

G. M. Lee
Department of Applied Mathematics,
Pukyong National University, Busan 608-737, Korea
gmlee@pknu.ac.kr

J. H. Lee
Department of Applied Mathematics,
Pukyong National University, Busan 608-737, Korea
mc7558@naver.com

Received: August 5, 2014
Revised manuscript received: December 4, 2014
Accepted: December 5, 2014

In this paper, we introduce the notion of ρ-SOS-convexity, extending the numerically checkable concept of SOS-convexity of a real polynomial. The class of ρ-SOS-convex polynomials includes the important class of (not necessarily convex) quadratic functions. We provide various characterizations of ρ-SOS-convexity in terms of SOS-convexity. Consequently, we establish strong duality results for classes of nonconvex polynomial optimization problems involving strong SOS-convex (where $\rho > 0$) and weak SOS-convex (where $\rho < 0$) polynomials. These classes of problems include some polynomial optimization problems, involving SOS-convex polynomials, minimax quadratic optimization problems with quadratic constraints, fractional programming problems and robust optimization problems. Our results also provide necessary and sufficient conditions for strong duality of some classes of minimax quadratic optimization problems and extended trust-region problems.

Keywords: Strong duality, ρ-SOS-convex polynomials, SOS-convex polynomials, non-convex quadratic optimization, extended trust-region problems

1991 Mathematics Subject Classification: 26A51, 90C25, 47N10

*The research of the first author was supported by a grant from the Australian Research Council. The second and third authors were supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2013R1A1A2005378)

ISSN 0944-6532 / $ 2.50

1. Introduction

Convexity of sets and functions has played a key role in the development of optimization and its applications. It has proved to be a powerful tool for Lagrange multiplier theory and numerical methods of optimization. However, verifying convexity is, in general, difficult numerically for real polynomials, especially when the degree of the polynomial is four or higher. The SOS-convexity ([1, 13]) is a numerically tractable relaxation of convexity for a real polynomial and it can easily be checked whether a real polynomial is SOS-convex or not by solving a related semi-definite programming problem.

Recent research on SOS-convexity and polynomial optimization has produced numerically attractive semi-definite programming (SDP) relaxation schemes for solving various classes of polynomial optimization problems [13, 17, 19, 21]. In particular, under mild conditions, SOS-convex optimization problems such as the minimization of SOS-convex polynomials subject to SOS-convex inequality constraints enjoy exact SDP relaxation [19, 21] (strong duality [18]) in the sense that the optimal values of the given SOS-convex problem and its sum-of-squares polynomial relaxation problem (dual semidefinite program) are equal and the relaxation problem (dual problem) attains its optimum.

Exact semidefinite programming relaxation or strong duality involving dual semidefinite programs is a highly desirable property particularly for nonconvex global polynomial or quadratic optimization problems because such problems can efficiently be solved via interior point methods. In recent years, it has received a great deal of attention in the literature due to its applications to emerging areas such as robust optimization (see [5, 7, 8, 16] and other references therein).

This has provided us with the motivation for introducing, in this paper, a numerically tractable relaxation of ρ-convexity (see [24, 25]), called ρ-SOS-convexity (see Definition 2.1), generalizing SOS-convexity for real polynomials. Using ρ-SOS-convexity, we examine strong duality results for various classes of nonconvex polynomial optimization problems which enjoy semidefinite programs as their dual programs.

We make the following contributions which extend the numerically attractive features of SOS-convex polynomials and the associated convex optimization problems to a broader class of ρ-SOS-convex polynomials and to the associated nonconvex optimization problems, involving strong SOS-convex polynomials (where $\rho > 0$) and weak SOS-convex polynomials (where $\rho < 0$).

(i) We introduce the notion of ρ-SOS-convexity for a real polynomial and show that a real polynomial is ρ-SOS-convex if and only if it is a sum of a SOS-convex polynomial and $\rho\|\cdot\|^2$, where $\rho \in \mathbb{R}$ and $\|\cdot\|$ is the Euclidian norm. This characterization shows us whether a real polynomial is ρ-SOS-convex or not can be numerically checked by solving a semidefinite programming problem. This is a key distinguishing feature of ρ-SOS-convexity that is generally not shared by ρ-convexity [10, 14, 24, 25, 26]). *Importantly, we obtain that every quadratic function is a ρ-SOS-convex polynomial and also derive a new characterization of SOS-convexity.*

(ii) Under an easily checkable Lagrange multiplier condition, we establish that strong duality holds with attainment of their optimal values for a class of non-convex polynomial optimization problems. Consequently, we derive corresponding strong duality results for minimax ρ-SOS-convex polynomial optimization problems. In particular, we present a *necessary and sufficient condition* for strong duality of (not necessarily convex) *minimax quadratic optimization problems with quadratic constraints*.

(iii) As an application, we establish strong duality results for the following classes of nonlinear programming problems: ρ-SOS-convex fractional programming problems and robust optimization problems.

(iv) Finally, we establish that the Lagrange multiplier condition completely characterizes strong duality for an important class of *extended trust-region problems* (see [5, 9, 16] for recent work to extended trust-region problems).

The outline of the paper is as follows. Section 2 gives various characterizations of ρ-SOS-convexity, including the characterization for the gap between ρ-SOS-convexity and ρ-convexity. Section 3 presents strong duality results for ρ-SOS-convex polynomial optimization problems and derives strong duality results for minimax polynomial optimization problems including quadratically constrained minimax quadratic programming problems. Consequently, in Section 4, we derive strong duality results for fractional optimization problems and robust optimization problems. Section 5 presents necessary and sufficient Lagrange multiplier condition for strong duality of a class of extended trust-region problems. Section 6 provides a discussion on a potential generalization and application of the present work to minimax trust-region problems, including the study of trust-region problems under data uncertainty.

2. Generalized SOS-Convex Polynomials

We begin this section by fixing notation and preliminaries of convex sets, functions and polynomials. Throughout this paper, $\mathbb{R}^n$ denotes the Euclidean space with dimension n. The nonnegative orthant of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$ and is defined by $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geqq 0\}$. Moreover, the Euclidean norm on $\mathbb{R}^n$ is denoted by $\|\cdot\|$.

We say A is convex whenever $\mu a_1 + (1-\mu)a_2 \in A$ for all $\mu \in [0,1]$, $a_1, a_2 \in A$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be convex if for all $\mu \in [0,1]$ and all $x, y \in \mathbb{R}^n$,

$$f((1-\mu)x + \mu y) \leqq (1-\mu)f(x) + \mu f(y).$$

The function f is said to be concave whenever $-f$ is convex. The positive semidefiniteness of an $n \times n$ matrix B, denoted by $B \succeq 0$, is defined by $\langle x, Bx \rangle \geqq 0$, for each $x \in \mathbb{R}^n$. Moreover, a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be ρ-convex

[10, 14, 24, 25, 26] if there exists some real number ρ such that for all $\mu \in [0,1]$ and all $x, y \in \mathbb{R}^n$,

$$f((1-\mu)x + \mu y) \leqq (1-\mu)f(x) + \mu f(y) - \rho\mu(1-\mu)\|x-y\|^2.$$

We say that a real polynomial f is sum of squares if there exist real polynomials f_j, $j = 1, \ldots, r$, such that $f = \sum_{j=1}^r f_j^2$. The set consisting of all sum of squares real polynomial is denoted by Σ^2. The degree of a polynomial f is denoted by deg$f$. Moreover, the set consisting all sum of squares real polynomial of degree at most d is denoted by Σ_d^2. We say a matrix polynomial $F \in \mathbb{R}[x]^{n\times n}$ is a SOS matrix polynomial if $F(x) = H(x)H(x)^T$ where $H(x) \in \mathbb{R}[x]^{n\times r}$ is a matrix polynomial for some $r \in \mathbb{N}$.

A real polynomial f on $\mathbb{R}^n$ is called *SOS-convex* [1, 13] if the Hessian matrix function $H : x \mapsto \nabla^2 f(x)$ is a SOS matrix polynomial, equivalently, for all $x, y \in \mathbb{R}^n$, and for all $\lambda \in [0,1]$,

$$\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y)$$

is a *sum of squares polynomial in* $\mathbb{R}[x; y]$. Moreover, we say f is SOS-concave if $-f$ is SOS-convex. For properties of SOS-convex polynomials and related optimization problems, see [1, 2, 13, 17, 19]. The class of SOS-convex polynomials includes separable convex polynomials and convex quadratic functions as their special cases. The important feature of SOS-convexity, which distinguishes from convexity for polynomials is that one can numerically check whether a polynomial is SOS-convex or not by solving a related semi-definite optimization (feasibility) problem [21]. Clearly, a SOS-convex polynomial is convex. The gap between SOS-convex polynomial and convex polynomial is completely characterized in [2].

A numerically tractable relaxation of the ρ-convexity for real polynomials is given as follows.

Definition 2.1 (ρ-SOS-Convexity). Let $\rho \in \mathbb{R}$. A real polynomial f on $\mathbb{R}^n$ is said to be ρ-SOS-convex if there exists some real number ρ such that for all $x, y \in \mathbb{R}^n$ and for all $\lambda \in [0,1]$,

$$\lambda f(x) + (1-\lambda)f(y) - \rho\lambda(1-\lambda)\|x-y\|^2 - f(\lambda x + (1-\lambda)y)$$

is a sum of squares polynomial in $\mathbb{R}[x; y]$. If $\rho > 0$, then f is called a strong SOS-convex polynomial with modulus ρ. If $\rho = 0$, then f is a SOS-convex polynomial. If $\rho < 0$, then f is called a weak SOS-convex polynomial with modulus ρ.

Clearly, it follows from the definition that strong SOS-convexity implies SOS-convexity which, in turn, implies weak SOS-convexity. Moreover, a ρ-SOS-convex polynomial is SOS-convex if and only if if $\rho \geq 0$. Now we give a basic characterization of the ρ-SOS-convex polynomial.

Theorem 2.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real polynomial and let $\rho \in \mathbb{R}$. Then following statements are equivalent:*

(*i*) *f is a ρ-SOS-convex polynomial;*

(*ii*) *There exists a SOS-convex polynomial $h : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) = h(x) + \rho\|x\|^2$.*

Proof. Let x_1, $x_2 \in \mathbb{R}^n$, $\lambda \in [0,1]$ and $x_\lambda = \lambda x_1 + (1-\lambda)x_2$. Note that $\|x_\lambda\|^2 = \lambda\|x_1\|^2 + (1-\lambda)\|x_2\|^2 - \lambda(1-\lambda)\|x_1 - x_2\|^2$. Suppose that f is ρ-SOS-convex polynomial. It means that, for all x_1, $x_2 \in \mathbb{R}^n$ and for all $\lambda \in [0,1]$,

$$\lambda f(x_1) + (1-\lambda)f(x_2) - \rho\lambda(1-\lambda)\|x_1 - x_2\|^2 - f(\lambda x_1 + (1-\lambda)x_2) \in \Sigma^2.$$

Equivalently, for all x_1, $x_2 \in \mathbb{R}^n$ and for all $\lambda \in [0,1]$,

$$\begin{aligned}\lambda f(x_1) + (1-\lambda)f(x_2) + \rho\|x_\lambda\|^2 - \rho\lambda\|x_1\|^2 \\ - \rho(1-\lambda)\|x_2\|^2 - f(\lambda x_1 + (1-\lambda)x_2) \in \Sigma^2,\end{aligned}$$

that is, for all x_1, $x_2 \in \mathbb{R}^n$ and for all $\lambda \in [0,1]$,

$$\begin{aligned}\lambda f(x_1) + (1-\lambda)f(x_2) - \rho\lambda\|x_1\|^2 - \rho(1-\lambda)\|x_2\|^2 \\ - \left[f(\lambda x_1 + (1-\lambda)x_2) - \rho\|\lambda x_1 + (1-\lambda)x_2\|^2\right] \in \Sigma^2.\end{aligned}$$

It means that $f - \rho\|\cdot\|^2$ is a SOS-convex function. □

We now show that every quadratic function is a ρ-SOS-convex polynomial where ρ is the least eigenvalue of the Hessian of the quadratic function. Recall first that, for a matrix A, the least eigenvalue of A is denoted by $\gamma_{\min}(A)$.

