Strong Duality for the CDT Subproblem: A Necessary and Sufficient Condition

Wenbao Ai *　　　Shuzhong Zhang †

October 2007; revised June 2008

Abstract

In this paper we consider the problem of minimizing a nonconvex quadratic function, subject to two quadratic inequality constraints. As an application, such quadratic program plays an important role in the trust region method for nonlinear optimization; such problem is known as the CDT subproblem in the literature. The Lagrangian dual of the CDT subproblem is a Semidefinite Program (SDP), hence convex and solvable. However, a positive duality gap may exist between the CDT subproblem and its Lagrangian dual because the CDT subproblem itself is nonconvex. In this paper, we present a necessary and sufficient condition to characterize when the CDT subproblem and its Lagrangian dual admits no duality gap (i.e., the strong duality holds). This necessary and sufficient condition is easy verifiable and involves only one (any) optimal solution of the SDP relaxation for the CDT subproblem. Moreover, the condition reveals that it is actually rare to render a positive duality gap for the CDT subproblems in general. Moreover, if the strong duality holds then an optimal solution for the CDT problem can be retrieved from an optimal solution of the SDP relaxation, by means of a matrix rank-one decomposition procedure. The same analysis is extended to the framework where the necessary and sufficient condition is presented in terms of the Lagrangian multipliers at a KKT point. Furthermore, we show that the condition is numerically easy to work with approximatively.

Keywords: Quadratically constrained quadratic programming, strong Lagrangian duality, CDT subproblem, SDP relaxation.

Mathematics Subject Classification: 90C33, 90C51, 90C05.

*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China. Email: wenbaoai@vip.sina.com. Research partially supported by Chinese NSFC Earmarked Grant Project10701016.

†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: zhang@se.cuhk.edu.hk. Research supported by Hong Kong RGC Earmarked Grant CUHK418505.

1 Introduction

In this paper we consider the following nonconvex quadratic optimization problem

$$(QP) \quad \begin{array}{ll} \text{minimize} & q_0(x) = x^T Q_0 x - 2b_0^T x \\ \text{subject to} & q_i(x) = x^T Q_i x - 2b_i^T x + c_i \le 0, \quad i = 1, ..., m. \end{array}$$

In case $m = 1$ and $Q_1 \succ 0$, the problem is known as the *trust region subproblem*, since in the trust region approach to unconstrained optimization such problems need to be solved repeatedly. In this context, the problem has been thoroughly studied. (For general information on the trust region method, see [7]). It is known that the trust region subproblem can be easily solved. A connection between the solution methods for the trust region subproblem and Semidefinite Programming (SDP) was established by Sturm and Zhang in [12]. By using a matrix rank-one decomposition procedure, Sturm and Zhang [12] showed that if $m = 1$ then the SDP relaxation of (QP) is tight, and an optimal solution for (QP) can be obtained from an optimal solution of its SDP relaxation. Furthermore, Ye and Zhang [14] showed that if $m = 2$ and certain additional conditions are satisfied then the SDP relaxation for (QP) can still be tight in many cases. In fact, the quadratic program (QP) with $m = 2$ has its own history as an extended trust region subproblem. In 1985, Celis, Dennis and Tapia [4] proposed a trust region method for constrained optimization, in which (QP) with $m = 2$ plays the role as a model for validating a trust region step. In this particular context, $Q_1 \succ 0$ and $Q_2 \succeq 0$, and the extended trust region subproblem is also referred to as the CDT subproblem. A number of papers have been devoted to studying the structure and the solution algorithms for the CDT subproblem; see e.g. [5, 6, 10, 11, 12, 14, 15, 16, 17].

A remarkable property which makes the CDT subproblem interesting and intriguing is that at a global optimal solution, the Hessian matrix of the Lagrangian function may not necessarily be positive semidefinite; however, it can have at most one negative eigenvalue (see Yuan [15]). In fact, it is quite rare to encounter examples where the Hessian of the Lagrangian function indeed has a negative eigenvalue at optimum. In 1991, Yuan [16] suggested an algorithm for the CDT subproblem under the assumption that the objective function is convex, and in 1992, Zhang [17] proposed an algorithm for the CDT subproblem under the assumption that the optimal Lagranian Hessian matrix is positive semidefinite. Chen and Yuan [6] presented a sufficient condition (termed as *Property $\mathcal{J}$* in [6]) under which the Lagrangian function of the CDT subproblem will have a positive semidefinite Hessian at optimal point. Recently, Beck and Eldar [2] used the complex valued SDP (thus relaxed) approach to come up with a similar sufficient condition to guarantee the nonnegativity of the Hessian matrix of the Lagrangian function at optimum. Beck and Eldar [2] reported that in their experiments on randomly generated instances, their sufficient condition was satisfied for an overwhelming majority of the random instances.

The current paper is concerned with the CDT type quadratic programs. In particular, we shall present a verifiable condition which indicates whether or not the SDP relaxation for the quadratic program is tight. Since the Lagrangian dual of a general quadratically constrained quadratic program is the dual of its SDP relaxation (see Chapter 13 of [13]), our result is equivalent to a necessary and sufficient condition for the strong duality to hold for this class of nonconvex quadratic programs. Our condition involves only the information of an optimal SDP solution, or alternatively, the information of a given KKT point. The paper is organized as follows. In Section 2, we shall formally establish the equivalence between the nonnegativity of the Hessian matrix of the Lagrangian function (of (QP)) at an arbitrary optimal solution and the fact that the SDP relaxation is tight. Section 3 is devoted to a specific problem related to the rank-one decomposition of a positive semidefinite matrix. This technical result is interesting in its own right, and it is used in Section 4 to derive a *necessary and sufficient* condition to check whether or not the SDP relaxation is indeed tight. Because the dual of the SDP relaxation coincides with the Lagrangian dual of (QP), a tight SDP relaxation manifests that the strong duality holds for (QP). Our necessary and sufficient condition is different from the other two sufficient conditions previously studied in [6] and [2]. In nonlinear programming, it is customary to use terminologies such as the Lagrangian multipliers or the Karush-Kuhn-Tucker (KKT) conditions. For this reason, we shall present our results in Section 5 both as an easy verifiable condition based on an optimal solution of the SDP relaxation, or alternatively, as an easy verifiable condition based on a KKT point in terms of the Lagrangian function and multipliers. An example is given in Section 6 to show that the information carried by the KKT solutions may not be useful for the optimal solution of the CDT Problem when the strong duality fails. In Section 7 we propose a numerical implementation of the necessary and sufficient condition. Our simulation results show that the condition is indeed numerically stable and easy to work with.

Throughout the paper, $\mathcal{S}^{n\times n}$ denotes the set of real $n \times n$ symmetric matrices; $\mathcal{S}^{n\times n}_{+}$ denotes the set of real $n \times n$ positive semi-definite matrices; $\mathcal{S}^{n\times n}_{++}$ denotes the set of real $n \times n$ positive definite matrices; for $A, B \in \mathcal{S}^{n\times n}$, $A \bullet B := \mathrm{tr}\, AB$ denotes the matrix inner-product between A and B.

2 Convex Lagrangian function and the strong duality

In the literature there are mainly two ways to solve a general quadratically constrained quadratic program (QP): either to use the Lagrangian function with some appropriately chosen multipliers, or to base the solution method on the SDP relaxation. In the latter case, the method works well if the SDP relaxation is tight, while in the former case the method works well if the Hessian of the Lagrangian function is positive semidefinite. It is therefore natural to believe that these two properties must be essentially identical. In this section we shall formally prove this point. The result is useful for our subsequent analysis.

First of all, following [12] we use the notation:

$$M(q_0) := \begin{bmatrix} 0 & -b_0^T \\ -b_0 & Q_0 \end{bmatrix}, \; M(q_i) := \begin{bmatrix} c_i & -b_i^T \\ -b_i & Q_i \end{bmatrix}, \text{ for } i = 1, ..., m.$$

Then, (QP) is equivalently written as

$$\begin{array}{lll} (QP) & \text{minimize} & M(q_0) \bullet \begin{bmatrix} t \\ x \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}^T = x^T Q_0 x - 2b_0^T x t \\ & \text{subject to} & M(q_i) \bullet \begin{bmatrix} t \\ x \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}^T = x^T Q_i x - 2b_i^T x t + c_i t^2 \leq 0, \quad i = 1, ..., m \\ & & t^2 = 1. \end{array}$$

The so-called SDP relaxation of (QP) is

$$\begin{array}{lll} (SP) & \text{minimize} & M(q_0) \bullet X \\ & \text{subject to} & M(q_i) \bullet X \leq 0, \quad i = 1, ..., m \\ & & I_{00} \bullet X = 1 \\ & & X \succeq 0, \end{array}$$

where $I_{00} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{0} \end{bmatrix} \in \mathcal{S}^{(n+1)\times(n+1)}$. The dual problem of (SP) is:

$$\begin{array}{lll} (SD) & \text{maximize} & y_0 \\ & \text{subject to} & Z = M(q_0) - y_0 I_{00} + \sum_{i=1}^m y_i M(q_i) \succeq 0 \\ & & y_i \geq 0, i = 1, ..., m. \end{array}$$

Note that (SD) is also the Lagrangian dual problem for (QP) ([13]). The following well-known facts regarding the relationship between (SP) and (SD) are either straightforward or well known:

1. (SP) satisfies the Slater condition if the original problem (QP) satisfies the Slater condition.

2. (SD) satisfies the Slater condition if at least one of the matrices Q_i's, $i = 0, 1, \cdots, m$, is positive definite.

3. If both (SP) and (SD) satisfy the Slater condition, then (SP) and (SD) have attainable optimal solutions. Moreover, a primal-dual feasible pair X and $(Z, y_0, y_1, \cdots, y_m)$ are optimal if and only if they satisfy the complementary conditions:

$$XZ = 0, \quad y_i M(q_i) \bullet X = 0, \quad i = 1, ..., m.$$

Throughout this paper we assume that $Q_1 \succ 0$ and that (QP) satisfied the Slater condition. Hence, (QP), (SP) and (SD) all have optimal solutions, which we shall denote respectively by x^*, $\hat{X}$ and $(\hat{Z}, \hat{y_0}, \hat{y_1}, \cdots, \hat{y_m})$, and their optimal values respectively by $v(QP)$, $v(SP)$ and $v(SD)$.

