MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

A Polynomial-Time Algorithm for Nonconvex Quadratic Optimization with Two Quadratic Constraints

Shinsaku SAKAUE, Yuji NAKATSUKASA,
Akiko TAKEDA and Satoru IWATA

METR 2015–03 January 2015

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html

A POLYNOMIAL-TIME ALGORITHM FOR NONCONVEX QUADRATIC OPTIMIZATION WITH TWO QUADRATIC CONSTRAINTS *

SHINSAKU SAKAUE , YUJI NAKATSUKASA , AKIKO TAKEDA , AND SATORU IWATA†

Abstract. We consider solving a nonconvex quadratic minimization problem with two quadratic constraints (2QCQP), one of which being convex. This problem includes the Celis-Denis-Tapia (CDT) problem as a special case. The CDT problem has been widely studied, but no polynomial-time algorithm was known until Bienstock's recent work. His algorithm solves the CDT problem in polynomial time with respect to the number of bits in data and $\log \epsilon^{-1}$ by admitting an ϵ error in the constraints. The algorithm, however, appears to be difficult to implement.

In this paper, we present a polynomial-time algorithm to solve 2QCQP exactly. Our algorithm is based on the one proposed by Iwata, Nakatsukasa and Takeda (2014) for computing the signed distance between overlapping ellipsoids. Our algorithm computes all the Lagrange multipliers of 2QCQP by solving a two-parameter linear eigenvalue problem, obtains the corresponding KKT points, and finds a global solution as the KKT point with the smallest objective value. The computational complexity of the algorithm is $O(n^6)$, where n is the number of variables.

Key words. quadratically constrained quadratic programming, nonconvex optimization, Celis-Dennis-Tapia problem, two-parameter eigenvalue problem

AMS subject classifications. 49M37, 65K05, 90C25, 90C30

1. Introduction.

In this paper, we consider solving the quadratic minimization problem with two quadratic constraints (2QCQP):

$$\underset{x}{\text{minimize}} \quad f(x) = x^\top Q_0 x + 2q_0^\top x + \gamma_0 \tag{1.1}$$

$$\text{subject to} \quad g_i(x) = x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2) \tag{1.2}$$

where $Q_i \in \mathbb{R}^{n\times n}$ is symmetric, $q_i \in \mathbb{R}^n$, and $\gamma_i \in \mathbb{R}$ for each $i = 0, 1, 2$. We assume that Q_1 is positive definite; we also make other minor assumptions as summarized in Section 2.1.

2QCQP includes the Celis-Dennis-Tapia (CDT) problem, which minimizes a nonconvex quadratic function over the intersection of two ellipsoids, as a special case where Q_1 is positive definite and Q_2 is positive semidefinite. The CDT problem was proposed by Celis, Dennis and Tapia [8] as a natural extension of the trust region subproblem (TRS), which has only one ellipsoidal constraint. Though TRS is nonconvex since Q_0 in the objective function is indefinite, its Lagrangian dual gives an exact semidefinite programming (SDP) reformulation of TRS [21]; an optimal solution for TRS can be obtained from an optimal solution of the polynomial-time solvable SDP problem. Moreover, the polynomial solvability property is extended by Sturm and Zhang [22] to the case of a single nonconvex quadratic constraint by proving that the Lagrangian dual of a quadratic minimization problem with one quadratic constraint is also tight.

The additional constraint makes the CDT problem substantially more challenging than TRS. The CDT problem can have a duality gap in general (see, e.g., [19]). Ai and Zhang [1] derived easily verifiable conditions to characterize when the CDT

*Version of January 26, 2015. This research is supported by Grants-in-Aid for Scientific Research (Challenging Exploring Research, No. 26540007) from Japan Society for Promotion of Science.

†Department of Mathematical Informatics, The University of Tokyo, Tokyo 113-8656, Japan ({shinsaku_sakaue,nakatsukasa,takeda,iwata}mist.i.u-tokyo.ac.jp).

problem has no duality gap, which is equivalent to when the SDP relaxation of the CDT problem is tight since the Lagrangian dual problem coincides with the SDP relaxation (see [18]). They also proved that the SDP relaxation is tight if and only if the Hessian of the Lagrangian is positive semidefinite at a global solution: a fact that we use in Section 6 to analyze numerical results. In addition, various properties of the CDT problem have been studied, e.g., necessary and sufficient conditions for the optimality of the CDT problem [6, 19] and the location of Lagrangian multipliers corresponding to a local minimizer [9].