Corollary 2.3. *Let the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ be given by $x^T Ax + a^T x + a_0$, where A is a symmetric $(n \times n)$ matrix, $a \in \mathbb{R}^n$ and $a_0 \in \mathbb{R}$. Then, f is ρ-SOS-convex with $\rho = \gamma_{\min}(A)$.*

Proof. The quadratic function f can be written as, for each $x \in \mathbb{R}^n$,

$$f(x) = x^T Ax + a^T x + a_0 = x^T(A - \gamma_{\min}(A)I_n)x + a^T x + a_0 + \gamma_{\min}(A)\|x\|^2.$$

Since $(A - \gamma_{\min}(A)I_n) \succeq 0$, the function $h(x) = x^T(A - \gamma_{\min}(A)I_n)x + a^T x + a_0$ is a convex quadratic function which is SOS-convex, and, for each $x \in \mathbb{R}^n$, $f(x) = h(x) + \gamma_{\min}(A)\|x\|^2$. It now follows from Theorem 2.2 that f is a ρ-SOS-convex function with $\rho = \gamma_{\min}(A)$. □

Let us give an example of a ρ-SOS convex function that is neither quadratic nor SOS-convex.

Example 2.4. Let $f(x_1, x_2) = x_1^8 + x_1x_2$, Then, f is not convex and so is not SOS-convex. However, $f(x_1, x_2) = h(x_1, x_2) - (x_1^2 + x_2^2)$, where $h(x_1, x_2) = x_1^8 + x_1x_2 + x_1^2 + x_2^2$ is SOS-convex. So, $f(x_1, x_2)$ is ρ-SOS-convex with $\rho = -1$.

The following characterizations of SOS-convex functions of [2] play an important role in obtaining corresponding characterizations for ρ-SOS-convex polynomials.

Lemma 2.5 ([2]). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real polynomial function. Then the following statements are equivalent:*

(*i*) f *is a SOS-convex polynomial;*
(*ii*) $f(x) - f(y) - \nabla f(y)^T(x-y)$ *is a sum of squares polynomial in* $\mathbb{R}[x;y]$*;*
(*iii*) $y^T\nabla^2 f(x)y$ *is a sum of squares polynomial in* $\mathbb{R}[x;y]$*;*
(*iv*) $\nabla^2 f(x)$ *is a SOS matrix.*

Corollary 2.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real polynomial function and let $\rho \in \mathbb{R}$. Then following statements are equivalent:*

(*i*) f *is a ρ-SOS-convex polynomial;*
(*ii*) $\langle \nabla f(x) - \nabla f(y), x-y\rangle - 2\rho\|x-y\|^2$ *is a sum of squares polynomial in* $\mathbb{R}[x;y]$.

Proof. By Theorem 2.2, (*i*) is equivalent to the following statement: There exists a SOS-convex polynomial $h : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) = h(x) + \rho\|x\|^2$. So, $\nabla f(x) = \nabla h(x) + 2\rho x$. By Lemma 2.5,

$$h(x) - h(y) - \langle \nabla h(y), x-y\rangle$$

is a sum of squares polynomial and $h(y) - h(x) - \langle \nabla h(x), y-x\rangle$ is a sum of squares polynomial. So, $\langle \nabla h(x) - \nabla h(y), x-y\rangle$ is a sum of squares polynomial. Notice that

$$\begin{aligned}
&\langle \nabla f(x) - \nabla f(y), x-y\rangle - 2\rho\|x-y\|^2\\
&= \langle \nabla h(x) + 2\rho x - \nabla h(y) - 2\rho y, x-y\rangle - 2\rho\|x-y\|^2\\
&= \langle \nabla h(x) - \nabla h(y), x-y\rangle.
\end{aligned}$$

Thus $\langle \nabla f(x) - \nabla f(y), x-y\rangle - 2\rho\|x-y\|^2$ is a sum of squares polynomial. Conversely, suppose that $\langle \nabla f(x) - \nabla f(y), x-y\rangle - 2\rho\|x-y\|^2 \in \Sigma^2$. Then

$$\langle \nabla f(x) - 2\rho x - (\nabla f(y) - 2\rho y), x-y\rangle \in \Sigma^2. \tag{1}$$

Let $h(x) = f(x) - \rho\|x\|^2$, $x \in \mathbb{R}^n$. Then (1) is equivalent to the following statement:

$$\langle \nabla h(x) - \nabla h(y), x-y\rangle \in \Sigma^2. \tag{2}$$

By the mean value theorem,

$$h(x) - h(y) = \langle \nabla h(z), x-y\rangle, \tag{3}$$

where $z = \lambda x + (1-\lambda)y$ for some $\lambda \in (0,1)$. By (2),

$$\begin{aligned}
\langle \nabla h(z) - \nabla h(y), z-y\rangle &= \langle \nabla h(z) - \nabla h(y), \lambda(x-y)\rangle\\
&= \lambda\langle \nabla h(z) - \nabla h(y), x-y\rangle \in \Sigma^2.
\end{aligned}$$

This implies that $\langle \nabla h(z) - \nabla h(y), x-y\rangle \in \Sigma^2$. By (3), we have $h(x) - h(y) - \langle \nabla h(y), x-y\rangle \in \Sigma^2$. So, by Lemma 2.5, h is a SOS-convex polynomial. Since $f(x) = h(x) + \rho\|x\|^2$, $x \in \mathbb{R}^n$, it follows from Theorem 2.2 that f is a ρ-SOS-convex polynomial. $\square$

The special case of $\rho = 0$ in the preceding corollary gives us a new characterization of SOS-convex polynomials as follows.

Corollary 2.7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real polynomial function. Then the following statements are equivalent:*

(*i*) *f is SOS-convex;*
(*ii*) *$\langle \nabla f(x) - \nabla f(y), x-y\rangle$ is a sum of squares polynomial in $\mathbb{R}[x;y]$.*

Proof. The conclusion follows easily from Corollary 2.6 by letting $\rho = 0$. $\square$

Corollary 2.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a ρ-SOS-convex polynomial. Suppose that there exists $\bar{x} \in \mathbb{R}^n$ such that $\nabla f(\bar{x}) = 2\rho\bar{x}$ and $f(\bar{x}) = \rho\|\bar{x}\|^2$. Then $f - \rho\|\cdot\|^2$ is a sum of squares polynomial.*

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a ρ-SOS-convex polynomial. Then by Theorem 2.2, there exists a SOS-convex polynomial $h : \mathbb{R}^n \to \mathbb{R}$ such that $f = h + \rho\|\cdot\|^2$. By assumption, $\nabla h(\bar{x}) = 0$ and $h(\bar{x}) = 0$. Since h is SOS-convex, it follows from Lemma 2.5, $h(x) - h(\bar{x}) - \nabla h(\bar{x})^T(x-\bar{x})$ is a sum of squares polynomial. Since $h(x) = h(x) - h(\bar{x}) - \nabla h(\bar{x})^T(x-\bar{x})$, h is a sum of squares polynomial and hence $f - \rho\|\cdot\|^2$ is a sum of squares polynomial. $\square$

From Lemma 2.5 and Theorem 2.2, we can get easily the following useful characterizations of ρ-SOS-convex functions.

Corollary 2.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real polynomial function. Then the following statements are equivalent:*

(*i*) *f is a ρ-SOS-convex polynomial;*
(*ii*) *$f(x) - f(y) - \nabla f(y)^T(x-y) - \rho\|x-y\|^2$ is a sum of squares polynomial;*
(*iii*) *$y^T(\nabla^2 f(x) - 2\rho I)y$ is a sum of squares polynomial, where I is an identity matrix;*
(*iv*) *$\nabla^2 f(x) - 2\rho I$ is a SOS matrix.*

Proof. [(*i*) $\Leftrightarrow$ (*ii*)] If f is ρ-SOS-convex then by Theorem 2.2, $f(\cdot) = h + \rho\|\cdot\|^2$, where h is SOS-convex. Let $x, y \in \mathbb{R}^n$ be arbitrary. From the SOS-convexity of h and Lemma 2.5,

$$h(x) - h(y) - \nabla h(y)^T(x-y) = \sigma(x,y),$$

where σ is a sum of squares polynomial in x and y. Multiplying by ρ the identity

$$\|x\|^2 - \|y\|^2 - 2y^T(x-y) - \|x-y\|^2 = 0$$

and adding it to the preceding equality yields (ii).

Conversely, suppose that (ii) holds. Let $x, y \in \mathbb{R}^n$ be arbitrary. Then

$$f(x) - f(y) - \nabla f(y)^T(x-y) - \rho\|x-y\|^2 = \sigma(x,y), \tag{4}$$

where σ is a sum of squares polynomial in x and y. Let $h = f - \rho\|\cdot\|^2$. Subtracting the identity

$$\rho\|x\|^2 - \rho\|y\|^2 - 2\rho y^T(x-y) - \rho\|x-y\|^2 = 0$$

from (4) yields

$$h(x) - h(y) - \nabla h(y)^T(x-y) = \sigma(x,y).$$

So, by Lemma 2.5, h is SOS-convex. Hence, $f = h + \rho\|\cdot\|^2$ and (i) follows from Theorem 2.2.

The conclusion follows because $(i) \Leftrightarrow (iii)$ and $(i) \Leftrightarrow (iv)$ follow easily from Theorem 2.2 and Lemma 2.5. $\square$

The set of convex and SOS-convex polynomials in n variables of degree d are denoted by $\tilde{C}_{n,d}$ and $\Sigma\tilde{C}_{n,d}$, respectively. Using this definition, Ahmadi and Parrilo [2] obtained the following characterization of the classes of SOS-convex and convex polynomials.

Lemma 2.10 ([2]). *$\tilde{C}_{n,d} = \Sigma\tilde{C}_{n,d}$ if and only if $n = 1$ or $d = 2$ or $(n,d) = (2,4)$.*

Now, denote the set of ρ-convex polynomials and ρ-SOS-convex polynomials in n variables of degree d with modulus ρ by $\tilde{C}^{\rho}_{n,d}$ and $\tilde{\Sigma C}^{\rho}_{n,d}$ respectively. As a consequence of Lemma 2.10 and Theorem 2.2, we obtain a complete characterization of the gap between ρ-convexity and ρ-SOS-convexity of polynomials.