Clearly, $v(SP) \le v(QP)$ since (SP) is a relaxation of (QP), and $v(SP) = v(SD)$ since both (SP) and (SD) satisfy the Slater condition. Therefore, the strong duality holds for (QP) if and only if the SDP relaxation for (QP) is tight; i.e., $v(SP) = v(QP)$. It is helpful to keep in mind that $\hat{Z}$ can also be rewritten as

$$\hat{Z} = \begin{bmatrix} -\hat{y}_0 + \sum_{i=1}^m \hat{y}_i c_i & -b_0^T - \sum_{i=1}^m \hat{y}_i b_i^T \\ -b_0 - \sum_{i=1}^m \hat{y}_i b_i & Q_0 + \sum_{i=1}^m \hat{y}_i Q_i \end{bmatrix}.$$

On the other hand, the Lagrangian function for (QP), with y_i being the multiplier for the constraint $q_i(x) \le 0$, $i = 1, ..., m$, is given as

$$L(x; y) := q_0(x) + \sum_{i=1}^m y_i q_i(x).$$

Clearly, since the function is quadratic in x for any fixed multiplier y, its Hessian matrix is $\nabla_{xx}^2 L(x; y) = Q_0 + \sum_{i=1}^m y_i Q_i$.

Theorem 2.1. $v(SP) = v(QP) \iff \nabla_{xx}^2 L(x; y) = Q_0 + \sum_{i=1}^m y_i Q_i \succeq 0$ *where* y *is the Lagrangian multiplier for an optimal solution of* (QP).

Proof. "$\Longrightarrow$": For any minimizer x^* of the original problem (QP), the matrix $X^* := \begin{bmatrix} 1 \\ x^* \end{bmatrix} \begin{bmatrix} 1 \\ x^* \end{bmatrix}^T$ is also an optimal solution for (SP). So the primal-dual optimal pair X^* and $(\hat{Z}, \hat{y})$ satisfy complementary conditions, where $(\hat{Z}, \hat{y})$ is optimal to (SD), i.e.

$$\hat{Z} X^* = 0, \quad \hat{y}_i M(q_i) \bullet X^* = 0, \quad i = 1, ..., m. \tag{2.1}$$

Since $\hat{Z} \succeq 0$, the relation $\hat{Z} X^* = 0$ is equivalent to $\hat{Z} \begin{bmatrix} 1 \\ x^* \end{bmatrix} = 0$, which implies that

$$(Q_0 + \sum_{i=1}^m \hat{y}_i Q_i) x^* = b_0 + \sum_{i=1}^m \hat{y}_i b_i.$$

Also, since $q_i(x^*) = M(q_i) \bullet X^*$ it follows from (2.1) that $\hat{y}_i q_i(x^*) = 0$, $i = 1, ..., m$. Therefore, x^* and $\hat{y}$ satisfy the KKT condition and $\hat{y}$ is the corresponding Lagrangian multiplier with the Hessian matrix being

$$\nabla_{xx}^2 \left(q_0(x) + \sum_{i=1}^m \hat{y}_i q_i(x) \right) \Bigg|_{x=x^*} = Q_0 + \sum_{i=1}^m \hat{y}_i Q_i = \hat{Z} \succeq 0.$$

"$\Longleftarrow$": Suppose that the original problem (QP) has an optimal solution x^* with a positive semidefinite Lagrangian Hessian matrix $Q_0 + \sum_{i=1}^m y_i^* Q_i$. Then x^* and $y_1^*, ..., y_m^*$ satisfy the following KKT condition:

$$(Q_0 + \sum_{i=1}^m y_i^* Q_i)x^* = b_0 + \sum_{i=1}^m y_i^* b_i, \quad y_i^* q_i(x^*) = 0, \quad y_i^* \geq 0, i = 1, ..., m.$$

Let

$$X^* := \begin{bmatrix} 1 \\ x^* \end{bmatrix} \begin{bmatrix} 1 \\ x^* \end{bmatrix}^T, \quad y_0^* := q_0(x^*), \text{ and } Z^* := M(q_0) - y_0^* I_{00} + \sum_{i=1}^m y_i^* M(q_i). \tag{2.2}$$

Next, we aim to show that X^* and (y^*, Z^*) are in fact optimal to (SP) and (SD). To this end, we need only to verify that $Z^* \succeq 0$ and $Z^* X^* = 0$.

Let the Lagrangian function be

$$L(x; y^*) = q_0(x) + \sum_{i=1}^m y_i^* q_i(x). \tag{2.3}$$

By the Taylor expansion at x^* and the KKT optimality condition, we have

$$L(x; y^*) = L(x^*; y^*) + (x - x^*)^T (Q_0 + \sum_{i=1}^m y_i^* Q_i)(x - x^*) \geq L(x^*; y^*) = q_0(x^*) = y_0^* \tag{2.4}$$

for any x, which implies that x^* is a global minimizer of $L(x; y^*)$. Consider any $(n+1)$-dimensional vector $\begin{bmatrix} t \\ x \end{bmatrix}$. If $t \neq 0$, then it follows from (2.2), (2.3) and (2.4) that

$$\begin{aligned} \begin{bmatrix} t \\ x \end{bmatrix}^T Z^* \begin{bmatrix} t \\ x \end{bmatrix} &= \left(M(q_0) - y_0^* I_{00} + \sum_{i=1}^m y_i^* M(q_i) \right) \bullet \begin{bmatrix} t \\ x \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}^T \\ &= t^2 \left(q_0(x/t) + \sum_{i=1}^m y_i^* q_i(x/t) - y_0^* \right) \\ &= t^2 \left(L(x/t; y^*) - y_0^* \right) \geq 0. \end{aligned}$$

If $t = 0$, then

$$\begin{bmatrix} t \\ x \end{bmatrix}^T Z^* \begin{bmatrix} t \\ x \end{bmatrix} = x^T (Q_0 + \sum_{i=1}^m y_i^* Q_i) x \geq 0.$$

Therefore, $Z^* \succeq 0$. Moreover,

$$Z^* \bullet X^* = \begin{bmatrix} 1 \\ x^* \end{bmatrix}^T Z^* \begin{bmatrix} 1 \\ x^* \end{bmatrix} = L(x^*; y^*) - y_0^* = 0,$$

which, together with $Z^* \succeq 0$ and $X^* \succeq 0$, implies that $Z^* X^* = 0$. $\square$

Theorem 2.1 implies that once an optimal solution for (QP) admits a Lagrangian multiplier (vector) with nonnegative Hessian matrix then $v(QP) = v(SP)$, which in turn implies that *every* optimal solution admits a Lagrangian multiplier with nonnegative Hessian matrix. We formalize the statement as follows.

Corollary 2.2. *If at one optimal solution of (QP) there is a Lagrangian multiplier with a nonnegative Hessian matrix, then it follows that at any optimal solution of (QP) there is a Lagrangian multiplier with a nonnegative Hessian matrix.*

We note that Theorem 2.1 can actually be used to bridge an easy provable fact to a less obvious one. For instance, it is relatively easy to show that if $m = 1$ then the Hessian of the Lagrangian function is nonnegative (e.g. Theorem 7.2.1 in [7]). Hence we can conclude $v(QP) = v(SP) = v(SD)$ in this case, simply using Theorem 2.1. On the other hand, if $m = 2$ and (QP) is homogeneous (i.e. $b_0 = b_1 = b_2 = 0$), then Ye and Zhang [14] (Section 2.2) showed that $v(QP) = v(SP)$. A less obvious fact is that the Lagrangian function always has a nonnegative Hessian matrix in this case.

3 A new matrix rank-one decomposition procedure

Sturm and Zhang [12] proposed a simple (polynomial-time) procedure to compute the following matrix rank-one decomposition problem. Given $X \in \mathcal{S}_+^{n\times n}$ and $A \in \mathcal{S}^{n\times n}$, find $x_j \in \Re^n$, $j = 1, ..., r$, where $r = \text{rank}(A)$, such that $X = \sum_{j=1}^r x_j x_j^T$ and $x_j^T A x_j = A \bullet X/r$, $j = 1, ..., r$. Huang and Zhang [9] extended the result to the case where the matrices in questions are all Hermitian.

The aim of this section is to study a further extension of such rank-one decomposition in the real symmetric case. Our result will then be applied in the next section to enable a method for (QP) when $m = 2$. Let $x_1 \in \Re^n$ and $X \in \mathcal{S}_+^{n\times n}$. As a convention we shall call matrix X to be *rank-one decomposable* at x_1 if there exist other $r-1$ vectors $x_2, ..., x_r$ such that $X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T$, where $r = \text{rank}(X)$. To find out when is a matrix rank-one decomposable at a given vector, we first note the following lemma.

Lemma 3.1. *Suppose that $X \in \mathcal{S}_+^{n\times n}$ with* $\text{rank}(X) = r$, *and $X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T$. Let $X_r = [x_1, \cdots, x_r]$. Then, $X = y_1y_1^T + y_2y_2^T + \cdots + y_ry_r^T$ holds with $Y_r = [y_1, \cdots, y_r]$ iff there exists an orthonormal matrix $P \in \Re^{r\times r}$ such that $Y_r = X_rP$.*

Proof. The sufficiency is obvious. To show the necessity of the condition, let us suppose $X = X_rX_r^T = Y_rY_r^T$, and consider $P = X_r^TY_r(Y_r^TY_r)^{-1}$. Clearly,

$$P^TP = (Y_r^TY_r)^{-1}Y_r^TX_rX_r^TY_r(Y_r^TY_r)^{-1} = (Y_r^TY_r)^{-1}Y_r^TY_rY_r^TY_r(Y_r^TY_r)^{-1} = I_r.$$

Hence P is an orthonormal matrix. At the same time,

$$X_r P = X_r X_r^T Y_r (Y_r^T Y_r)^{-1} = Y_r Y_r^T Y_r (Y_r^T Y_r)^{-1} = Y_r.$$

$\square$

Since for any given unit vector one can always construct an orthonormal matrix with this unit vector as the first column, this leads to the following characterization of the rank-one decomposability at a given vector.

Proposition 3.2. *Suppose that $X \in \mathcal{S}_+^{n\times n}$ with $\operatorname{rank}(X) = r$, and $X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T$. Let $X_r = [x_1, \cdots, x_r]$. Then, X is rank-one decomposable at $y \in \Re^n$ if and only if there is $u \in \Re^r$ with $\|u\| = 1$ and $y = X_r u$.*

The next result plays an important role in this paper.