Theorem 2.11. *$\tilde{C}^{\rho}_{n,d} = \Sigma\tilde{C}^{\rho}_{n,d}$ if and only if $n = 1$ or $d = 2$ or $(n,d) = (2,4)$.*

Proof. It is clear that $\Sigma\tilde{C}^{\rho}_{n,d} \subset \tilde{C}^{\rho}_{n,d}$. We consider the cases that $n = 1$ or $d = 2$ or $(n,d) = (2,4)$. Let $p \in \tilde{C}^{\rho}_{n,d}$. Then $p - \rho\|\cdot\|^2 \in \tilde{C}_{n,d}$. If $n = 1$ or $(n,d) = (2,4)$, from Lemma 2.10, $p - \rho\|\cdot\|^2 \in \Sigma\tilde{C}_{n,d}$. Thus by Theorem 2.2, $p \in \Sigma\tilde{C}^{\rho}_{n,d}$. If $d = 2$, $p - \rho\|\cdot\|^2 \in \tilde{C}_{n,2} \cup \tilde{C}_{n,1} \cup \mathbb{R}_+$, from Lemma 2.10, $p - \rho\|\cdot\|^2 \in \Sigma\tilde{C}_{n,2} \cup \Sigma\tilde{C}_{n,1} \cup \mathbb{R}_+$, and hence from Theorem 2.2, $p \in \Sigma\tilde{C}^{\rho}_{n,2}$.

Assume that $n \neq 1$, $d \neq 2$ and $(n,d) \neq (2,4)$. Then from Lemma 2.10, there exists a polynomial $h \in \tilde{C}_{n,d}$ which is not in $\Sigma\tilde{C}_{n,d}$. Thus $h + \rho\|\cdot\|^2$ is a ρ-convex polynomial. Suppose that $h + \rho\|\cdot\|^2$ is a ρ-SOS-convex polynomial. Then from Theorem 2.2, h is a SOS-convex polynomial. This is a contradiction. So, $h + \rho\|\cdot\|^2$ is not a ρ-SOS-convex polynomial. Hence $\Sigma\tilde{C}^{\rho}_{n,d} \subset \tilde{C}^{\rho}_{n,d}$, but $\Sigma\tilde{C}^{\rho}_{n,d} \neq \tilde{C}^{\rho}_{n,d}$. $\square$

3. Strong Duality with SDP Duals

Consider the polynomial optimization problem:

$$(\mathbf{P}) \quad \begin{array}{ll} \inf & f(x) \\ \text{s.t.} & g_i(x) \leqq 0, \ i = 1, \ldots, m, \end{array}$$

where f and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ are polynomials and the feasible set of $(\mathbf{P})$ is given by $K := \{x \in \mathbb{R}^n : g_i(x) \leqq 0, \ i = 1, \ldots, m\}$. We assume that $K \neq \emptyset$.

Let

$$\operatorname{argmin}(\mathbf{P}) := \{x \in K \ : \ f(x) \leqq f(y), \forall y \in K\}.$$

The point $\bar{x} \in K$ is said to be a minimizer of $(\mathbf{P})$ if for any $x \in K$, $f(x) \geqq f(\bar{x})$.

We say that the Kuhn-Tucker condition holds at $\bar{x} \in K$ whenever

$$(\exists \bar{\lambda} \in \mathbb{R}^m_+) \ \ \nabla f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla g_i(\bar{x}) = 0, \quad \bar{\lambda}_i g_i(\bar{x}) = 0, \ i = 1, \ldots, m. \tag{5}$$

It is well known that if $\bar{x} \in K$ is a minimizer of $(\mathbf{P})$ and a suitable constraint qualification holds, then the Kuhn-Tucker condition holds at $\bar{x}$.

The Lagrangian dual problem for $(\mathbf{P})$ is given by

$$\sup_{\lambda \in \mathbb{R}^m_+} \inf_{x \in \mathbb{R}^n} \left\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) \right\},$$

which can equivalently be written as

$$\sup_{(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}} \left\{ \mu : \ f(x) + \sum_{i=1}^m \lambda_i g_i(x) - \mu \geqq 0, \forall x \in \mathbb{R}^n \right\}.$$

The dual problem for $(\mathbf{P})$ is the sum-of-squares-polynomial relaxation of the preceding problem which is given by

$$(\mathbf{D}) \quad \sup_{(\lambda, \mu) \in \mathbb{R}^m_+ \times \mathbb{R}} \left\{ \mu : \ f + \sum_{i=1}^m \lambda_i g_i - \mu \in \Sigma^2_d \right\},$$

where d is the smallest even number such that $d \geqq \max\{\deg f, \max_{i \in \{1,\ldots,m\}} \deg g_i\}$. Note that the dual problem $(\mathbf{D})$ is a form of the Lasserre SDP relaxation problem [21].

It is well known (see [4, 21]) that problem $(\mathbf{D})$ can be equivalently rewritten as a semi-definite linear programming problem which can be solved efficiently. For instance, consider the quadratic polynomial optimization problem:

$$(\mathbf{QP_1}) \quad \begin{array}{ll} \text{Minimize} & x^T Q_0 x + c_0^T x + \alpha_0 \\ \text{subject to} & x^T Q_i x + c_i^T x + \alpha_i \leqq 0, \ i = 1, \ldots, m. \end{array}$$

We assume throughout the paper that Q_0 and Q_i, $i = 1,\ldots,m$ are $n \times n$ symmetric matrices, $c_i \in \mathbb{R}^n$ and $\alpha_i \in \mathbb{R}$.

The dual problem $\mathbf{D}$ for $(\mathbf{QP_1})$ collapses to the semidefinite program

$$(\mathbf{QD_1}) \sup_{\mu\in\mathbb{R},\lambda_i\geqq 0} \left\{\mu : \begin{pmatrix} 2(Q_0+\sum_{i=1}^m \lambda_i Q_i) & c_0+\sum_{i=1}^m \lambda_i c_i \\ (c_0+\sum_{i=1}^m \lambda_i c_i)^T & 2(\alpha_0+\sum_{i=1}^m \lambda_i\alpha_i-\mu) \end{pmatrix} \succeq 0\right\}.$$

In passing, observe that, for $(\mathbf{QP_1})$, the dual problem $(\mathbf{QD_1})$ is equivalent to the Lagrangian dual problem of $(\mathbf{QP_1})$.

Recall that optimal values of $(\mathbf{P})$ and $(\mathbf{D})$ are denoted by $\inf(\mathbf{P})$ and $\sup(\mathbf{D})$. When the optimal values of $(\mathbf{P})$ and $(\mathbf{D})$ are attained at some feasible points they are denoted by $\min(\mathbf{P})$ and $\max(\mathbf{D})$ respectively.

Theorem 3.1 (Strong Duality: Necessary as well as Sufficient Condition). *For the problem* $(\mathbf{P})$*, let* f *and* $g_i, i = 1,2,\ldots,m$ *be polynomials,* **and let** $d = \max\{\deg f, \deg g_1,\ldots,\deg g_m\}$*. Then, the following statements hold:*

(*i*) *If the Kuhn-Tucker condition* (5) *holds for* $(\mathbf{P})$ *at a local minimizer* $\bar{x}$ *with the Lagrange multipliers* $\bar{\lambda}_i \geqq 0, i = 1,2,\ldots,m$*, and if* f *and* g_i *are respectively* ρ_0*-SOS-convex and* ρ_i*-SOS-convex for* $i = 1,2,\ldots,m$ *such that* $\rho_0+\sum_{i=1}^m \bar{\lambda}_i\rho_i \geqq 0$*, then argmin*$(\mathbf{P}) \neq \emptyset$, $\min(\mathbf{P}) = \max(\mathbf{D})$ *and* $(\bar{\lambda}, f(\bar{x}))$ *is a maximizer of* $(\mathbf{D})$*;*

(*ii*) *If argmin*$(\mathbf{P}) \neq \emptyset$, $\min(\mathbf{P}) = \max(\mathbf{D})$ *and* $(\mathbf{D})$ *attains its maximum, then the Kuhn-Tucker condition* (5) *holds at each minimizer of* $(\mathbf{P})$*.*

Proof. (*i*) We first note from the definition of ρ-SOS-convexity that $f+\sum_{i=1}^m \bar{\lambda}_i g_i$ is $\rho_0+\sum_{i=1}^m \bar{\lambda}_i\rho_i$-SOS-convex. Then, $f+\sum_{i=1}^m \bar{\lambda}_i g_i$ is SOS-convex if and only if $\rho_0+\sum_{i=1}^m \bar{\lambda}_i\rho_i \geqq 0$. So, it follows from Lemma 2.5 that

$$f(x)+\sum_{i=1}^m \bar{\lambda}_i g_i(x) - f(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i g_i(\bar{x}) - \left[\nabla f(\bar{x})+\sum_{i=1}^m \bar{\lambda}_i \nabla g_i(\bar{x})\right]^T (x-\bar{x})$$

is a sum of squares polynomial. It then follows from the Kuhn-Tucker condition at $\bar{x}$ that $f+\sum_{i=1}^m \bar{\lambda}_i g_i - f(\bar{x})$ is a sum of squares polynomial and so,

$$f+\sum_{i=1}^m \bar{\lambda}_i g_i - f(\bar{x}) \in \Sigma_d^2. \tag{6}$$

So, from (6), we obtain that argmin$(\mathbf{P}) \neq \emptyset$ as $\bar{x}$ is a global minimizer of $(\mathbf{P})$. Moreover, $\max_{\mu\in\mathbb{R},\lambda_i\geqq 0}\{\mu : \ f+\sum_{i=1}^m \lambda_i g_i - \mu \in \Sigma_d^2\} \geqq f(\bar{x})$. This together with the inequality, $\inf(\mathbf{P}) \geqq \sup(\mathbf{D})$, which holds by construction, yields the required equality,

$$f(\bar{x}) = \max_{\mu\in\mathbb{R},\lambda_i\geqq 0}\left\{\mu : \ f+\sum_{i=1}^m \lambda_i g_i - \mu \in \Sigma_d^2\right\},$$

and $(\bar{\lambda}, f(\bar{x}))$ is a maximizer of **(D)**.

(ii) Assume that argmin$(\mathbf{P}) \neq \emptyset$, $\min(\mathbf{P}) = \max(\mathbf{D})$ and **(D)** attains its maximum. Let $\bar{x} \in K$ be an arbitrary minimizer of **(P)**. Then,

$$f(\bar{x}) = \max_{\mu \in \mathbb{R}, \lambda_i \geqq 0} \left\{ \mu : \ f + \sum_{i=1}^{m} \lambda_i g_i - \mu \in \Sigma_d^2 \right\}.$$

So, there exist $\bar{\lambda}_i \geq 0, i = 1, 2, \ldots, m$ and $\mu \in \mathbb{R}$ such that $f(\bar{x}) = \mu$ and $f + \sum_{i=1}^{m} \bar{\lambda}_i g_i - \mu \in \Sigma_d^2$. This gives us that $f + \sum_{i=1}^{m} \bar{\lambda}_i g_i - f(\bar{x}) \in \Sigma_d^2$. Thus, there exists $\sigma \in \Sigma_d^2$ such that, for each $x \in \mathbb{R}^n$, $f(x) + \sum_{i=1}^{m} \bar{\lambda}_i g_i(x) - f(\bar{x}) = \sigma(x)$. Letting $x = \bar{x}$, we see that $\sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}) = \sigma(\bar{x}) \geq 0$. As $\bar{x} \in K$ and $\bar{\lambda}_i \geq 0, i = 1, 2, \ldots, m$, $\sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}) = 0$ and $\sigma(\bar{x}) = 0$. This shows us that $\bar{x}$ is a minimizer of σ since for each $x \in \mathbb{R}^n$, $\sigma(x) \geq 0 = \sigma(\bar{x})$. Then, by the necessary condition for a minimizer of σ and $\bar{x}$, we obtain that $\nabla\sigma(\bar{x}) = 0$. Hence, $\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0$ and so the Kuhn-Tucker condition (5) is satisfied at $\bar{x}$. $\square$

Now we provide an example of a non-convex polynomial optimization for which strong duality holds with a SDP dual program, verifying statement (i) of Theorem 3.1.

Example 3.2 (Non-convex polynomial problem with strong duality). Consider the problem $(\mathbf{P_0})$:

$$\begin{array}{lll} (\mathbf{P_0}) & \text{Minimize} & x_1^8 + x_1 x_2 \\ & \text{s.t.} & x_1^2 + 4x_1 + 3 + x_2^2 \leqq 0, \ x_2 \leqq 0. \end{array}$$

Let $f(x_1, x_2) = x_1^8 + x_1 x_2$, $g_1(x_1, x_2) = x_1^2 + 4x_1 + 3 + x_2^2$ and $g_2(x_1, x_2) = x_2$. As shown in Example 2.4, f is (-1)-SOS-convex. Clearly, g_1 is 1-SOS-convex and g_2 is SOS-convex. The feasible set of $(\mathbf{P_0})$ is $K := \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 + 2)^2 + x_2^2 - 1 \leqq 0, \ x_2 \leqq 0\}$. Note that, for any $x \in K$, $x_1^8 + x_1 x_2 \geqq x_1^8 \geqq 1 = f(-1, 0)$. Thus, $(\bar{x}_1, \bar{x}_2) = (-1, 0)$ is a minimizer of $(\mathbf{P_0})$.

Let $\bar{\lambda}_1 = 4$ and $\bar{\lambda}_2 = 1$, $\rho_0 = -1$, $\rho_1 = 1$ and $\rho_2 = 0$. Then $\nabla f(\bar{x}_1, \bar{x}_2) + \bar{\lambda}_1 \nabla g_1(\bar{x}_1, \bar{x}_2) + \bar{\lambda}_2 \nabla g_2(\bar{x}_1, \bar{x}_2) = 0$, $\bar{\lambda}_1 g_1(\bar{x}_1, \bar{x}_2) = 0$ and $\bar{\lambda}_2 g_2(\bar{x}_1, \bar{x}_2) = 0$, that is, the Kuhn-Tucker condition holds at $(\bar{x}_1, \bar{x}_2)$ with the multipliers $\bar{\lambda}_1 = 4$ and $\bar{\lambda}_2 = 1$, and $\rho_0 + \bar{\lambda}_1 \rho_1 + \bar{\lambda}_2 \rho_2 = 3 \geqq 0$. It is easy to check that $f(\bar{x}_1, \bar{x}_2) \geqq \sup_{\mu \in \mathbb{R}, \lambda_1 \geqq 0, \lambda_2 \geqq 0} \{\mu : \ f + \lambda_1 g_1 + \lambda_2 g_2 - \mu \in \Sigma_8^2\}$. Moreover, we see that

$$\begin{aligned} & f(x_1, x_2) + \bar{\lambda}_1 g_1(x_1, x_2) + \bar{\lambda}_2 g_2(x_1, x_2) - f(\bar{x}_1, \bar{x}_2) \\ &= x_1^8 + x_1 x_2 + 4x_1^2 + 4x_2^2 + 16x_1 + x_2 + 11 \\ &= (x_1^4 - 1)^2 + 2(x_1^2 - 1)^2 + 5(x_1 + 1)^2 + (x_1 + x_2 + 1)^2 \\ &\qquad + (x_1 - x_2 + 1)^2 + \left(x_1 + \frac{1}{2}x_2 + 1\right)^2 + \frac{7}{4}x_2^2. \end{aligned}$$

So, $f(x_1,x_2) + \bar{\lambda}_1 g_1(x_1,x_2) + \bar{\lambda}_2 g_2(x_1,x_2) - f(\bar{x}_1,\bar{x}_2) \in \Sigma^2_8$. Hence $\max_{\mu\in\mathbb{R},\lambda_1\geqq 0,\lambda_2\geqq 0}\{\mu : f + \lambda_1 g_1 + \lambda_2 g_2 - \mu \in \Sigma^2_8\} \geqq f(\bar{x}_1,\bar{x}_2)$. So, $f(\bar{x}_1,\bar{x}_2) = \max_{\mu\in\mathbb{R},\lambda_1\geqq 0,\lambda_2\geqq 0}\{\mu : f + \lambda_1 g_1 + \lambda_2 g_2 - \mu \in \Sigma^2_8\}$ and $(\bar{\lambda}_1,\bar{\lambda}_2,\mu) = (4,1,1)$ is a maximizer of the dual problem of $(\mathbf{P_0})$. Thus the statement (i) of Theorem 3.1 holds.

Minimax Polynomial Optimization

Consider now the following minimax programming problem

$$\begin{array}{lll}(\mathbf{P_1}) & \inf\limits_{x\in\mathbb{R}^n} & \max\limits_{j\in\{1,\dots,r\}} p_j(x) \\ & \text{s.t.} & g_i(x) \leqq 0,\ i=1,\dots,m,\end{array}$$

where p_j and g_i are polynomials on $\mathbb{R}^n$ for all $j=1,\dots,r$ and $i=1,\dots,m$. Let d be the smallest even number such that $d \geqq \max\{\max_{j\in\{1,\dots,r\}} \deg p_j, \max_{i\in\{1,\dots,m\}} \deg g_i\}$. Let $K = \{x\in\mathbb{R}^n \mid g_i(x)\leqq 0,\ i=1,\dots,m\} \neq \emptyset$.

Notice $(\mathbf{P_1})$ is equivalent to the following optimization problem

$$\begin{array}{lll}(\tilde{\mathbf{P}}_\mathbf{1}) & \inf\limits_{x\in\mathbb{R}^n,t\in\mathbb{R}} & t \\ & \text{s.t.} & p_j(x) \leqq t,\ j=1,\dots,r, \\ & & g_i(x)\leqq 0,\ i=1,\dots,m.\end{array}$$

Let x^* be a local minimizer of $(\mathbf{P_1})$, then (x^*,t^*) is a local minimizer of $(\tilde{\mathbf{P}}_\mathbf{1})$, where $t^* = \max_{j\in\{1,\dots,r\}} p_j(x^*)$. If a suitable constraint qualification for $(\tilde{\mathbf{P}}_\mathbf{1})$ holds such as the Mangasarian-Fromovitz constraint qualification for $(\tilde{\mathbf{P}}_\mathbf{1})$, that is, there exist $d_1\in\mathbb{R}^n$ and $d_2\in\mathbb{R}$ such that $\nabla p_j(x^*)^T d_1 < d_2$ $(j\in I(x^*,t^*))$ and $\nabla g_i(x^*)^T d_1 < 0$ $(i \in I(x^*,t^*))$, where $I(x^*,t^*) = \{j : p_j(x^*) - t^* = 0\} \cup \{i : g_i(x^*) = 0\}$, then the following Kuhn-Tucker condition hold: there exist $\bar{\delta}_j \geqq 0$, $j=1,\dots,r$ and $\bar{\lambda}_i \geqq 0$, $i=1,\dots,m$ such that

$$\sum_{j=1}^r \bar{\delta}_j \nabla p_j(x^*) + \sum_{i=1}^m \bar{\lambda}_i \nabla g_i(x^*) = 0,\quad \bar{\lambda}_i g_i(x^*) = 0,\ i=1,\dots,m, \tag{7}$$

for some $\bar{\delta}\in\Delta$ with $\bar{\delta}_j = 0$ whenever $p_j(x^*) \neq \max_{j\in\{1,\dots,r\}} p_j(x^*)$, and $\bar{\lambda}\in\mathbb{R}^m_+$, where

$$\Delta = \left\{\delta\in\mathbb{R}^r \mid \delta_i \geqq 0,\ i=1,\dots,r,\ \sum_{i=1}^r \delta_i = 1\right\}.$$

We associate with $(\mathbf{P_1})$ the following dual problem:

$$(\mathbf{D_1}) \sup_{\substack{\delta\in\Delta,\lambda\in\mathbb{R}^m_+,\\ \mu\in\mathbb{R}}} \left\{\mu \ :\ \sum_{j=1}^r \delta_j p_j + \sum_{i=1}^m \lambda_i g_i - \mu \in \Sigma^2_d\right\}.$$

Consider the minimax quadratic polynomial optimization problem **(QP)**:

$$(\mathbf{QP}) \quad \inf_{x\in\mathbb{R}^n} \max_{j\in\{1,\ldots,r\}} \{x^T A_j x + a_j^T x + \beta_j\}$$
$$\text{s.t.} \quad x^T Q_i x + c_i^T x + \alpha_i \leqq 0, \; i = 1,\ldots,m,$$

where $A_j, j = 1,2,\ldots,r$ and Q_i, $i = 1,\ldots,m$ are $(n\times n)$ symmetric matrices, $a_j, j = 1,2,\ldots,r$ and c_i, $i = 1,\ldots,m$ are vectors and $\beta_j, j = 1,2,\ldots,r$ and α_i, $i = 1,\ldots,m$ are scalars. The corresponding SDP dual problem of **(QP)** becomes:

$$(\mathbf{QD}) \sup_{\delta\in\Delta,\lambda\in\mathbb{R}^m_+,\mu\in\mathbb{R}}$$
$$\left\{\mu : \begin{pmatrix} 2(\sum_{j=1}^r \delta_j A_j + \sum_{i=1}^m \lambda_i Q_i) & \sum_{j=1}^r \delta_j a_j + \sum_{i=1}^m \lambda_i c_i \\ (\sum_{j=1}^r \delta_j a_j + \sum_{i=1}^m \lambda_i c_i)^T & 2(\sum_{j=1}^r \delta_j \beta_j + \sum_{i=1}^m \lambda_i \alpha_i - \mu) \end{pmatrix} \succeq 0\right\}.$$

Theorem 3.3 (Minimax problems and strong duality). *For the problem* $(\mathbf{P_1})$*, let* $p_j, j = 1,2,\ldots,r$ *and* $g_i, i = 1,2,\ldots,m$ *be polynomials. Let argmin*$(\mathbf{P_1}) \neq \emptyset$*. Suppose that the Kuhn-Tucker condition* (7) *holds at a local minimizer* x^* *of* $(\mathbf{P_1})$ *with multipliers* $\bar\delta \in \Delta$ *and* $\bar\lambda \in \mathbb{R}^m_+$*. Then, the following statements hold:*

(*i*) *If* p_j *and* g_i *are respectively* $\overline{\rho}_j$*-SOS-convex and* $\underline{\rho}_i$*-SOS-convex for* $j = 1,2,\ldots,r$ *and* $i = 1,2,\ldots,m$ *such that* $\sum_{j=1}^r \bar\delta_j \overline{\rho}_j + \sum_{i=1}^m \bar\lambda_i \underline{\rho}_i \geqq 0$*, then* $\min(\mathbf{P_1}) = \max(\mathbf{D_1})$ *and* $(\bar\delta, \bar\lambda, \max_{j\in\{1,\ldots,r\}} p_j(x^*))$ *is a maximizer of* $(\mathbf{D_1})$*;*

(*ii*) *If* $p_j(x) = x^T A_j x + a_j^T x + \beta_j$*, for each* $j = 1,2,\ldots,r$ *and* $g_i(x) = x^T Q_i x + c_i^T x + \alpha_i$ *with* $\sum_{j=1}^r \bar\delta_j A_j + \sum_{i=1}^m \bar\lambda_i Q_i \succeq 0$*, then* $\min(\mathbf{QP}) = \max(\mathbf{QD})$ *and* $(\bar\delta, \bar\lambda, \bar q)$ *is a maximizer of* $(\mathbf{QD})$*, where* $\bar q := \max_{j\in\{1,\ldots,r\}} \bar x^T A_j \bar x + a_j^T \bar x + \beta_j$.