Lemma 3.3. *Let $A_1, A_2 \in \mathcal{S}^{n\times n}$. Suppose that $X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T$, where $r \geq 3$. If*

$$\begin{array}{l} A_1 \bullet x_1x_1^T = A_1 \bullet x_2x_2^T = \delta_1, \\ (A_2 \bullet x_1x_1^T - \delta_2)(A_2 \bullet x_2x_2^T - \delta_2) < 0, \end{array} \tag{3.1}$$

then in the real-number computation sense (viz. the BSS model [3]), one can find in polynomial-time a vector $y \in \Re^n$ such that X is rank-one decomposable at y and

$$\begin{array}{l} A_1 \bullet yy^T = \delta_1, \\ A_2 \bullet yy^T = \delta_2. \end{array} \tag{3.2}$$

Proof. Without loss of generality, we assume that

$$A_2 \bullet x_1x_1^T - \delta_2 > 0, \text{ and } A_2 \bullet x_2x_2^T - \delta_2 < 0. \tag{3.3}$$

For given real values α_i, $i = 1, 2, 3$, with $(\alpha_1, \alpha_2, \alpha_3) \neq 0$, define

$$y = \frac{\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}}. \tag{3.4}$$

By Proposition 3.2, X is rank-one decomposable at y. Let us substitute (3.4) into (3.2) and consider the following system of equations with respect to the unknown real variables α_1, α_2 and α_3:

$$0 = \alpha_3^2(A_1 \bullet x_3x_3^T - \delta_1) + 2\alpha_1\alpha_2 A_1 \bullet x_1x_2^T + 2\alpha_1\alpha_3 A_1 \bullet x_1x_3^T + 2\alpha_2\alpha_3 A_1 \bullet x_2x_3^T, \tag{3.5}$$

$$\begin{aligned} 0 = \; & \alpha_1^2(A_2 \bullet x_1x_1^T - \delta_2) + \alpha_2^2(A_2 \bullet x_2x_2^T - \delta_2) + \alpha_3^2(A_2 \bullet x_3x_3^T - \delta_2) \\ & + 2\alpha_1\alpha_2 A_2 \bullet x_1x_2^T + 2\alpha_1\alpha_3 A_2 \bullet x_1x_3^T + 2\alpha_2\alpha_3 A_2 \bullet x_2x_3^T. \end{aligned} \tag{3.6}$$

In fact, it follows from Finsler's lemma [8] that equations (3.5) and (3.6) admit a real-valued solution $(\alpha_1, \alpha_2, \alpha_3)$. However, Finsler's lemma is a pure existence result. Below we shall construct such solutions. We proceed by considering two cases.

Case 1. $A_1 \bullet x_1 x_2^T = 0$.

We choose $\alpha_1 = 1, \alpha_3 = 0$. Then equation (3.5) is trivially satisfied for any values of α_2, and equation (3.6) can be rewritten as follows:

$$(A_2 \bullet x_1 x_1^T - \delta_2) + \alpha_2^2 (A_2 \bullet x_2 x_2^T - \delta_2) + 2\alpha_2 A_2 \bullet x_1 x_2^T = 0,$$

which is a quadratic equation in α_2 and must have two distinct real roots because of (3.3), one is positive, and another is negative. Let $\bar{\alpha}_2$ be one of the roots. Then $(\alpha_1, \alpha_2, \alpha_3) = (1, \bar{\alpha}_2, 0)$ is a solution for (3.5) and (3.6).

Case 2. $A_1 \bullet x_1 x_2^T \neq 0$.

We choose $\alpha_3 = 1$. Then (3.5) and (3.6) become

$$\begin{aligned} 0 &= 2\alpha_2 \left(\alpha_1 A_1 \bullet x_1 x_2^T + A_1 \bullet x_2 x_3^T \right) + 2\alpha_1 A_1 \bullet x_1 x_3^T + (A_1 \bullet x_3 x_3^T - \delta_1), && (3.7) \\ 0 &= \alpha_1^2 (A_2 \bullet x_1 x_1^T - \delta_2) + \alpha_2^2 (A_2 \bullet x_2 x_2^T - \delta_2) + 2\alpha_1 \alpha_2 A_2 \bullet x_1 x_2^T \\ &\quad + 2\alpha_1 A_2 \bullet x_1 x_3^T + 2\alpha_2 A_2 \bullet x_2 x_3^T + (A_2 \bullet x_3 x_3^T - \delta_2). && (3.8) \end{aligned}$$

Solving (3.7) yields

$$\alpha_2 = -\frac{2\alpha_1 A_1 \bullet x_1 x_3^T + (A_1 \bullet x_3 x_3^T - \delta_1)}{2(\alpha_1 A_1 \bullet x_1 x_2^T + A_1 \bullet x_2 x_3^T)} \triangleq p(\alpha_1). \quad (3.9)$$

Moreover, let us denote

$$\begin{aligned} g(\alpha_1, \alpha_2) \quad &:= \quad \alpha_1^2 (A_2 \bullet x_1 x_1^T - \delta_2) + \alpha_2^2 (A_2 \bullet x_2 x_2^T - \delta_2) + 2\alpha_1 \alpha_2 A_2 \bullet x_1 x_2^T \\ &\quad + 2\alpha_1 A_2 \bullet x_1 x_3^T + 2\alpha_2 A_2 \bullet x_2 x_3^T + (A_2 \bullet x_3 x_3^T - \delta_2), \end{aligned} \quad (3.10)$$

and define

$$t_1 := -\frac{A_1 \bullet x_2 x_3^T}{A_1 \bullet x_1 x_2^T}.$$

We consider the following two possible subcases.

Case 2.1. $\det \begin{bmatrix} 2A_1 \bullet x_1 x_3^T & A_1 \bullet x_3 x_3^T - \delta_1 \\ A_1 \bullet x_1 x_2^T & A_1 \bullet x_2 x_3^T \end{bmatrix} \neq 0.$

Since

$$\begin{aligned} & \left. \left(2\alpha_1 A_1 \bullet x_1 x_3^T + (A_1 \bullet x_3 x_3^T - \delta_1) \right) \right|_{\alpha_1 = t_1} \\ = \quad & -\det \begin{bmatrix} 2A_1 \bullet x_1 x_3^T & A_1 \bullet x_3 x_3^T - \delta_1 \\ A_1 \bullet x_1 x_2^T & A_1 \bullet x_2 x_3^T \end{bmatrix} \Big/ A_1 \bullet x_1 x_2^T \neq 0, \end{aligned}$$

the function $p(\alpha_1)$ has the the properties that

$$\lim_{\alpha_1 \to t_1} p(\alpha_1) = \infty \tag{3.11}$$

and

$$\lim_{\alpha_1 \to \infty} p(\alpha_1) = -\frac{A_1 \bullet x_1 x_3^T}{A_1 \bullet x_1 x_2^T}. \tag{3.12}$$

Substituting (3.9) into (3.10) we obtain an equation in α_1,

$$\begin{aligned} g(\alpha_1, p(\alpha_1)) \quad := \quad & \alpha_1^2(A_2 \bullet x_1 x_1^T - \delta_2) + p(\alpha_1)^2(A_2 \bullet x_2 x_2^T - \delta_2) + 2\alpha_1 p(\alpha_1) A_2 \bullet x_1 x_2^T \\ & +2\alpha_1 A_2 \bullet x_1 x_3^T + 2p(\alpha_1) A_2 \bullet x_2 x_3^T + (A_2 \bullet x_3 x_3^T - \delta_2) \\ = \quad & 0, \end{aligned}$$

which is essentially a quartic polynomial equation in α_1. Since

$$\lim_{\alpha_1 \to t_1} g(\alpha_1, p(\alpha_1)) = -\infty$$

and

$$\lim_{\alpha_1 \to \infty} g(\alpha_1, p(\alpha_1)) = +\infty$$

due to (3.3), (3.11) and (3.12) it follows that $g(\alpha_1, p(\alpha_1))$ has at least one real root $\bar{\alpha}_1$ in the interval $(t_1, +\infty)$. Moreover, such root can be found by solving a quartic polynomial equation with the standard root-finding formula, which can be regarded as a constant operation in the BSS computational model. Substituting back, we derive $(\bar{\alpha}_1, p_1(\bar{\alpha}_1), 1)$ as a solution for (3.5) and (3.6).

Case 2.2. $\det \begin{bmatrix} 2A_1 \bullet x_1 x_3^T & A_1 \bullet x_3 x_3^T - \delta_1 \\ A_1 \bullet x_1 x_2^T & A_1 \bullet x_2 x_3^T \end{bmatrix} = 0.$

The above implies that there exists k such that

$$(2A_1 \bullet x_1 x_3^T, A_1 \bullet x_3 x_3^T - \delta_1) = k(A_1 \bullet x_1 x_2^T, A_1 \bullet x_2 x_3^T).$$

Thus (3.7) becomes

$$(\alpha_1 A_1 \bullet x_1 x_2^T + A_1 \bullet x_2 x_3^T)(2\alpha_2 + k) = 0,$$

for which the roots are

$$\alpha_1 = -\frac{A_1 \bullet x_2 x_3^T}{A_1 \bullet x_1 x_2^T} =: t_1, \quad \alpha_2 \text{ arbitrary;}$$

and

$$\alpha_2 = -k/2 =: t_2, \quad \alpha_1 \text{ arbitrary.}$$

Substituting them back into (3.8), it suffices to solve either $g(t_1, \alpha_2) = 0$ or $g(\alpha_1, t_2) = 0$, which are quadratic equations in α_2 and α_1 respectively. If $g(t_1, \alpha_2)$ has a real root $\bar{\alpha}_2$ then $(t_1, \bar{\alpha}_2, 1)$ is a solution to (3.5) and (3.6); otherwise, we have

$$g(t_1, \alpha_2) < 0 \text{ for all } \alpha_2$$

as $\lim_{\alpha_2\to+\infty} g(t_1,\alpha_2) = -\infty$ due to (3.3). In particular,

$$g(t_1,t_2) < 0.$$

Thus $g(\alpha_1,t_2)$ has a real zero point $\bar{\alpha}_1$ for α_1 on the interval $(t_1,+\infty)$ as $\lim_{\alpha_1\to+\infty} g(\alpha_1,t_2) = +\infty$ due to (3.3). Then $(\bar{\alpha}_1,t_2,1)$ is a solution to (3.5) and (3.6). $\square$

Remark that in Lemma 3.3, we require that $r=3$. This condition cannot be removed. Consider the following example:

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix},\ X = x_1x_1^T + x_2x_2^T = \begin{bmatrix} 1 \\ -1 \end{bmatrix}[1,-1] + \begin{bmatrix} 1 \\ 1 \end{bmatrix}[1,1].$$

Clearly,

$$\begin{array}{l} A_1 \bullet x_1x_1^T = A_1 \bullet x_2x_2^T = 0, \\ A_2 \bullet x_1x_1^T = -2 < 0, A_2 \bullet x_2x_2^T = 2 > 0. \end{array}$$

However, for any nonzero $x \in \Re^2$, $A_1 \bullet xx^T = 0$ if and only if x is either parallel to x_1 or to x_2, which implies that there is no nontrivial x satisfying both $A_1 \bullet xx^T = 0$ and $A_2 \bullet xx^T = 0$ simultaneously.

Using the above lemma we now show the following theorem.

Theorem 3.4. *Let $A_1, A_2 \in \mathcal{S}^{n\times n}$ and $X \in \mathcal{S}_+^{n\times n}$ with*

$$A_1 \bullet X = \delta_1, \quad A_2 \bullet X = \delta_2.$$

If $r := \mathrm{rank}(X) \geq 3$ then in polynomial-time (real-number computation) one finds a rank-one decomposition for X,

$$X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T,$$

such that

$$\begin{array}{l} A_1 \bullet x_ix_i^T = \delta_1/r \text{ for } i = 1,...,r \\ A_2 \bullet x_ix_i^T = \delta_2/r \text{ for } i = 1,...,r-2. \end{array}$$

Proof. We shall achieve the desired decomposition by the following steps. Initially, we set $X_0 := \emptyset$ and $X_1 := X$. By Lemma 2.2 of [14], one finds a rank-one decomposition for X_1,

$$X_1 = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T,$$

such that $A_1 \bullet x_ix_i^T = \delta_1/r$ for $i = 1,...,r$. Introduce an index set

$$I_0 := \{i \mid A_2 \bullet x_ix_i^T = \delta_2/r, i = 1,...,r\}$$

and then update X_0 and X_1 by setting

$$X_0 := X_0 + \sum_{i\in I_0} x_ix_i^T, \quad X_1 := X_1 - \sum_{i\in I_0} x_ix_i^T.$$

If $\text{rank}(X_1) < 3$ then the procedure is completed; otherwise, i.e., $\text{rank}(X_1) \geq 3$, using Lemma 3.3 we find y for which X_1 is rank-one decomposable at y, such that

$$A_1 \bullet yy^T = \delta_1/r, \quad A_2 \bullet yy^T = \delta_2/r.$$

Update X_0 and X_1 by letting

$$X_0 := X_0 + yy^T, \quad X_1 := X_1 - yy^T.$$

In this case, $\text{rank}(X_1)$ is reduced by 1. Repeat the above procedure until $\text{rank}(X_1) < 3$. □

4 Strong duality: a necessary and sufficient condition

In this section we consider (QP) with $m = 2$, which shall be denoted $(QP)_2$ hereafter. Without loss of generality, we assume $q_1(x) = x^T x - 1$; i.e.,

$$\begin{array}{lll} (QP)_2 & \text{minimize} & q_0(x) = x^T Q_0 x - 2b_0^T x \\ & \text{subject to} & q_1(x) = x^T x - 1 \leq 0 \\ & & q_2(x) = x^T Q_2 x - 2b_2^T x + c_2 \leq 0. \end{array}$$

The above problem is slightly more general than the CDT subproblem, in that Q_2 above can be indefinite. The central issue to be considered here is when the corresponding SDP relaxation for $(QP)_2$ is tight, which is shown in Section 2 to be equivalent to a strong Lagrangian duality (alternatively, it is also equivalent to the fact that the Lagrangian function has a positive semidefinite Hessian matrix at optimum due to Theorem 2.1). As before we assume throughout the discussion that the Slater condition is satisfied by $(QP)_2$.

Let $(SP)_2$ be the SDP relxation for $(QP)_2$ and $(SD)_2$ be the dual of $(SP)_2$; that is,

$$\begin{array}{lll} (SP)_2 & \text{minimize} & M(q_0) \bullet X \\ & \text{subject to} & M(q_1) \bullet X \leq 0 \\ & & M(q_2) \bullet X \leq 0 \\ & & I_{00} \bullet X = 1 \\ & & X \succeq 0, \end{array}$$

where

$$M(q_0) := \begin{bmatrix} 0 & -b_0^T \\ -b_0 & Q_0 \end{bmatrix}, M(q_1) := \begin{bmatrix} -1 & 0 \\ 0 & I_n \end{bmatrix}, M(q_2) := \begin{bmatrix} c_2 & -b_2^T \\ -b_2 & Q_2 \end{bmatrix}, I_{00} := \begin{bmatrix} 1 & 0 \\ 0 & O_n \end{bmatrix}.$$

As we observed earlier, $(SD)_2$ is also the Lagrangian dual of $(QP)_2$. Let $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ be a pair of optimal solutions to $(SP)_2$ and to $(SD)_2$ respectively. It turns out that the following property of $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ is important, which we shall call *Property* $\mathcal{I}$ for ease of reference.

Definition 4.1. *For $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$, a given pair of optimal solutions for $(SP)_2$ and $(SD)_2$ respectively, we say that this pair has Property $\mathcal{I}$ if:*

(1) $\hat{y}_1\hat{y}_2 \neq 0$*;*

(2) $\text{rank}\,(\hat{Z}) = n - 1$*;*

(3) $\text{rank}\,(\hat{X}) = 2$*, and there is a rank-one decomposition of* $\hat{X}$*,* $\hat{X} = \hat{x}_1\hat{x}_1^T + \hat{x}_2\hat{x}_2^T$*, such that* $M(q_1) \bullet \hat{x}_i\hat{x}_i^T = 0$*,* $i = 1, 2$*, and* $(M(q_2) \bullet \hat{x}_1\hat{x}_1^T)(M(q_2) \bullet \hat{x}_2\hat{x}_2^T) < 0$*.*

We remark here that it is easy to verify Property $\mathcal{I}$, once $(SP)_2$ and $(SD)_2$ are solved. The first two conditions being straightforward, the last one, due to Proposition 3.2, can be reduced to verifying the condition on a single parameter satisfying a quadratic equation (any 2 by 2 orthonormal matrix can be completely characterized by polar coordinates in a single parameter).

Theorem 4.2. *Consider $(QP)_2$ where the Slater condition is satisfied. Suppose that $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ are a pair of optimal solutions for its SDP relaxation problem $(SP)_2$ and the dual $(SD)_2$ respectively. Then, $v((SP)_2) < v((QP)_2)$ holds if and only if the pair $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ has Property $\mathcal{I}$.*

Proof. We shall complete the proof in two parts. They are, Part 1: if Property $\mathcal{I}$ does not hold then the SDP relaxation is tight; and Part 2: if Property $\mathcal{I}$ holds then the relaxation must not be tight.

In Part 1, we enumerate four exhaustive (but not mutually exclusive) possibilities, to be denoted by Part 1.i, with $i = 1, 2, 3, 4$.

Part 1.1. $\hat{y}_1\hat{y}_2 = 0$.

The proof that the SDP relaxation is tight in this case can be found in Ye and Zhang [14].

Part 1.2. $\hat{y}_1\hat{y}_2 \neq 0$ and $\text{rank}\,(\hat{X}) \neq 2$.

$\hat{y}_1\hat{y}_2 \neq 0$ implies by the complementary conditions that

$$\hat{Z}\hat{X} = 0, \quad M(q_1) \bullet \hat{X} = 0, \quad M(q_2) \bullet \hat{X} = 0.$$

Let $r := \text{rank}\,(\hat{X})$. Obviously, $r > 0$ since $I_{00} \bullet \hat{X} = 1$, and if $r = 1$ then the theorem is already true. Therefore we need only to consider the nontrivial case $r \geq 3$. By Theorem 3.4 there is a rank-one decomposition of $\hat{X}$ satisfying

$$\begin{aligned} &X = x_1x_1^T + x_2x_2^T + \cdots + x_rx_r^T \\ &M(q_1) \bullet x_ix_i^T = 0, \text{ for } i = 1, ..., r \\ &M(q_2) \bullet x_ix_i^T = 0, \text{ for } i = 1, ..., r - 2. \end{aligned}$$

Thus $x_1 x_1^T / t_1^2$ satisfies the complementary conditions hence optimal to $(SP)_2$. This implies that x_1/t_1 is a homogenized optimal solution to (QP), where t_1 denotes the first element of x_1, which must be nonzero because $M(q_1) \bullet x_1 x_1^T = 0$, $x_1 \neq 0$, and $Q_1 \succ 0$.

Part 1.3. $\hat{y}_1 \hat{y}_2 \neq 0$ and $\text{rank}\,(\hat{X}) = 2$, and $M(q_2) \bullet \hat{x}_1 \hat{x}_1^T = M(q_2) \bullet \hat{x}_2 \hat{x}_2^T = 0$.

In this case, both $\hat{x}_1 \hat{x}_1^T / \hat{t}_1^2$ and $\hat{x}_2 \hat{x}_2^T / \hat{t}_2^2$ are optimal to $(SP)_2$. Thus both $\hat{x}_1/\hat{t}_1$ and $\hat{x}_2/\hat{t}_2$ are optimal solutions for $(QP)_2$, where $\hat{t}_1$ and $\hat{t}_2$ are the first elements of $\hat{x}_1$ and $\hat{x}_2$ respectively, which are both nonzero as argued before.

Part 1.4. $\hat{y}_1 \hat{y}_2 \neq 0$ and $\text{rank}\,(\hat{X}) = 2$, $\left(M(q_2) \bullet \hat{x}_1 \hat{x}_1^T\right) \left(M(q_2) \bullet \hat{x}_2 \hat{x}_2^T\right) < 0$, and $\text{rank}\,(\hat{Z}) \neq n-1$.

Since $\text{rank}\,(\hat{Z}) + \text{rank}\,(\hat{X}) \leq n+1$ and $\text{rank}\,(\hat{X}) = 2$, it follows that $\text{rank}\,(\hat{Z}) \leq n-1$, and therefore in this particular case $\text{rank}\,(\hat{Z}) < n-1$. Now $\hat{X} + \hat{Z}$ is singular and both $\hat{X}$ and $\hat{Z}$ are positive semidefinite, so there must be a nontrivial y in the intersection of the null spaces of $\hat{X}$ and $\hat{Z}$. Let

$$X := \hat{X} + yy^T = \hat{x}_1 \hat{x}_1^T + \hat{x}_2 \hat{x}_2^T + yy^T.$$

Obviously, $\text{rank}\,(X) = 3$ and $\hat{Z}X = 0$. Since

$$M(q_1) \bullet \hat{x}_1 \hat{x}_1^T = M(q_1) \bullet \hat{x}_1 \hat{x}_1^T = 0, \tag{4.1}$$

$$(M(q_2) \bullet \hat{x}_1 \hat{x}_1^T)(M(q_2) \bullet \hat{x}_1 \hat{x}_1^T) < 0. \tag{4.2}$$

by applying Lemma 3.3 we obtain x such that X is rank-one decomposable at x and that

$$M(q_1) \bullet xx^T = 0, \quad M(q_2) \bullet xx^T = 0.$$

Since x is in the range space of X, it must be in the null space of $\hat{Z}$. That is, $\hat{Z} \bullet xx^T = 0$, implying that xx^T/t^2 is an optimal solution to $(SP)_2$ and x/t is an optimal solution to $(QP)_2$, where t is the first component of x (which must be nonzero as argued before).