Proof. (*i*) Let x^* be a minimizer of $(\mathbf{P_1})$ at which Kuhn-Tucker condition (7) holds with $\bar\delta \in \Delta$ and $\bar\lambda \in \mathbb{R}^m_+$ such that $\bar\lambda_i g_i(\bar x) = 0$, $i = 1,\ldots,m$ and $\bar\delta_j = 0$ whenever $p_j(x^*) \neq \max_{j\in\{1,\ldots,r\}} p_j(x^*)$. Let $t^* = \max_{j\in\{1,\ldots,r\}} p_j(x^*)$ and $h(x) = \sum_{j=1}^r \bar\delta_j p_j(x) + \sum_{i=1}^m \bar\lambda_i g_i(x) - t^*$, $x \in \mathbb{R}^n$. Then h is SOS-convex, $\nabla h(x^*) = 0$ and $h(x^*) = 0$. As in the proof of Theorem 3.1, $\sum_{j=1}^r \bar\delta_j p_j + \sum_{i=1}^m g_i - t^* \in \Sigma^2_d$ and hence $(\bar\delta, \bar\lambda, t^*)$ is a feasible solution of $(\mathbf{D_1})$. Thus $\max(\mathbf{D_1}) \geqq t^*$. As the opposite inequality, $t^* \geqq \max(\mathbf{D_1})$, always holds, we get that $\min(\mathbf{P_1}) = \max(\mathbf{D_1})$ and $(\bar\delta, \bar\lambda, t^*)$ is a maximizer of $(\mathbf{D_1})$.

(*ii*) The conclusion follows using the same line of arguments as in (*i*) by noting the fact that $\sum_{j=1}^r \bar\delta_j p_j(x) + \sum_{i=1}^m \bar\lambda_i g_i(x)$ is SOS-convex if and only if $\sum_{j=1}^r \bar\delta_j A_j + \sum_{i=1}^m \bar\lambda_i Q_i \succeq 0$. □

We now show that the existence of Lagrange multipliers, satisfying the Kuhn-Tucker condition at each minimizer of **(QP)** together with an eigenvalue condition, is necessary for strong duality of **(QP)**.

Theorem 3.4. *Minimax QPs: necessary and sufficient condition For the problem* **(QP)***, let* A_j *and* Q_i *be* $(n \times n)$ *symmetric matrices for* $j = 1, \ldots, r$ *and* $i = 1, \ldots, m$. *Let argmim*(**QP**) *be non-empty. Then,* $\min(\mathbf{QP}) = \max(\mathbf{QD})$ *if and only if at each minimizer of* **(QP)** *the Kuhn-Tucker condition* (7) *holds with* $\bar{\delta} \in \Delta$ *and* $\bar{\lambda} \in \mathbb{R}^m_+$ *and* $\sum_{j=1}^r \bar{\delta}_j A_j + \sum_{i=1}^m \bar{\lambda}_i Q_i \succeq 0$.

Proof. Assume that $\min(\mathbf{QP}) = \max(\mathbf{QD})$. Let $\bar{x}$ be a minimizer of **(QP)** with $\bar{q} := \max_{j\in\{1,\ldots,r\}} f_j(\bar{x})$, where $f_j(x) = x^T A_j x + a_j^T x + \beta_j$. Then,

$$\begin{aligned}\bar{q} &= \max(\mathbf{QD}) \\ &= \max_{\substack{\delta\in\Delta,\lambda\in\mathbb{R}^m_+,\\ \mu\in\mathbb{R}}} \left\{ \mu : \sum_{j=1}^p \delta_j f_j(x) + \sum_{i=1}^m \lambda_i (x^T Q_i x + c_i^T x + \alpha_i) - \mu \geqq 0, \forall x \in \mathbb{R}^n \right\}.\end{aligned}$$

As a non-negative quadratic function is sum of squares, it follows that there exist $\bar{\delta} \in \Delta$, $\bar{\lambda} \in \mathbb{R}^m_+$ and $\sigma \in \Sigma_2^2$ such that, for each $x \in \mathbb{R}^n$,

$$\sum_{j=1}^r \bar{\delta}_j f_j(x) + \sum_{i=1}^m \bar{\lambda}_i (x^T Q_i x + c_i^T x + \alpha_i) - \bar{q} = \sigma(x). \tag{8}$$

Letting $x = \bar{x}$ in (8), we get

$$\sum_{j=1}^r \bar{\delta}_j f_j(\bar{x}) - \bar{q} = \sigma(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i (\bar{x}^T Q_i \bar{x} + c_i^T \bar{x} + \alpha_i) \geqq \sigma(\bar{x}) \geqq 0.$$

On the other hand, by definition, $\bar{q} \geqq \sum_{j=1}^r \bar{\delta}_j f_j(\bar{x})$. So, $\bar{q} = \sum_{j=1}^r \bar{\delta}_j f_j(\bar{x})$ and hence,

$$\sum_{i=1}^m \bar{\lambda}_i (x^T Q_i x + c_i^T x + \alpha_i) = \sigma(\bar{x}) \geqq 0.$$

As $\bar{x}$ is feasible for **(QP)** and $\bar{\lambda} \in \mathbb{R}^m_+$, $\sum_{i=1}^m \bar{\lambda}_i (\bar{x}^T Q_i \bar{x} + c_i^T \bar{x} + \alpha_i) = 0$ and $\sigma(\bar{x}) = 0$. This shows us that $\bar{x}$ is a minimizer of σ since for each $x \in \mathbb{R}^n$, $\sigma(x) \geq 0 = \sigma(\bar{x})$. Then, by the first-order necessary optimality condition, we obtain that $\nabla\sigma(\bar{x}) = 0$. Thus, $\sum_{j=1}^r \bar{\delta}_j \nabla f_j(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i (2Q_i \bar{x} + c_i) = 0$ and so the Kuhn-Tucker condition (7) is satisfied at $\bar{x}$. Note that $\bar{\delta}_j = 0$ whenever $\bar{q} \neq f_j(\bar{x})$ because $\sum_{J=1}^r \bar{\delta}_j (f_j(\bar{x}) - \bar{q}) = 0$. Moreover, by the second-order necessary optimality condition, $\nabla^2\sigma(\bar{x}) \succeq 0$, which is equivalent to the condition that $\sum_{j=1}^r \bar{\delta}_j A_j + \sum_{i=1}^m \bar{\lambda}_i Q_i \succeq 0$.

The converse implication follows from the same line of arguments as in Theorem 3.1 (i) as $\sum_{j=1}^r \bar{\delta}_j A_j + \sum_{i=1}^m \bar{\lambda}_i Q_i \succeq 0$ if and only if $\sum_{j=1}^r \bar{\delta}_j (x^T A_j x + a_j^T x + \beta_j) + \sum_{i=1}^m \bar{\lambda}_i (x^T Q_i x + c_i^T x + \alpha_i)$ is an SOS-convex polynomial. $\square$

Now consider again the quadratic polynomial optimization problem:

$$\begin{array}{lll} (\mathbf{QP_1}) & \text{Minimize} & x^T Q_0 x + c_0^T x + \alpha_0 \\ & \text{subject to} & x^T Q_i x + c_i^T x + \alpha_i \leqq 0, \ i = 1, \ldots, m, \end{array}$$

and its dual problem

$$(\mathbf{QD_1}) \sup_{\mu\in\mathbb{R},\lambda_i\geqq 0}\left\{\mu : \begin{pmatrix} 2(Q_0+\sum_{i=1}^m \lambda_i Q_i) & c_0+\sum_{i=1}^m \lambda_i c_i \\ (c_0+\sum_{i=1}^m \lambda_i c_i)^T & 2(\alpha_0+\sum_{i=1}^m \lambda_i \alpha_i - \mu) \end{pmatrix} \succeq 0\right\}.$$

As a consequence of Theorem 3.4, we obtain the following characterization of strong duality for quadratic programs. A related result for the model $(\mathbf{QP_1})$ was given in [3] under the Slater condition.

Corollary 3.5. *For the problem* $(\mathbf{QP_1})$*, let* Q_i *be* $(n\times n)$ *symmetric matrices for* $i=0,\dots,m$*. Let argmin*$(\mathbf{QP_1})$ *be non-empty. Then,* $\min(\mathbf{QP_1}) = \max(\mathbf{QD_1})$ *if and only if at each minimizer of* $(\mathbf{QP_1})$ *the Kuhn-Tucker condition holds with* $\bar\lambda_i \geq 0, i=1,2,\dots,m$ *and* $Q_0+\sum_{i=1}^m \bar\lambda_i Q_i \succeq 0$*.*

Proof. The conclusion follows from Theorem 3.4 by letting $r=1$. □

4. Strong Duality for Classes of Nonlinear Programs

In this Section we derive SDP dual problems for classes of nonlinear programming problems involving ρ-SOS-convex polynomials. We begin with fractional programming problems.

Fractional Programs

Now, consider the following fractional programming problem

$$(\mathbf{FP}) \quad \inf_{x\in\mathbb{R}^n} \frac{p(x)}{q(x)} \quad \text{s.t.} \quad g_i(x)\leqq 0,\ i=1,\dots,m,$$

where p, q and g_i, for $i=1,\dots,m$, are real polynomials on $\mathbb{R}^n$, and $p(x)\geqq 0$, and $q(x)>0$ over the feasible set. Let $K:=\{x\in\mathbb{R}^n \mid g_i(x)\leqq 0,\ i=1,\dots,m\}\neq\emptyset$. Let $x^*\in K$ be an optimal solution of **(FP)** and let $\mu^*=\frac{p(x^*)}{q(x^*)}$. Then x^* is an optimal solution of the following optimization problem:

$$\inf_{x\in\mathbb{R}^n} \{p(x)-\mu^* q(x)\} \quad \text{s.t.} \quad g_i(x)\leqq 0,\ i=1,\dots,m.$$

Thus if the Mangasarian-Fromovitz constraint qualification holds for **(FP)** holds, that is, there exists $d\in\mathbb{R}^n$ such that $\nabla g_i(x^*)^T d<0$ $(i\in I(x^*))$, where $I(x^*)=\{i : g_i(x^*)=0\}$, then

$$0=\nabla p(x^*)+\sum_{i=1}^m \bar\lambda_i\nabla g_i(x^*)-\mu^*\nabla q(x^*) \tag{9}$$

for some $\bar\lambda\in\mathbb{R}^m_+$ with $\bar\lambda_i g_i(x^*)=0$ for all $i=1,\dots,m$.

We associate with **(FP)** the following SDP dual problem

$$(\mathbf{FD}) \sup_{\lambda\in\mathbb{R}^m_+,\ \mu\in\mathbb{R}} \left\{ \mu :\ p + \sum_{i=1}^{m} \lambda_i g_i - \mu q \in \Sigma^2_d \right\},$$

where d is the smallest even number such that $d \geqq \max\{\deg p, \deg g, \max_{i\in\{1,\dots,m\}} \deg g_i\}$.

Theorem 4.1. *Let p, q and g_i be $\overline{\rho}_0$-SOS-convex, $\underline{\rho}_0$-SOS-convex and ρ_i-SOS-convex polynomials respectively for each $i = 1,\dots,m$. Assume that $K := \{x \in \mathbb{R}^n : g_i(x) \leqq 0, i = 1,\dots,m\} \neq \emptyset$. Let $x^* \in K$ at which the Kuhn-Tucker condition* (9) *holds with multiplier $\bar{\lambda} \in \mathbb{R}^m_+$ and $\mu^* = \frac{p(x^*)}{q(x^*)}$. If $\overline{\rho}_0 + \sum_{i=1}^m \bar{\lambda}_i\rho_i - \mu^*\underline{\rho}_0 \geqq 0$, then x^* is a minimizer of* $(\mathbf{FP})$ *and* $\frac{p(x^*)}{q(x^*)} = \max(\mathbf{FD})$.

Proof. Let $x^* \in K$ be an optimal solution of **(FP)** at which the Kuhn-Tucker condition (9) holds with multiplier $\bar{\lambda} \in \mathbb{R}^m_+$ and $\mu^* = \frac{p(x^*)}{q(x^*)}$. Let $h(x) = p(x) + \sum_{i=1}^p \bar{\lambda}_i g_i(x) - \mu^* q(x)$, $x \in \mathbb{R}^n$. Then $\nabla h(x^*) = 0$ and $h(x^*) = 0$. Since h is $(\overline{\rho}_0 + \sum_{i=1}^m \bar{\lambda}_i\rho_i - \mu^*\underline{\rho}_0)$-SOS-convex and $\overline{\rho}_0 + \sum_{i=1}^m \bar{\lambda}_i\rho_i - \mu^*\underline{\rho}_0 \geqq 0$, h is a SOS-convex polynomial. Hence, it follows from Lemma 2.5 that $h \in \Sigma^2_d$. So, $(\bar{\lambda}, \mu^*)$ is a feasible point of $(\mathbf{FD})$ and hence $\mu^* \leqq \sup(\mathbf{FD})$. Hence, we conclude that x^* is a minimizer of $(\mathbf{FP})$ and $\min(\mathbf{FP}) = \max(\mathbf{FD})$ as the opposite inequality always holds. □

Robust Optimization Problems

In this subsection, we show how our Theorem 3.1 can be used to derive SDP dual problems for optimization problems **(P)** in the face of data uncertainty. The problem **(P)** in the face of data uncertainty in the constraints can be captured by the following nonlinear programming problem:

$$\begin{aligned} (\mathbf{UP}) \quad \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x, v_i) \leqq 0,\ i = 1,\dots,m, \end{aligned}$$

where v_i is an uncertain parameter and $v_i \in \mathcal{V}_i$ for some convex compact set $\mathcal{V}_i$ in $\mathbb{R}^q$ and $g_i : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$ is a function such that for each fixed $v_i \in \mathcal{V}_i$, $g_i(\cdot, v_i)$ is a polynomial.

We only consider the simple case of finite uncertainty, where $\mathcal{V}_i = \{v_i^1, \dots, v_i^{s_i}\}$ for any $i \in \{1,\dots,m\}$. Then, the robust counterpart of **(UP)** takes the form

$$(\mathbf{RP}) \quad \min_{x\in\mathbb{R}^n} \{f(x) : g_i^j(x, v_i^j) \leq 0,\ j = 1,\dots,s_i,\ i = 1,\dots,m\},$$

where the uncertain constraints are enforced for every possible value of the parameters within their prescribed uncertainty sets $\mathcal{V}_i$, $i = 1,\dots,m$.

The SDP dual problem for (**UP**) is given by

$$(\mathbf{DP}) \qquad \sup_{\lambda_i^j \geq 0, \mu \in \mathbb{R}} \left\{ \mu : f + \sum_{i=1}^{m} \sum_{j=1}^{s_i} \lambda_i^j g_i^j(\cdot, v_i^j) - \mu \in \Sigma_d^2 \right\},$$

where d is the smallest even number such that $d \geq k_0 = \max_{1 \leq i \leq m, 1 \leq j \leq s_i}\{\deg f, \deg g_i(\cdot, v_i^j)\}$.

Theorem 4.2. *Let f be ρ_0-SOS-convex polynomial and for each $v_i^j \in \mathcal{V}_i$, $g_i^j(\cdot, v_i^j)$ be ρ_i^j-SOS-convex polynomial. Let $\mathcal{F} := \{x \in \mathbb{R}^n \mid g_i^j(\cdot, v_i^j) \leqq 0,\ i = 1, \cdot, m,\ j = 1, 2, \ldots, s_i\} \neq \emptyset$. Assume that $x^* \in \mathcal{F}$ at which the Kuhn-Tucker condition holds with multipliers $\bar{\lambda}_i^j$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, s_i$. If $\rho_0 + \sum_{i=1}^m \sum_{i=1}^{s_i} \bar{\lambda}_i \rho_i^j \geqq 0$, then* $\min(\mathbf{UP}) = \max(\mathbf{DP})$.

Proof. The conclusion follows using the same line of arguments as in Theorem 3.1. □

Related duality results for robust SOS-convex problems under various classes of uncertainty sets, such as the ellipsoidal uncertainty sets and norm uncertainty sets, can be found in [18, 19, 20].

5. A Class of Extended Trust-Region Problems

Although the Lagrangian conditions for strong duality in Sections 3 are easy to check for quadratic optimization problems, it would be useful, in practice, to know whether or not these conditions can be guaranteed for some classes of quadratic problems under suitable conditions. To examine this, let us consider the extended trust-region problem with linear inequality constraints

$$(\mathbf{QT}) \ \min_{x \in \mathbb{R}^n} \{x^T A x + a^T x \ : \ \|x - x_0\|^2 \leq \alpha, \ b_i^T x \leq \beta_i, \ i = 2, \ldots, m\},$$

where A is a symmetric $(n \times n)$ matrix, $a, b_i, x_0 \in \mathbb{R}^n$ and $\alpha, \beta_i \in \mathbb{R}, \alpha > 0$, $i = 2, \ldots, m$. This extended trust-region model was examined recently in [16] and they arise from the application of the trust region method for solving constrained optimization problems [11], such as nonlinear programming problems with linear inequality constraints, nonlinear optimization problems with discrete variables [9] and robust optimization problems [7] under matrix norm or polyhedral uncertainty.

Our SDP relaxation problem for **(QT)** becomes

$$(\mathbf{DQT}) \ \sup_{\mu \in \mathbb{R}, \lambda_i \geqq 0}$$
$$\left\{ \mu : \begin{pmatrix} 2A + 2\lambda_1 I_n & a - 2\lambda_1 x_0 + \sum_{i=2}^m \lambda_i b_i \\ (a - 2\lambda_1 x_0 + \sum_{i=2}^m \lambda_i b_i)^T & 2(\lambda_1 \|x_0\|^2 - \lambda_1 \alpha - \sum_{i=2}^m \lambda_i \beta_i - \mu) \end{pmatrix} \succeq 0 \right\}.$$

Recently, the problem **(QT)** was examined in [16] where a sufficient condition was given for the equality between the values of **(QT)** and a related SDP relaxation problem that was formulated as a rank one relaxation problem. In the following we obtain a necessary and sufficient condition for the equality between the values **(QT)** and its dual problem **(DQT)** with attainment of their optimal values.

Let $f_0(x) = x^TAx+a^Tx$, $g_1(x) = \|x-x_0\|^2-\alpha$ and $g_i(x) = b_i^Tx-\beta_i$, $i = 2, \dots, m$. Define

$$U(f_0, g_1, \dots, g_m) := \{(f_0(x), g_1(x), \dots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}_+^{m+1}.$$

As seen in [15, 16], the epigraphical set $U(f_0, g_1, \dots, g_m)$ plays a key role in the study of the model problem **(QT)**. Clearly, if $A \succeq 0$, then the set $U(f_0, g_1, \dots, g_m)$ is convex.

The set $U(f_0, g_1, \dots, g_m)$ has been shown to be a convex set for some (not necessarily convex) quadratic functions f. For instance, if $g_i = 0$, for $i = 2 \dots, m$, then $U(f_0, g_1)$ is always a convex set (see [23]). Note also that if the joint-range set $\{(f_0(x), g_1(x), \dots, g_m(x)) : x \in \mathbb{R}^n\}$ is convex then $U(f_0, g_1, \dots, g_m)$ is also convex. For sufficient conditions for convexity of the joint-range set involving quadratic functions, see [5, 6]. Sufficient conditions using Z-matrices for convexity of the epigraphical set of the form $U(f_1, f_2, \dots, f_m)$, involving quadratic functions $f_i, i = 1, 2, \dots, m,$, may also be found in [15].

More recently, the following sufficient condition for convexity of $U(f_0, g_1, \dots, g_m)$ was given under the dimension condition that

$$\dim \operatorname{Ker}(A - \gamma_{\min}(A)I_n) \geqq s + 1,$$

where $\dim \operatorname{span}\{b_2, \dots, b_m\} = s$. Recall that the kernel of A is denoted by $\operatorname{Ker}(A) := \{d \in \mathbb{R}^n : Ad = 0\}$. The dimension of a subspace, L, is denoted by $\dim L$.

Lemma 5.1 (Theorem 2.1 in [16]). *Let $f_0(x) = x^TAx + a^Tx + r$, $g_1(x) = \|x - x_0\|^2 - \alpha$ and $g_i(x) = b_i^Tx - \beta_i$, $i = 2, \dots, m$. Let A be a symmetric $(n \times n)$ matrix, a, x_0, $b_i \in \mathbb{R}^n$ and r, α, $\beta_i \in \mathbb{R}$. Let $\dim \operatorname{span}\{b_2, \dots, b_m\} = s$. Assume that*

$$\dim \operatorname{Ker}(A - \gamma_{\min}(A)I_n) \geqq s + 1. \tag{10}$$

Then $U(f_0, g_1, \dots, g_m) := \{(f_0(x), g_1(x), \dots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}_+^{m+1}$ is a convex set in $\mathbb{R}_+^{m+1}$.

The example below shows that the dimension condition may fail for a non-convex quadratic program while the set $U(f_0, g_1, \dots, g_m)$ is convex.

Example 5.2. Consider the extended trust-region problem $(\mathbf{QT_0})$:

$$\begin{array}{lll} (\mathbf{QT_0}) & \text{Minimize} & -x_1^2 - x_2 \\ & \text{s.t.} & (x_1+3)^2 + x_2^2 \leqq 1, \\ & & x_2 \leqq 0. \end{array}$$

Let $f(x_1, x_2) = -x_1^2 - x_2$, $g_1(x_1, x_2) = (x_1 + 3)^2 + x_2^2 - 1$ and $g_2(x_1, x_2) = x_2$. Then

$$f_0(x_1, x_2) = (x_1, x_2)\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (0, -1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

$$g_1(x_1, x_2) = \|(x_1, x_2) - (-3, 0)\|^2 - 1 \quad \text{and} \quad g_2(x_1, x_2) = (0, 1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Let $A = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $a = \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)$, $x_0 = \left(\begin{smallmatrix} -3 \\ 0 \end{smallmatrix}\right)$, $\alpha = 1$, $b_1 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ and $\beta_1 = 0$.
Since $\dim \operatorname{span}\{b_1\} = 1$ and $\dim\operatorname{Ker}(A - \gamma_{\min}(A)I_2) = 1$, $\dim\operatorname{Ker}(A - \gamma_{\min}(A)I_2) < 2$, and hence the dimension condition fails.