This concludes Part 1.

Next we proceed to Part 2, in which we shall prove that if Property $\mathcal{I}$ holds then there is definitely a gap between $(QP)_2$ and $(SP)_2$, i.e., $v((SP_2) < v((QP)_2)$. To see why this is true, we use a contradiction argument. Suppose that Property $\mathcal{I}$ holds, while $v((SP_2) = v((QP)_2)$. Let x^* be an optimal solution of $(QP)_2$ (we extend the dimension of x^* to be $(n+1)$ dimensional by putting 1 in the first component). Then, since $v((SP_2) = v((QP)_2)$, $x^*(x^*)^T$ must be an optimal solution to $(SP)_2$. Consequently, $x^*(x^*)^T$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ must satisfy the complementarity condition; i.e.,

$$\hat{Z}x^*(x^*)^T = 0, \quad M(q_1) \bullet x^*(x^*)^T = 0, \quad M(q_2) \bullet x^*(x^*)^T = 0. \tag{4.3}$$

This implies that x^* must be in the null space of $\hat{Z}$, which is two-dimensional in this case. In other words, it must be a linear combination of $\hat{x}_1$ and $\hat{x}_2$. Let us assume that there are two numbers α and β such that

$$x^* = \alpha \hat{x}_1 + \beta \hat{x}_2. \tag{4.4}$$

Substituting (4.4) into the equations $M(q_1) \bullet x^*(x^*)^T = 0$ and $M(q_2) \bullet x^*(x^*)^T = 0$, and noting (4.1) and (4.2), we obtain

$$\alpha\beta\hat{x}_1^T M(q_1)\hat{x}_2 = 0, \tag{4.5}$$

$$\alpha^2 M(q_2) \bullet \hat{x}_1\hat{x}_1^T + 2\alpha\beta M(q_2) \bullet \hat{x}_1\hat{x}_2^T + \beta^2 M(q_2) \bullet \hat{x}_2\hat{x}_2^T = 0. \tag{4.6}$$

Due to (4.2), neither α nor β can be zero. (E.g., if $\alpha = 0$ then by (4.6) and (4.2) it necessarily follows that $\beta = 0$, and vice versa). Thus, from (4.5) it follows that that $\hat{x}_1^T M(q_1)\hat{x}_2 = 0$. Let $\hat{x}_1 = (t_1, u^T)^T$ and $\hat{x}_2 = (t_2, v^T)^T$ where $u, v \in \Re^n$. We have $0 = M(q_1) \bullet \hat{x}_1\hat{x}_1^T = t_1^2 - \|u\|^2$ and $0 = M(q_1)\bullet\hat{x}_2\hat{x}_2^T = t_2^2 - \|v\|^2$. Now, $\hat{x}_1^T M(q_1)\hat{x}_2 = 0$ leads to $0 = t_1t_2 - u^Tv$, and so $(u^Tv)^2 = \|u\|^2\|v\|^2$. By the Cauchy-Schwartz inequality, this is only possible when u is a multiple of v. Consequently, $\hat{x}_1$ and $\hat{x}_2$ must be linearly dependent, a contradiction to the fact that $2 = \operatorname{rank}(\hat{X}) = \hat{x}_1\hat{x}_1^T + \hat{x}_2\hat{x}_2^T$. $\square$

5 The Lagrangian function and the KKT condition

It is intuitively clear that the Lagrangian function must be related to the SDP relaxation, as Theorem 2.1 has already indicated, primarily due to the fact that the Lagrangian dual of quadratically constrained quadratic program (QP) is identical to the dual of its SDP relaxation. It is, however, useful to translate Property $\mathcal{I}$ using the terms of the Lagrangian function and the KKT conditions explicitly due to its relevance in nonlinear programming community.

First, let us formally introduce an analog of Property $\mathcal{I}$ in the context of Lagrangian multipliers.

Definition 5.1. *For given Lagrangian multipliers λ and μ for the quadratic program $(QP)_2$, we say that they have Property $\mathcal{I}'$ if:*

(1) $\lambda > 0$ *and* $\mu > 0$*;*

(2) $H(\lambda, \mu) = Q_0 + \lambda I + \mu Q_2 \succeq 0$ *and* $\operatorname{rank}(H(\lambda, \mu)) = n - 1$*;*

(3) *The system of linear equations $H(\lambda, \mu)x = b_0 + \mu b_2$ has two solutions x_1 and x_2 satisfying $x_i^T x_i = 1$, $i = 1, 2$, and $q_2(x_1)q_2(x_2) < 0$.*

Theorem 5.2. *Suppose that $(QP)_2$ satisfies the Slater condition. Then, $(QP)_2$ has no strong duality if and only if there exist multipliers λ and μ such that Property $\mathcal{I}'$ holds.*

Proof. The Slater condition for $(QP)_2$ implies $v((SP)_2) = v((SD)_2) = v((QD)_2)$, where $(QD)_2$ denotes the dual problem of $(QP)_2$. Therefore, '$(QP)_2$ has no strong duality' is equivalent to

'$v((SP)_2) < v((QP)_2)$'. By Theorem 4.2, it is again equivalent to 'Property $\mathcal{I}$ holds'. What remains to show is that Property $\mathcal{I}$ holds if and only if the above Property $\mathcal{I}'$ holds. To put things in perspective, we restate Property $\mathcal{I}$ as follows: There exist three numbers y_0, λ, μ and two linearly independent $(n+1)$-dimensional vectors $\hat{x}_1 = [t_1, x_1^T]^T$ and $\hat{x}_2 = [t_2, x_2^T]^T$ such that

$$\begin{cases} x_1^T x_1 - t_1^2 = x_2^T x_2 - t_2^2 = 0, \\ M(q_2) \bullet \hat{x}_1 \hat{x}_1^T + M(q_2) \bullet \hat{x}_2 \hat{x}_2^T = 0, \\ (M(q_2) \bullet \hat{x}_1 \hat{x}_1^T)(M(q_2) \bullet \hat{x}_2 \hat{x}_2^T) < 0, \\ t_1^2 + t_2^2 = 1, \\ (\lambda, \mu) > 0, \\ Z := M(q_0) - y_0 I_{00} + \lambda M(q_1) + \mu M(q_2) \succeq 0, \\ \operatorname{rank}(Z) = n-1, \\ Z\hat{x}_1 = Z\hat{x}_2 = 0. \end{cases} \tag{5.1}$$

"Property $\mathcal{I} \Longrightarrow$ Property $\mathcal{I}'$":

First we note that from the 6th equation in (5.1) we may write Z as

$$Z = \begin{bmatrix} -y_0 + \lambda + \mu c_2 & -b_0^T - \mu b_2^T \\ -b_0 - \mu b_2 & Q_0 + \lambda I + \mu Q_2 \end{bmatrix}. \tag{5.2}$$

By $x_1^T x_1 - t_1^2 = x_2^T x_2 - t_2^2 = 0$ and the linear independence of $\hat{x}_1$ and $\hat{x}_2$, we have $t_1 t_2 \neq 0$. Let

$$\bar{x}_1 := x_1/t_1,\ \bar{x}_2 := x_2/t_2.$$

By (5.1), it immediately follows that $\bar{x}_1$ and $\bar{x}_2$ must satisfy the following:

$$\begin{array}{l} \|\bar{x}_1\| = \|\bar{x}_2\| = 1, \\ q_2(\bar{x}_1) q_2(\bar{x}_2) < 0, \\ (\lambda, \mu) > 0, \\ Q_0 + \lambda I + \mu Q_2 \succeq 0, \\ (Q_0 + \lambda I + \mu Q_2)\bar{x}_i = b_0 + \mu b_2,\ i = 1, 2 \\ \bar{x}_1, \bar{x}_2 \text{ are linearly indepedent.} \end{array}$$

It now remains only to check if $\operatorname{rank}(Q_0 + \lambda I + \mu Q_2) = n-1$. By using $Z\hat{x}_1$=0 and (5.2) we have

$$\begin{bmatrix} -y_0 + \lambda + \mu c_2 \\ -b_0 - \mu b_2 \end{bmatrix} = - \begin{bmatrix} -b_0^T - \mu b_2^T \\ Q_0 + \lambda I + \mu Q_2 \end{bmatrix} \bar{x}_1$$

which implies that

$$\begin{aligned} n-1 &= \operatorname{rank}(Z) = \operatorname{rank}\left(\begin{bmatrix} -y_0 + \lambda + \mu c_2 & -b_0^T - \mu b_2^T \\ -b_0 - \mu b_2 & Q_0 + \lambda I + \mu Q_2 \end{bmatrix}\right) \\ &= \operatorname{rank}\left(\begin{bmatrix} -b_0^T - \mu b_2^T \\ Q_0 + \lambda I + \mu Q_2 \end{bmatrix}\right) = \operatorname{rank}(Q_0 + \lambda I + \mu Q_2). \end{aligned}$$

"Property $\mathcal{I}'$ $\Longrightarrow$ Property $\mathcal{I}$":

Let us assume, without loss of generality, that $q_2(x_1) < 0, q_2(x_2) > 0$, and let us define

$$\begin{array}{rcl} y_0 & := & q_0(x_1) + \lambda q_1(x_1) + \mu q_2(x_1) \\ t_1 & := & \sqrt{\dfrac{-q_2(x_2)}{q_2(x_1) - q_2(x_2)}} \\ t_2 & := & \sqrt{\dfrac{q_2(x_1)}{q_2(x_1) - q_2(x_2)}} \\ \hat{x}_1 & := & t_1 \begin{bmatrix} 1 \\ x_1 \end{bmatrix} \\ \hat{x}_2 & := & t_2 \begin{bmatrix} 1 \\ x_2 \end{bmatrix} \\ Z & := & M(q_0) - y_0 I_{00} + \lambda M(q_1) + \mu M(q_2). \end{array}$$

Then, it can be straightforwardly checked that

$$\begin{array}{l} M(q_1) \bullet \hat{x}_1 \hat{x}_1^T = M(q_1) \bullet \hat{x}_2 \hat{x}_2^T = 0, \\ M(q_2) \bullet \hat{x}_1 \hat{x}_1^T + M(q_2) \bullet \hat{x}_2 \hat{x}_2^T = 0, \\ M(q_2) \bullet \hat{x}_1 \hat{x}_1^T < 0, \ M(q_2) \bullet \hat{x}_2 \hat{x}_2^T > 0, \\ t_1^2 + t_2^2 = 1, \\ (\lambda, \mu) > 0. \end{array}$$

To complete the proof, one needs only to show that $Z \succeq 0$, $Z\hat{x}_1 = Z\hat{x}_2 = 0$, and $\text{rank}\,(Z) = n - 1$.

Consider the Lagrangian function

$$L(x; \lambda, \mu) := q_0(x) + \lambda q_1(x) + \mu q_2(x),$$

whose Hessian matrix, $H(\lambda, \mu) = Q_0 + \lambda I + \mu Q_2$, is semidefinite, due to **(2)** of Property $\mathcal{I}'$. This implies that $L(x; \lambda, \mu)$ is a convex quadratic function in x. Furthermore, **(2)** and **(3)** of Property $\mathcal{I}'$ imply that the minimizers of $L(x; \lambda, \mu)$ consist of all the points on the straight line connecting x_1 and x_2. Consequently,

$$y_0 = L(x_1; \lambda, \mu) = L(x_2; \lambda, \mu) = \min_{x \in \Re^n} L(x; \lambda, \mu).$$

Consider any $(n+1)$-dimensional vector $(t, x^T)^T$ where $t \in \Re^1$ and $x \in \Re^n$. If $t = 0$ then

$$\begin{array}{rcl} \begin{bmatrix} t \\ x \end{bmatrix}^T Z \begin{bmatrix} t \\ x \end{bmatrix} & = & \begin{bmatrix} 0 \\ x \end{bmatrix}^T \begin{bmatrix} -y_0 + \lambda + \mu c_2 & -b_0^T - \mu b_2^T \\ -b_0 - \mu b_2 & Q_0 + \lambda I + \mu Q_2 \end{bmatrix} \begin{bmatrix} 0 \\ x \end{bmatrix} \\ & = & x^T (Q_0 + \lambda I + \mu Q_2) x \geq 0. \end{array}$$

Otherwise, if $t \neq 0$, then

$$\begin{aligned}
\begin{bmatrix} t \\ x \end{bmatrix}^T Z \begin{bmatrix} t \\ x \end{bmatrix} &= \begin{bmatrix} t \\ x \end{bmatrix}^T (M(q_0) - y_0 I_{00} + \lambda M(q_1) + \mu M(q_2)) \begin{bmatrix} t \\ x \end{bmatrix} \\
&= t^2 q_0(x/t) - t^2 y_0 + \lambda t^2 q_1(x/t) + \mu t^2 q_2(x/t) = t^2(L(x/t; \lambda, \mu) - y_0) \\
&\geq t^2(L(x_1; \lambda, \mu) - L(x_1; \lambda, \mu)) = 0.
\end{aligned}$$

Moreover, $\hat{x}_1^T Z \hat{x}_1 = t_1^2(L(x_1; \lambda, \mu) - y_0) = 0$ and $\hat{x}_2^T Z \hat{x}_2 = t_2^2(L(x_2; \lambda, \mu) - y_0) = t_2^2(L(x_2; \lambda, \mu) - L(x_1; \lambda, \mu)) = 0$. Therefore, $Z\hat{x}_1 = 0$ and $Z\hat{x}_2 = 0$ because $Z \succeq 0$. Since $\hat{x}_1$ and $\hat{x}_2$ are linearly independent, it follows that $\operatorname{rank}(Z) \leq n-1$. On the other hand, $\operatorname{rank}(Z) \geq \operatorname{rank}(H(\lambda, \mu)) = n-1$, leading to $\operatorname{rank}(Z) = n - 1$. □

Property $\mathcal{I}'$ is closely related to Property $\mathcal{J}$ studied in Chen and Yuan [6] for the CDT subproblem. Since Chen and Yuan [6] considered the CDT subproblem, they considered problem $(QP)_2$ with an additional condition that $Q_2 \succeq 0$. To put things in perspective, their Property $\mathcal{J}$ can be stated as:

Definition 5.3. *For given Lagrangian multipliers λ and μ for the quadratic program $(QP)_2$, we say that they have Property $\mathcal{J}$ if:*

(1) $\lambda > 0$ *and* $\mu > 0$*;*

(2) $H(\lambda, \mu) = Q_0 + \lambda I + \mu Q_2 \succeq 0$ *and* $\operatorname{rank}(H(\lambda, \mu)) = n - 1$*;*

(3) *The following 'surrogate' problem*

$$(P)_{\frac{\lambda}{\lambda+\mu}} \quad \begin{array}{ll} \textit{minimize} & q_0(x) \\ \textit{subject to} & \frac{\lambda}{\lambda+\mu} q_1(x) + \frac{\mu}{\lambda+\mu} q_2(x) \leq 0 \end{array}$$

has two solutions x_1 and x_2 satisfying

$$H(\lambda, \mu)x = b_0 + \mu b_2,$$

and $x_1^T x_1 < 1$ and $x_2^T x_2 > 1$.

The above Property $\mathcal{J}$ ([6]) is based on the idea of surrogate representation of the constraints, hence different from ours. Moreover, Chen and Yuan in ([6]) proved just only that if $(QP)_2$ with $Q_2 \succeq 0$ has no strong duality then the Property $\mathcal{J}$ holds, in other words the converse proposition can not be proved so far. However, the appearances of Property $\mathcal{J}$ and Property $\mathcal{I}'$ are quite similar indeed. Despite of this, below we shall show that they are not identical in all circumstances. Before our discussion, we shall first remark that the existence of multipliers satisfying Property $\mathcal{J}$ cannot be directly verified, while Property $\mathcal{I}$ can be checked in polynomial-time by solving a pair of SDP problems.

Proposition 5.4. *If $Q_2 \succeq 0$ then Property $\mathcal{J}$ is equivalent to Property $\mathcal{I}'$. If $Q_2 \not\succeq 0$ then Property $\mathcal{J}$ is not identical to Property $\mathcal{I}'$, the latter being a necessary and sufficient condition for $(QP)_2$ to admit a gap with its SDP relaxation.*

Proof. First consider the situation when $Q_2 \succeq 0$. We shall prove in this case that Property $\mathcal{J}$ leads to Property $\mathcal{I}'$.

Restricting the quadratic function $\frac{\lambda}{\lambda+\mu}q_1(x) + \frac{\mu}{\lambda+\mu}q_2(x)$ on the line connecting x_1 and x_2 we obtain a univariate function

$$g(t) := \tfrac{\lambda}{\lambda+\mu}q_1((1-t)x_1 + tx_2) + \tfrac{\mu}{\lambda+\mu}q_2((1-t)x_1 + tx_2), \quad t \in \Re.$$

Since $\frac{\lambda}{\lambda+\mu}q_1(x) + \frac{\mu}{\lambda+\mu}q_2(x)$ is strictly convex and quadratic, we have

$$g(t)\begin{cases} = 0, & \text{if } t = 0 \text{ and } 1; \\ < 0, & \text{if } 0 < t < 1; \\ > 0, & \text{else.} \end{cases} \tag{5.3}$$

Similarly, $h(t) := q_1((1-t)x_1+tx_2) = \|(1-t)x_1+tx_2\|^2$ is also a strictly convex quadratic function of t. Therefore, $h(0) < 0$ and $h(1) > 0$ lead to the existence of two numbers $t_1 \in (-\infty, 0)$ and $t_2 \in (0, 1)$, such that $h(t_1) = h(t_2) = 0$. Denote $x_3 = (1-t_1)x_1 + t_1x_2$ and $x_4 = (1-t_2)x_1 + t_2x_2$. Based on (5.3), we have

$$q_1(x_3) = q_1(x_4) = 0, q_2(x_3) > 0, q_2(x_4) < 0,$$

which means that Property $\mathcal{I}'$ holds.

Now consider the case where $Q_2 \not\succeq 0$. We shall prove our assertion by the following example:

$$\begin{array}{ll} \text{minimize} & q_0(x) = x_1^2 - 3x_1 \\ \text{subject to} & q_1(x) = x_1^2 + x_2^2 - 1 \le 0 \\ & q_2(x) = -x_1^2 - x_2^2 + 2x_1 \le 0 \end{array}$$

where $x = (x_1, x_2)^T$. It is easy to see that the two circles $q_1(x) = 0$ and $q_2(x) = 0$ intersect at two points, P_1 with coordinates $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and P_2 with coordinates $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$. It is easy to see that P_1 and P_2 are two unique optimal solutions for this problem, for which the corresponding multipliers are $\lambda = \mu = 1$ with the corresponding Hessian matrix of the Lagrangian function being

$$H(\lambda, \mu) = Q_0 + \lambda Q_1 + \mu Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

which is positive semidefinite with rank $n-1$ ($n = 2$). So this problem has optimal solutions with positive semidefinite Lagrangian Hessian matrices. The KKT points satisfy

$$\begin{cases} 1x_1 = \frac{1}{2} \\ 0x_2 = 0, \end{cases}$$

which lie on the line connecting P_1 and P_2. In this case, however, by Theorem 5.2 we know that Property $\mathcal{I}'$ is violated. We shall see below that Property $\mathcal{J}$ still holds nevertheless. Choose, for instance, $x^{(1)} = (\frac{1}{2}, 0)^T$ and $x^{(2)} = (\frac{1}{2}, 1)^T$, and $\lambda = \mu = 1$. We have $\frac{\lambda}{\lambda+\mu} q_1(x) + \frac{\mu}{\lambda+\mu} q_2(x) = 0$ for $x = x^{(1)}$ and $x = x^{(2)}$, and $\|x^{(1)}\| < 1$ and $\|x^{(2)}\| > 1$. After checking the conditions we see that Property $\mathcal{J}$ is indeed satisfied in this case; however, Property $\mathcal{I}'$ is violated as we have observed. □

Another related result is due to Beck and Eldar [2]. Their approach is based on a comparison between the real and the complex valued SDP relaxations. They showed that if the dimension of the null space of $H(\lambda, \mu)$ is not equal to 1, or equivalently, $\text{rank}\,(H(\lambda, \mu)) \neq n - 1$ then the SDP relaxation is tight. In the context of Theorem 5.2, this is clear, since this sufficient condition guarantees that Property $\mathcal{I}'$ does not hold and hence the SDP relaxation must be tight.

Since in Property $\mathcal{I}'$ of Theorem 5.2 the constraint $q_2(x) \leq 0$ plays a role only in the last part, the following corollary is immediate.

Consider

$$\begin{array}{lll} (Q(\rho))_2 & \text{minimize} & q_0(x) = x^T Q_0 x - 2b_0^T x \\ & \text{subject to} & q_1(x) = x^T x - 1 \leq 0 \\ & & q_2(x) - \rho \leq 0, \end{array}$$

where ρ is a parameter.