However, it is worth noting that the related convex minimization problem

$$\min_{(x_1,x_2)\in\mathbb{R}^2}\{f(x_1, x_2) - (-1)\|(x_1, x_2) - (-3, 0)\|^2 : g_1(x_1, x_2) \leqq r,\ g_2(x_1, x_2) \leqq s_2\}$$

attains its minimum on the ball $(\bar{x}_1 + 3)^2 + \bar{x}_2^2 = 1 + r$ with $\bar{x}_2 \leqq s_2$, for each $(r, s) \in D := \{(u, v) \in \mathbb{R}^2 : g_1(x_1, x_2) \leqq u,\ g_2(x_1, x_2) \leqq v\}$ as shown in Example 5.5 below.

This prompts us to give the following more general sufficient condition for convexity of the set $U(f_0, g_1, \ldots, g_m)$.

Lemma 5.3 (Epigraphical convexity in extended trust-region problems). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a ρ_0-SOS-convex polynomial, $x_0 \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ and $g_1(x) = \|x - x_0\|^2 - \alpha$. Let $g_i(x)$, $i = 2, \ldots, m$ be convex functions. Assume that $D := \{(r, s_2, \ldots, s_m) \mid g_1(x) \leqq r,\ g_i(x) \leqq s_i,\ i = 2, \ldots, m,\ for\, some\, x \in \mathbb{R}^n\} \neq \emptyset$. If for each $(r, s_2, \ldots, s_m) \in D$, the (SOS-) convex minimization problem*

$$\min_{x\in\mathbb{R}^n}\{f(x) - \rho_0\|x - x_0\|^2 : g_1(x) \leqq r,\ g_i(x) \leqq s_i,\ i = 2, \ldots, m\}$$

attains its minimum at some $\bar{x} \in \mathbb{R}^n$ with $g_1(\bar{x}) = r$ and $g_i(\bar{x}) \leqq s_i$, $i = 2, \ldots, m$, then $U(f, g_1, \ldots, g_m) := \{(f(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}^{m+1}_+$ is a convex set in $\mathbb{R}^{m+1}$.

Proof. If $\rho_0 \geqq 0$, then f is convex and hence $U(f, g_1, \ldots, g_m)$ is convex. Assume that $\rho_0 < 0$. Define an optimal value function $h : \mathbb{R}^{m+1} \to \mathbb{R} \cup \{+\infty\}$ by

$$h(r, s_2, \ldots, s_m) = \begin{cases} \min\limits_{x\in\mathbb{R}^n}\{f(x) : \|x - x_0\|^2 - \alpha \leqq r, & \\ \qquad\qquad g_i(x) \leqq s_i,\ i = 2, \ldots, m\}, & (r, s_2, \ldots, s_m) \in D \\ +\infty, & \text{otherwise.} \end{cases}$$

We can easily check that the set D is convex and that the set $U(f, g_1, \ldots, g_m)$ is the same as the epigraph of the function h. Let $(r, s_2, \ldots, s_m) \in D$. Then by assumption, there exists $\bar{x} \in \mathbb{R}^n$ such that $g_1(\bar{x}) = r$, $g_i(\bar{x}) \leqq s_i$, $i = 2, \dots, m$ and

$$\min_{x\in\mathbb{R}^n} \{f(x) - \rho_0\|x - x_0\|^2 : \|x - x_0\|^2 - \alpha \leqq r,\ g_i(x) \leqq s_i,\ i = 2, \dots, m\}$$
$$= f(\bar{x}) - \rho_0(\alpha + r).$$

Moreover, we have,

$$\begin{aligned} & h(r, s_2, \dots, s_m) \\ = & \min_{x\in\mathbb{R}^n} \{f(x) - \rho_0\|x - x_0\|^2 + \rho_0\|x - x_0\|^2 : \|x - x_0\|^2 - \alpha \leqq r, \\ & \quad g_i(x) \leqq s_i,\ i = 2, \dots, m\} \\ \geqq & \min_{x\in\mathbb{R}^n} \{f(x) - \rho_0\|x - x_0\|^2 + \rho_0(\alpha + r) : \|x - x_0\|^2 - \alpha \leqq r, \\ & \quad g_i(x) \leqq s_i,\ i = 2, \dots, m\} \\ = & f(\bar{x}) - \rho_0(\alpha + r) + \rho_0(\alpha + r) \\ = & f(\bar{x}) \\ \geqq & h(r, s_2, \dots, s_m). \end{aligned}$$

Hence $h(r, s_2, \dots, s_m) = \min_{x\in\mathbb{R}^n}\{f(x) - \rho_0\|x - x_0\|^2 + \rho_0(\alpha + r) : \|x - x_0\|^2 - \alpha \leqq r,\ g_i(x) \leqq s_i,\ i = 2, \dots, m\}$. For any $x \in \mathbb{R}^n$, $f(x) - \rho_0\|x - x_0\|^2 + \rho_0(\alpha + r) = f(x) - \rho_0\|x\|^2 - 2\rho_0 x_0^T x - \rho_0\|x_0\|^2 + \rho_0(\alpha + r)$. Since f is ρ_0-SOS-convex, $f - \rho_0\|\cdot\|^2$ is convex and hence $f(x) - \rho_0\|x - x_0\|^2 + \rho_0(\alpha + r)$ is also convex. Hence, the function h is convex (see Lemma 1 in [12]) and then the epigraph of h is convex in $\mathbb{R}^{m+1}$, and so $U(f, g_1, \dots, g_m)$ is convex in $\mathbb{R}^{m+1}$. □

It is worth noting that it is easy to show that if f is a quadratic function then the dimension condition (10) guarantees our hypothesis of Lemma 5.3.

We now establish a necessary and sufficient condition for strong duality of extended trust-region problems, extending recent work in [5, 16].

Theorem 5.4 (Extended trust-region problems: optimality and duality). *For the problem* **(QT)**, *A be a symmetric $(n \times n)$ matrix, $a, b_i, x_0 \in \mathbb{R}^n$ and $\alpha, \beta_i \in \mathbb{R}, \alpha > 0$, $i = 2, \dots, m$. Let $f(x) = x^T Ax + a^T x$, $g_1(x) = \|x - x_0\|^2 - \alpha$ and let $g_i(x) = b_i^T x - \beta_i$, $i = 2, \dots, m$. Assume that $U(f, g_1, \dots, g_m)$ is a convex set and that there exists $\overline{x} \in \mathbb{R}^n$ with $\|\overline{x} - x_0\|^2 < \alpha$ and $b_i^T \overline{x} < \beta_i$, $i = 1, \dots, m$. Then the following statements are equivalent:*

(i) *The feasible point x^* is a minimizer of* **(QT)**;

(ii) *The point x^* is a feasible point of* **(QT)** *such that there exists $(\lambda_1^*, \dots, \lambda_m^*) \in \mathbb{R}_+^m$ satisfying*

$$\begin{cases} 2Ax^* + a + 2\lambda_1^*(x^* - x_0) + \sum_{i=2}^m \lambda_i^* b_i = 0, \\ \lambda_1^*(\|x^* - x_0\|^2 - \alpha) = 0,\ \lambda_i(b_i^T x^* - \beta_i) = 0, i = 2, \dots, m, \\ A + \lambda_1^* I_n \succeq 0; \end{cases} \tag{11}$$

(*iii*) *The point* x^* *is a feasible point of* **(QT)**, $\min(\mathbf{QT}) = \max(\mathbf{DQT})$ *and* $(\lambda^*, f(x^*))$ *is a maximizer of* **(DQT)**.

Proof. $[(i) \Rightarrow (ii)]$ Assume that x^* is a global minimizer of **(P)**. Then, the following inequality system has no solution:

$$g_i(x) \leqq 0, \quad i = 1, \ldots, m, \ f(x) < f(x^*).$$

So, $(f(x^*), 0) \notin \operatorname{int} U(f, g_1, \ldots, g_m)$, where

$$\operatorname{int} U(f, g_1, \ldots, g_m) = \{(f(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \operatorname{int} \mathbb{R}^{m+1}_+$$

is a convex set, where $\operatorname{int} A$ is the interior of a set A in $\mathbb{R}^n$, as f and g_i are all continuous functions and

$$U(f, g_1, \ldots, g_m) := \{(f(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\} + \mathbb{R}^{m+1}_+$$

is a convex set, by assumption.

Now, by the convex separation theorem, there exists $(\mu_0^*, \mu_1^*, \ldots, \mu_m^*) \in \mathbb{R}^{m+1}_+ \backslash \{0\}$ such that, for all $x \in \mathbb{R}^n$,

$$\mu_0^*(f(x) - f(x^*)) + \sum_{i=1}^m \mu_i^* g_i(x) \geqq 0.$$

The strict feasibility assumption guarantees that $\mu_0^* \neq 0$. So, for each $x \in \mathbb{R}^n$,

$$f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) \geqq f(x^*).$$

where $\lambda_i^* = \frac{\mu_i^*}{\mu_0^*}$, $i = 1, \ldots, m$. Letting $x = x^*$, we see that $\sum_{i=1}^m \lambda_i^* g_i(x^*) \geq 0$. Since x^* is feasible for **(P)**, it follows that $\lambda_i^* g(x^*) = 0$, $i = 1, \ldots, m$. Let $h(x) := f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) - f(x^*)$. Then, x^* is a global minimizer of h, and so, $\nabla h(x^*) = 0$ and $\nabla^2 h(x^*) \succeq 0$. Now, $\nabla h(x^*) = 0$ give us that $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$. On the other hand, $\nabla^2 h(x^*) \succeq 0$ is equivalent to the condition that $A + \lambda_1^* I_n \succeq 0$. Hence (ii) holds.

$[(ii) \Rightarrow (iii)]$ This implication directly follows from statement (iii) of Theorem 3.1.

$[(iii) \Rightarrow (i)]$ As x^* is a feasible point of **(QT)**, it now follows from (iii) that $\min(\mathbf{QT}) = \max(\mathbf{QDT}) = f(x^*)$. Hence, x^* is a global minimizer of **(QT)**. □

Example 5.5 (Extended trust-region with strong duality). Consider again the extended trust-region problem, examined in Example 5.2

$$(\mathbf{QT_0}) \quad \begin{array}{ll} \text{Minimize} & -x_1^2 - x_2 \\ \text{s.t.} & (x_1+3)^2 + x_2^2 \leqq 1, \\ & x_2 \leqq 0. \end{array}$$

Let $f(x_1, x_2) = -x_1^2 - x_2$, $g_1(x_1, x_2) = (x_1 + 3)^2 + x_2^2 - 1$ and $g_2(x_1, x_2) = x_2$. Then

$$f(x_1, x_2) = (x_1, x_2) \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (0, -1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

$$g_1(x_1, x_2) = \|(x_1, x_2) - (-3, 0)\|^2 - 1 \quad \text{and} \quad g_2(x_1, x_2) = (0, 1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Let $A = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 0 \end{smallmatrix} \right)$, $a = \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix} \right)$, $x_0 = \left(\begin{smallmatrix} -3 \\ 0 \end{smallmatrix} \right)$, $\alpha = 1$, $b_1 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right)$ and $\beta_1 = 0$.