Corollary 5.5. *Suppose that Property $\mathcal{I}'$ holds for $(QP)_2$ with x_1 and x_2 being the two solutions in* **(3)** *of Property $\mathcal{I}'$ satisfying $q_2(x_1) < 0 < q_2(x_2)$. Then for any $\rho \in (q_2(x_1), q_2(x_2))$, problem $(Q(\rho))_2$ will not have a positive semidefinite Hessian for its Lagrangian function at any optimal solution.*

6 The optimal line of the dual problem

As shown in the previous sections, if Property $\mathcal{I}'$ holds for a CDT subproblem then there exists a gap between the optimal values of the primal and the dual problems. In case of Property $\mathcal{I}'$, we obtain two dual optimal solutions x_1 and x_2, one of which is feasible for the primal problem, say x_1. It can be easily proved that each point of the entire line connecting x_1 and x_2 is also an optimal solution to the dual problem. Let us call this line *the optimal line of the dual problem.* Naturally, we may wish to minimize the original quadratic function along this line to obtain a better approximate solution than x_1 for the primal problem. It is tempting to conjecture that this will always lead to an improvement. However, below we shall give an example to show that this approach may not yield a solution with any quality assurance.

Example 6.1.

$$\begin{array}{ll} \text{minimize} & q_0(x_1, x_2) = x_1(p - x_1) \\ \text{subject to} & q_1(x_1, x_2) = x_1^2 + x_2^2 \le \frac{17}{16}p^2, \\ & q_2(x_1, x_2) = (x_1 - 2p)^2 + (x_2 - p)^2 \le \frac{73}{16}p^2, \end{array}$$

where p is a positive parameter. The global optimal solution for this problem is $x^* \approx \begin{bmatrix} -0.1359p \\ 1.0218p \end{bmatrix}$, which is one of two intersection points of the circles $q_1(x_1, x_2) = 0$ and $q_2(x_1, x_2) = 0$, and the corresponding optimal value is $v^* \approx -0.1544p^2$. The system $(Q_0 + \lambda Q_1 + \mu Q_2)x = b_0 + \mu b_2$ is in this case:

$$\begin{array}{l} (-2 + 2\lambda + 2\mu)x_1 = (4\mu - 1)p \\ (2\lambda + 2\mu)x_2 = 2\mu p. \end{array}$$

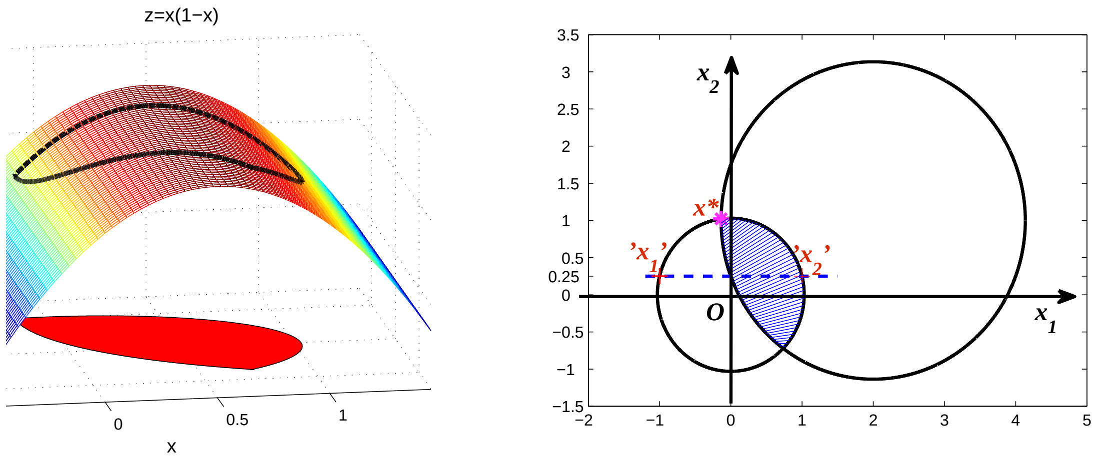

Figure 6.1: The graph of $z = x_1(1 - x_1)$(on the left) and the feasible domain (on the right) at $p = 1$.

One easily verifies that Property $\mathcal{I}'$ holds at $(\lambda, \mu) = (0.75, 0.25)$, and the solutions '$x_1$' and '$x_2$' in **(3)** of Property $\mathcal{I}'$ are $\begin{bmatrix} p \\ 0.25p \end{bmatrix}$ and $\begin{bmatrix} -p \\ 0.25p \end{bmatrix}$ respectively (see Figure 6.1). The optimal value of the SDP relaxation $(SP)_2$ is $y_0 = -0.75p^2$, and the gap between y_0 and v^* is $v^* - y_0 \approx 0.5926p^2$. The line segment that connects 'x_1' and 'x_2' and is contained in the feasible domain can be expressed by

$$\left\{ \begin{bmatrix} tp \\ 0.25p \end{bmatrix} \middle| \, 0 \le t \le 1 \right\}.$$

On this line segment, the optimal value of $q_0(x_1, x_2)$ is identically 0 for any p, which can be attained at the point 'x_2'. This shows that there cannot be any bound, in neither absolute nor relative sense of error measurements, regarding the quality of the solution obtained by the heuristic method of searching along the line segment. It remains to be a challenge to solve $(QP)_2$ efficiently, if, after solving its SDP relaxation it turns out that Property $\mathcal{I}$ indeed holds, although numerical experiments in [2] suggest that this is highly unlikely for randomly generated instances.

7 Testing Property $\mathcal{I}$ numerically

In its direct form, Property $\mathcal{I}$ requires the knowledge of an exact solution for the SDP relaxation. As is well known, in general it is impossible to solve an SDP problem exactly. It is therefore natural to test its predictive power if one uses the necessary and sufficient condition involving Property $\mathcal{I}$ in an approximative sense. In other words, if we use an ε_1-approximation solution of the SDP relaxation, then a similarly relaxed Property $\mathcal{I}$ can be verified, leading to the conclusion whether or not the original CDT subproblem satisfies the strong duality within an ε_2 error tolerance. The question is: how does the approximation work in practice?

First, we need to relax the requirement on the optimal solution. Applying an SDP solver (such as SeDuMi) to solve the SDP relaxation will return with a solution $\bar{X} \succeq 0$ and a dual solution $(\bar{Z}, \bar{y}_0, \bar{y}_1, \bar{y}_2)$ with $\bar{Z} \succeq 0$. Of course, these solutions might however violate the equality and inequality constraints of the primal-dual feasibility requirements, say by an amount no more than ε_1. Then, to purify the ranks of $\bar{X}$ and $\bar{Z}$, we may operate a spectral decomposition on $\bar{X}$ and $\bar{Z}$: $\bar{X} = Q_1^T \Lambda_1 Q_1$ and $\bar{Z} = Q_2^T \Lambda_2 Q_2$, where Q_i is orthonormal and $\Lambda_i = \mathrm{diag}(\lambda_{i1}, ..., \lambda_{in})$ with $\lambda_{ij} \geq 0$, $j = 1, ..., n$, $i = 1, 2$. Introduce

$$\hat{\lambda}_{ij} := \begin{cases} \lambda_{ij}, & \text{if } \lambda_{ij} \geq \varepsilon_2 \\ 0, & \text{if } \lambda_{ij} < \varepsilon_2, \end{cases}$$

for $j = 1, ..., n$, $i = 1, 2$, and let us purify the solutions by using $\hat{X} := Q_1^T \mathrm{diag}(\hat{\lambda}_{11}, ..., \hat{\lambda}_{1n}) Q_1$ and $\hat{Z} := Q_2^T \mathrm{diag}(\hat{\lambda}_{21}, ..., \hat{\lambda}_{2n}) Q_2$ instead of $\bar{X}$ and $\bar{Z}$, while keeping $\hat{y}_i := \bar{y}_i$, $i = 0, 1, 2$. We call $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ to be *a pair of purified $(\varepsilon_1, \varepsilon_2)$-approximate optimal solutions.*

Definition 7.1. *Suppose that $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ are a pair of purified $(\varepsilon_1, \varepsilon_2)$-approximate optimal solutions for $(SP)_2$ and $(SD)_2$ respectively. We call this pair has Property $\mathcal{I}(\varepsilon_2)$ if:*

(1) $\hat{y}_1 > \varepsilon_2$ *and* $\hat{y}_2 > \varepsilon_2$*;*

(2) $\mathrm{rank}\,(\hat{Z}) = n - 1$*;*

(3) $\mathrm{rank}\,(\hat{X}) = 2$*, and there is a rank-one decomposition of $\hat{X}$, $\hat{X} = \hat{x}_1 \hat{x}_1^T + \hat{x}_2 \hat{x}_2^T$, such that $M(q_1) \bullet \hat{x}_i \hat{x}_i^T = M(q_1) \bullet \hat{X}/2$, $i = 1, 2$, and $M(q_2) \bullet \hat{x}_1 \hat{x}_1^T < -\varepsilon_2$ and $M(q_2) \bullet \hat{x}_2 \hat{x}_2^T > \varepsilon_2$.*

Below we shall introduce a polynomial-time procedure to test the strong duality for the CDT problem, based on the ε_1-optimal SDP relaxation solution, Property $\mathcal{I}(\varepsilon_2)$, and the matrix decomposition technique.

Algorithm 7.2. *Input* ε_2*,* $M(q_0)$*,* $M(q_1)$ *and* $M(q_2)$*.*

Step 1. *Let* $\hat{X}$ *and* $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ *be the purified* $(\varepsilon_1, \varepsilon_2)$*-approximate solutions for* $(SP)_2$ *and its dual.*

Step 2. *Test whether or not Property* $\mathcal{I}(\varepsilon_2)$ *is satisfied by checking Definition 7.1, which runs in polynomial-time.*

Step 3. *If Property* $\mathcal{I}(\varepsilon_2)$ *is violated, then use the matrix decomposition technique presented in the previous sections to obtain an approximate solution to the original CDT problem; otherwise, get an approximate solution by searching along the optimal line of the dual problem.*

We now use SeDuMi to test this procedure by numerical simulations. Throughout our tests, we let $\varepsilon_2 = 10^{-4}$, and ε_1 be set as the default precision of SeDuMi. For a given positive integer n, our MATLAB code would generate two $(n+1) \times (n+1)$ matrices $M(q_0)$ and $M(q_2)$, of which the upper triangular part (including diagonal) of the entries are uniformly generated random numbers on the interval [-50, 50] (the lower part takes the values by symmetry). In order to guarantee that $(SP)_2$ have an interior feasible solution, we first solve

$$\begin{array}{ll} \text{minimize} & M(q_2) \bullet X \\ \text{subject to} & M(q_1) \bullet X \leq 0, \\ & I_{00} \bullet X = 1, \\ & X \succeq 0. \end{array}$$

Let f^* denote its optimal value. If $f^* > -10^{-4}$, we decrease the first entry (the (1,1)th position) of $M(q_2)$ by the amount $f^* + 10^{-4}$. This ensures that the Slater condition is satisfied. We apply Algorithm 7.2 on 90 randomly generated instances. The numerical results are summarized in Tables 1, 2 and 3, where 'n' denotes the dimension of the CDT problem, 'value 1' is equal to $M(q_0) \bullet \bar{X}$, i.e. the ε_1-optimal value of the SDP relaxation solution returned by SeDuMi, 'value 2' denotes the objective value of the feasible solution for the CDT problem generated by Algorithm 7.2, and 'gap' indicates the difference between 'value 1' and 'value 2' (gap = value 2 − value 1), which reflects the eventual performance of Algorithm 7.2. Finally, 'rank' indicates the rank of $\hat{X}$, and at the column '$\mathcal{I}(\varepsilon_2)$' the symbol 'V' denotes that Property $\mathcal{I}(\varepsilon_2)$ is *violated*, and 'H' signifies that Property $\mathcal{I}(\epsilon_2)$ *holds*.