It is clear that f is (-1)-SOS-convex.

Now we show that $U(f, g_1, g_2)$ is a convex set. Let $D = \{(r, s_2) \in \mathbb{R}^2 : g_1(x_1, x_2) \leqq r,\ g_2(x_1, x_2) \leqq s_2\}$ and let $(r, s_2) \in D$ be any fixed vector. Consider the convex minimization problem:

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \{f(x_1, x_2) - (-1)\|(x_1, x_2) - (-3, 0)\|^2 : g_1(x_1, x_2) \leqq r,\ g_2(x_1, x_2) \leqq s_2\}.$$

Since its feasible set is compact, the minimization problem has its minimizer $(\bar{x}_1, \bar{x}_2) \in \mathbb{R}^2$. By Fritz-John necessary optimality theorem in [22], there exist $\lambda_0 \geqq 0$, $\lambda_1 \geqq 0$ and $\lambda_2 \geqq 0$, not all zero, such that

$$\lambda_0 \begin{pmatrix} 6 \\ 2\bar{x}_2 - 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2\bar{x}_1 + 6 \\ 2\bar{x}_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{12}$$

$$\lambda_1((\bar{x}_1 + 3)^2 + \bar{x}_2^2 - 1 - r) = 0, \tag{13}$$

$$\lambda_2(\bar{x}_2 - s_2) = 0, \tag{14}$$

$$(\bar{x}_1 + 3)^2 + \bar{x}_2^2 \leqq 1 + r, \tag{15}$$

$$\bar{x}_2 \leqq s_2. \tag{16}$$

If $(\bar{x}_1 + 3)^2 + \bar{x}_2^2 - 1 - r < 0$, from (13), $\lambda_1 = 0$ and so from (12), $\lambda_0 = 0$, and so from (12), $\lambda_2 = 0$, and consequently, $(\lambda_0, \lambda_1, \lambda_2) = (0, 0, 0)$. This is a contradiction. Hence $(\bar{x}_1 + 3)^2 + \bar{x}_2^2 = 1 + r$ and $\bar{x}_2 \leqq s_2$, that is , $g_1(\bar{x}_1, \bar{x}_2) = r$ and $g_2(\bar{x}_1, \bar{x}_2) = s_2$. Hence by Lemma 5.3, $U(f, g_1, g_2)$ is a convex set in $\mathbb{R}^3$.
For any feasible point (x_1, x_2) of $(\mathbf{QT_0})$, $f(x_1, x_2) \geqq -x_1^2 \geqq -16$, and hence $x^* := (x_1^*, x_2^*) = (-4, 0)$ is an optimal solution of $(\mathbf{QT_0})$. Let $\lambda_1^* = 4$ and $\lambda_2^* = 1$. Then

$$\begin{aligned} &2Ax^* + a + 2\lambda_1^*(x^* - x_0) + \lambda_2^* b_1 \\ &= 2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 8 \left(\begin{pmatrix} -4 \\ 0 \end{pmatrix} - \begin{pmatrix} -3 \\ 0 \end{pmatrix} \right) + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned}$$

$$\lambda_1^*(\|x^* - x_0\|^2 - \alpha) = 0, \qquad \lambda_2^* x_2^* = 0,$$

and

$$A + \lambda_1^* I_2 = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \succeq 0.$$

The SDP dual problem $(\mathbf{QT_0})$ becomes

$$(\mathbf{DQT_0}) \quad \sup_{\substack{\mu\in\mathbb{R}\\ \lambda_1\geqq 0,\lambda_2\geqq 0}} \left\{ \mu : \begin{pmatrix} 2A + 2\lambda_1 I_2 & a - 2\lambda_1 x_0 + \lambda_2 b_1 \\ (a - 2\lambda_1 x_0 + \lambda_2 b_1)^T & 2(\lambda_1 \|x_0\|^2 - \lambda_2 \beta_1 - \mu) \end{pmatrix} \succeq 0 \right\}.$$

Thus $(\mathbf{DQT_0})$ is

$$\sup_{\substack{\mu\in\mathbb{R}\\ \lambda_1\geqq 0,\lambda_2\geqq 0}} \left\{ \mu : \begin{pmatrix} 2\lambda_1 - 2 & 0 & 6\lambda_1 \\ 0 & 2\lambda_1 & \lambda_2 - 1 \\ 6\lambda_1 & \lambda_2 - 1 & 18\lambda_1 - 2\mu \end{pmatrix} \succeq 0 \right\}.$$

Since $\sup(\mathbf{DQT_0}) = \sup_{\substack{\mu\in\mathbb{R}\\ \lambda_1\geqq 0,\lambda_2\geqq 0}} \{\mu : f + \lambda_1 g_1 + \lambda_2 g_2 - \mu \in \Sigma_2^2\}$, $f(x_1^*, x_2^*) \geqq \sup(\mathbf{DQT_0})$. Let $\mu^* = f(x_1^*, x_2^*) = -16$. Then we can check that

$$\begin{pmatrix} 2\lambda_1^* - 2 & 0 & 6\lambda_1^* \\ 0 & 2\lambda_1^* & \lambda_2^* - 1 \\ 6\lambda_1^* & \lambda_2^* - 1 & 18\lambda_1^* - 2\mu^* \end{pmatrix} \succeq 0$$

and hence $(\lambda_1^*, \lambda_2^*, \mu^*)$ is a maximizer of $(\mathbf{DQT_0})$ and so, $\min(\mathbf{QT_0}) = \max(\mathbf{DQT_0})$. Hence, Theorem 5.4 holds.

6. Conclusion and Further Research

In this paper we used the generalized notion of SOS-convexity, called ρ-SOS-convexity, to obtain strong duality results for classes of non-convex polynomial optimization problems. As a result, we were able to provide necessary as well as sufficient conditions for strong duality of quadratic optimization problems. The approach in the present work suggests that our strong duality results may extend to minimax trust-region problems with linear inequalities under an appropriate condition such as an extended dimension condition of the form in (10). This will permit one to study trust-region problems in the face of data uncertainty by way of examining robust counterparts [7, 16, 18, 20] of the uncertain problems as they are often formulated as minimax optimization problems.

On the other hand, our notion of ρ-SOS-convexity can also be further extended in the sense that a real polynomial $p(x)$ of even degree d on $\mathbb{R}^n$ is ρ-SOS-convex if and only if $p(x) - \rho(\sum_{i=1}^n x_i^2)^{d/2}$ is SOS-convex for some $\rho \in \mathbb{R}$. This notion may allow us to treat a broader class of non-convex optimization problems involving higher degree polynomials than the class of non-convex quadratic optimization problems. They will be examined in a forthcoming study.

Acknowledgements. The authors are grateful to Dr Guoyin Li, University of New South Wales, for his comments on the preliminary version of the paper and to the referees for their valuable comments and suggestions that have contributed to the final preparation of the paper.

References

[1] A. A. Ahmadi, P. A. Parrilo: A convex polynomial that is not SOS-convex, Math. Program. 135 (2012) 275–292.

[2] A. A. Ahmadi, P. A. Parrilo: A complete characterization of the gap between convexity and SOS-convexity, SIAM J. Optim. 23(2) (2013) 811–833.

[3] W. Ai, S. Zhang: Strong duality for the CDT subproblem: a necessary and sufficient condition, SIAM J. Optim. 19(4) (2009) 1735–1756.

[4] M. Anjos, J. B. Lasserre: Handbook on Semidefinite, Conic and Polynomial Optimization, Springer, New York (2012).

[5] A. Beck, Y. C. Eldar: Strong duality in nonconvex quadratic optimization with two quadratic constraints, SIAM J. Optim. 17 (2006) 844–860.

[6] A. Beck: Convexity properties associated with nonconvex quadratic matrix functions and applications to quadratic programming, J. Optim. Theory Appl. 142 (2009) 1–29.

[7] A. Ben-Tal, L. E. Ghaoui, A. Nemirovski: Robust Optimization, Princeton Series in Applied Mathematics, Princeton University Press, Princeton (2009).

[8] A. Ben-Tal, A. Nemirovski: Lectures on Modern Convex Optimization. Analysis, Algorithms and Engineering Applications, SIAM-MPS, Philadelphia (2000).

[9] S. Burer, K. M. Anstreicher: Second-order-cone constraints for extended trust-region subproblems, SIAM J. Optim. 23(1) (2013) 432–451.

[10] M. Castellani, M. Pappalardo: Characterizations of ρ-convex functions, in: Generalized Convexity, Generalized Monotonicity: Recent Results (Luminy, 1996), J.-P. Crouzeix et al. (ed.), Nonconvex Optim. Appl. 27, Kluwer, Dordrecht (1998) 219–233.

[11] A. R. Conn, N. I. M. Gould, P. L. Toint: Trust-Region Methods, MPS-SIAM Series in Optimization 1, SIAM, Philadelphia (2000).

[12] A. M. Geoffrion: Duality in nonlinear programming: a simplified applications-oriented develoment, in: Perspectives on Optimization, A. M. Geoffrion (ed.), Addison-Wesley, Reading (1972) 65–101.

[13] J. W. Helton, J. W. Nie: Semidefinite representation of convex sets, Math. Program. 122(1) (2010) 21–64.

[14] V. Jeyakumar: On subgradient duality with strong and weak convex functions, J. Aust. Math. Soc., Ser. A 40 (1986) 143–152.

[15] V. Jeyakumar, G. M. Lee, G. Y. Li: Alternative theorems for quadratic inequality systems and global quadratic optimization, SIAM J. Optim. 20(2) (2009) 983–1001.

[16] V. Jeyakumar, G. Li: Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization, Math. Program. 147 (2014) 171–206.

[17] V. Jeyakumar, G. Li: Exact SDP relaxations for classes of nonlinear semidefinite programming problems, Oper. Res. Lett. 40 (2012) 529–536.

[18] V. Jeyakumar, G. Li, J. Vicente-Pérez: Robust SOS-convex polynomial optimization problems: exact SDP relaxations, Optim. Lett. 9(1) 2015 1–18.

[19] V. Jeyakumar, J. Vicente-Pérez: Dual semidefinite programs without duality gaps for a class of convex minimax programs, J. Optim. Theory Appl. 162 (2014) 735–753.

[20] V. Jeyakumar, G. Li, G. M. Lee: Robust duality for generalized convex programming problems under data uncertainty, Nonlinear Anal. 75 (2012) 1362–1373.

[21] J. B. Lasserre: Moments, Positive Polynomials and Their Applications, Imperial College Press, London (2009).

[22] O. L. Mangasarian: Nonlinear Programming, SIAM, Philadelphia (1994).

[23] B. T. Polyak: Convexity of quadratic transformations and its use in control and optimization, J. Optim. Theory Appl. 99 (1998) 563–583.

[24] J.-Ph. Vial: Strong and weak convexity of sets and functions, Math. Oper. Res. 8 (1983) 231–259.

[25] J.-Ph. Vial: Strong convexity of sets and functions, J. Math. Econ. 9 (1982) 187–205.

[26] Z. Y. Wu: Sufficient optimality conditions for weakly convex minimization problems, J. Glob. Optim. 39 (2007) 427–440.