Among 90 runs summarized in Tables 1 through 3, there are 87 instances violating Property $\mathcal{I}(\varepsilon_2)$ and only 3 cases holding Property $\mathcal{I}(\varepsilon_2)$. For all these 87 instances, the gaps between 'value 1' and 'value 2' are far less than the tolerance ε_2, which show that Algorithm 7.2 is indeed effective. Furthermore, the rank of the purified solution $\hat{X}$ for the 87 instances are all actually one, meaning that the eigenvector of $\hat{X}$ is the approximate optimal solution for the original CDT problem. We also made a test for two different values of the dimension: $n = 5$ and $n = 50$. Tables 2 and 3 show that it is less likely for Property $\mathcal{I}(\varepsilon_2)$ to hold for the larger n.

Table 1: Numerical results

n	value 1	value 2	gap	rank	$I(\varepsilon_2)$	n	value 1	value 2	gap	rank	$I(\varepsilon_2)$
1	31.8310	31.8310	-1.9315e-008	1	V	16	-288.2241	-288.2241	8.9968e-008	1	V
2	-61.6350	-61.6350	1.0267e-007	1	V	17	-180.2632	-180.2632	1.5133e-007	1	V
3	-92.6195	-92.6195	1.5046e-008	1	V	18	-257.0321	-257.0321	1.1875e-007	1	V
4	-64.3479	-64.3479	6.3392e-009	1	V	19	-307.8921	-307.8921	7.1101e-008	1	V
5	-76.0429	-76.0429	1.2039e-007	1	V	20	-250.2240	-250.2240	2.4064e-008	1	V
6	-148.3942	-148.3942	8.4647e-008	1	V	21	-216.6837	-216.6837	1.5005e-007	1	V
7	-149.2147	-149.2147	1.3788e-007	1	V	22	-285.2257	-285.2257	1.1723e-006	1	V
8	-165.2366	-165.2366	2.2856e-007	1	V	23	-305.7068	-305.7068	1.1012e-007	1	V
9	-146.7020	-146.7020	6.5012e-010	1	V	24	-273.7716	-273.7716	2.2697e-008	1	V
10	-193.3607	-193.3607	1.0247e-007	1	V	25	-305.1200	-305.1200	2.6449e-010	1	V
11	-194.9409	-194.9409	4.3410e-006	1	V	26	-311.0972	-311.0972	9.3392e-008	1	V
12	-131.2606	-131.2606	3.4186e-009	1	V	27	-269.2598	-269.2598	1.5854e-008	1	V
13	-174.0891	-174.0891	4.6756e-008	1	V	28	-349.2378	-349.2378	1.3295e-009	1	V
14	-215.5152	-215.5152	2.8498e-008	1	V	29	-280.3103	-280.3103	7.1443e-007	1	V
15	-232.2548	-232.2548	1.1953e-007	1	V	30	-322.0861	-322.0861	1.9794e-008	1	V

Table 2: Numerical results for $n = 5$

Inst.	value 1	value 2	gap	rank	$I(\varepsilon_2)$	Ins.	value 1	value 2	gap	rank	$I(\varepsilon_2)$
1	-72.2487	-72.2487	-9.0962e-010	1	V	16	-195.9235	-195.9235	4.2069e-008	1	V
2	-78.8733	-78.8733	3.1875e-007	1	V	17	-91.5627	-91.5627	1.7774e-009	1	V
3	-129.3945	-129.3945	2.4719e-009	1	V	18	-149.4562	-149.4562	2.0514e-007	1	V
4	-78.6061	-78.6061	3.1858e-007	1	V	19	-199.7809	-199.7809	4.9602e-010	1	V
5	-87.7781	-87.7781	4.0048e-009	1	V	20	-96.7141	-96.7141	2.0592e-006	1	V
6	-162.4757	-162.4757	3.2261e-009	1	V	21	-193.2582	-193.2582	1.0298e-006	1	V
7	-181.4192	-181.4192	1.2105e-006	1	V	22	-121.9034	-121.9034	1.3054e-009	1	V
8	-148.9920	-131.6450	17.3470	2	H	23	-132.7388	-132.7388	5.9610e-008	1	V
9	-84.6160	-84.6160	1.2004e-007	1	V	24	-221.9654	-221.9654	-7.9771e-010	1	V
10	-106.1400	-106.1400	2.6063e-007	1	V	25	-69.0899	-69.0899	9.3646e-006	1	V
11	-80.2952	-80.2952	8.1327e-010	1	V	26	-48.9339	-38.6602	10.2737	2	H
12	-93.9455	-37.5482	56.3973	2	H	27	-204.1014	-204.1014	1.4712e-007	1	V
13	-182.7852	-182.7852	7.6042e-008	1	V	28	-50.4021	-50.4021	1.8484e-006	1	V
14	-47.4945	-47.4945	3.5781e-008	1	V	29	-95.7052	-95.7052	4.4413e-008	1	V
15	-107.2132	-107.2132	7.4877e-008	1	V	30	-162.5680	-162.5680	6.1157e-009	1	V

Table 3: Numerical results for $n = 50$

ins.	value 1	value 2	gap	rank	$I(\varepsilon_2)$	Ins.	value 1	value 2	gap	rank	$I(\varepsilon_2)$
1	-329.1350	-329.1350	1.3564e-008	1	V	16	-353.2036	-353.2036	1.1395e-007	1	V
2	-418.0411	-418.0411	1.3500e-010	1	V	17	-422.6912	-422.6912	1.7024e-007	1	V
3	-334.9108	-334.9108	8.4879e-010	1	V	18	-373.7865	-373.7865	5.1733e-010	1	V
4	-314.3538	-314.3538	4.0116e-007	1	V	19	-356.4418	-356.4418	4.0084e-007	1	V
5	-406.6970	-406.6970	1.8738e-008	1	V	20	-449.4164	-449.4164	1.8588e-010	1	V
6	-376.4849	-376.4849	6.9003e-009	1	V	21	-363.3087	-363.3087	6.4648e-010	1	V
7	-436.8686	-436.8686	1.1316e-009	1	V	22	-422.4459	-422.4459	3.0531e-009	1	V
8	-456.1419	-456.1419	1.0745e-009	1	V	23	-376.0524	-376.0524	1.4611e-007	1	V
9	-420.0406	-420.0406	2.3637e-009	1	V	24	-399.0962	-399.0962	3.5397e-007	1	V
10	-443.0921	-443.0921	2.8577e-009	1	V	25	-428.4575	-428.4575	1.6672e-010	1	V
11	-398.1299	-398.1299	1.2683e-008	1	V	26	-422.2624	-422.2624	2.8901e-009	1	V
12	-381.3000	-381.3000	2.2239e-009	1	V	27	-422.8571	-422.8571	5.3685e-009	1	V
13	-400.2680	-400.2680	1.5546e-007	1	V	28	-344.5267	-344.5267	7.2918e-007	1	V
14	-337.3982	-337.3982	4.3128e-008	1	V	29	-448.3855	-448.3855	1.4571e-008	1	V
15	-433.0168	-433.0168	1.5800e-007	1	V	30	-403.9283	-403.9283	4.7542e-009	1	V

References

[1] Y. Au-Yeung and Y. Poon. *A remark on convexity and positive definiteness concerning Hermitian matrices*, Southeast Asian Bull. Math. 3 (1979), 85–92.

[2] A. Beck and Y. Eldar. *Strong Duality in Nonconvex Quadratic Optimization with Two Quadratic Constraints*, SIAM Journal on Optimization, 17 (2006), 844–860.

[3] L. Blum, M. Shub, S. Smale. *On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines.* Bull. Amer. Math. Soc. 21 (1989), 1–46.

[4] M.R. Celis, J.E. Dennis, and R.A. Tapia. A trust region algorithm for nonlinear equality constrained optimization, in R.T. Boggs, R.H. Byrd, and R.B. Schnabel eds., *Numerical Optimization*, SIAM, Philadelphia, pp. 71–82, 1985.

[5] X. Chen and Y. Yuan. *On local solutions of the Celis-Dennis-Tapia subproblem.* SIAM Journal on Optimization, 10 (2000), 359–383.

[6] X. Chen and Y. Yuan. *On maxima of dual function of the CDT subproblem*, J. Comp. Math., 19, 113–124, 2001.

[7] A.R. Conn, N.L.M. Gould and P.L. Toint. *Trust-Region Methods*, MPS/SIAM Series on Optimization, Philadelphia, 2000.

[8] P. Finsler. *Über das Vorkommen definiter und semidefiniteter Formen in Scharen quadratischer Formen.* Comment. Math. Helv., 9 (1937), 188–192.

[9] Y.W. Huang and S. Zhang. *Complex matrix decomposition and quadratic programming*, Mathematics of Operations Research, 32 (2007), 758–768.

[10] J.M. Martinez. *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optm., 4 (1994), 159–176.

[11] J. Peng and Y. Yuan. *Optimality conditions for the minimization of a quadratic with two quadratic constraints*, SIAM Journal on Optimization, 7 (1997), 579–594.

[12] J.F. Sturm and S. Zhang. *On cones of nonnegative quadratic functions*, Mathematics of Operations Research, 28 (2003), 246–267.

[13] H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds. *Handbook on Semidefinite Programming: Theory, Algorithms, and Applications*. Kluwer Academic Publishers, Dordrecht, 2000.

[14] Y. Ye and S. Zhang. *New results on quadratic minization*, SIAM Journal on Optimization, 28 (2003), 246–267.

[15] Y. Yuan. *On a subproblem of trust region algorithms for constrained optimization*, Mathematical Programming, 47 (1990), 53–63.

[16] Y. Yuan. *A dual algorithm for minimizing a quadratic function with two quadratic constraints*, J. Comp. Math., 9 (1991), 348–359.

[17] Y. Zhang. *Computing a Celis-Dennis-Tapia trust-region step for equality constrained optimization*, Mathematical Programming, 55 (1992),109–124.