Li and Yuan [16] proposed an algorithm that finds a global solution for the CDT problem with no duality gap, i.e., if the Hessian of the Lagrangian is positive semidefinite at a global solution. As Yuan [25] proved, however, it is possible that the Hessian of the Lagrangian in the CDT problem has one negative eigenvalue at a global solution, which means that Li and Yuan's algorithm does not always find a global solution. Burer and Anstreicher [7] provided a tighter relaxation problem by adding second-order cone constraints to the usual SDP relaxation, but the resulting problem still has a relaxation gap. Yang and Burer [24] reformulated the special case of the CDT problem with two variables into an exact SDP formulation by adding valid constraints. In general, however, the complexity of the CDT problem had been open for a long time until Bienstock [5] recently proved its polynomial-time solvability.

Bienstock's proof in fact provides a polynomial-time algorithm for general quadratic optimization problems with an arbitrary fixed number of quadratic constraints. His algorithm makes a sequence of calls to a polynomial-time feasibility algorithm based on Barvinok's construction [3]; the length of the sequence is polynomial in the number of bits in the data and $\log \epsilon^{-1}$. The algorithm returns an ϵ-feasible solution, that is, a solution guaranteed to satisfy the relaxed constraints: $x^\top Q_i x + 2q_i^\top x + \gamma_i \le \epsilon$. Unfortunately, however, Bienstock's polynomial-time algorithm does not appear to be very practical, because the polynomial-time feasibility algorithm looks difficult to implement.

In this paper, we provide a polynomial-time algorithm for a generalization, 2QCQP, of the CDT problem. To the best of our knowledge, no polynomial-time algorithm has been implemented and used to solve large-scale problems. An efficient CDT algorithm also provides an efficient algorithm for equality constrained optimization, since solving a sequence of CDT problems is required in the Powell-Yuan trust-region algorithm [20] for equality constrained optimization.

Our algorithm is based on the one developed in [13] for computing the signed distance between overlapping ellipsoids via solving a special case of 2QCQP. We generalize the algorithm to solve 2QCQP. The approach is to find the Lagrange multipliers of 2QCQP from the Karush-Kuhn-Tucker (KKT) conditions. The KKT conditions of 2QCQP result in rational equations of Lagrange multipliers. We convert the rational equations into polynomial equations by constructing certain bivariate matrix pencils whose zeros of determinants are the zeros of the rational equations. This reduces the problem to a two-parameter linear eigenvalue problem, which can be solved via a single-parameter linear eigenvalue problem of large (squared) size, for which reliable algorithms are available. As we shall see, in nongeneric cases our algorithm encounters singular matrix pencils, and we also discuss how to handle such issues. The overall computational complexity of our algorithm for 2QCQP is $O(n^6)$.

This paper is organized as follows. In Section 2, we derive the KKT conditions of 2QCQP and express them as two generalized eigenvalue problems and a two-parameter linear eigenvalue problem with certain polynomial matrix pencils, whose solutions include the Lagrange multipliers. Section 3 discusses the solution method of the two-parameter eigenvalue problem and show that our algorithm works in generic cases. In Section 4, we analyze the case in which our algorithm faces difficulty and describe how to handle such a case by employing certain preprocessing techniques. In Section 5, we summarize our algorithm and show that our algorithm solves 2QCQP in $O(n^6)$ time. Finally, in Section 6, we present numerical experiments to demonstrate the practical performance of our algorithm.

Notation. Throughout this paper, we denote a zero vector in $\mathbb{R}^k$ by $\mathbf{0}_k$, or just by 0 when the dimension is clear. The unit matrix of size k is denoted by I_k. For a pair of symmetric matrices X and Y, we write $X \succ Y$ if $X - Y$ is positive definite and $X \succeq Y$ if $X - Y$ is positive semidefinite.

2. Finding the KKT points.

2.1. Assumptions.

We impose the following two assumptions throughout this paper.

ASSUMPTION 2.1. *The symmetric matrix Q_1 is positive definite.*

Assumption 2.1 enables us to check the feasibility of 2QCQP. Since $Q_1 \succ O$, Q_1 is nonsingular and the constraint $g_1(x) \leq 0$ can be written as

$$(x + Q_1^{-1}q_1)^\top Q_1(x + Q_1^{-1}q_1) \leq q_1^\top Q_1^{-1}q_1 - \gamma_1.$$

Therefore, noting $Q_1 \succ O$, 2QCQP is infeasible if $q_1^\top Q_1^{-1}q_1 - \gamma_1 < 0$. If $q_1^\top Q_1^{-1}q_1 - \gamma_1 = 0$, the only possible solution is $x = -Q_1^{-1}q_1$ and we check whether $g_2(-Q_1^{-1}q_1) \leq 0$ holds or not. If $q_1^\top Q_1^{-1}q_1 - \gamma_1 > 0$, we see the constraint $g_1(x) \leq 0$ is strictly feasible (i.e., $\exists \hat{x}$ satisfying $g_1(\hat{x}) < 0$). We then check the feasibility of 2QCQP by solving the Lagrangian dual of

$$\begin{aligned} &\underset{x}{\text{minimize}} && g_2(x) \\ &\text{subject to} && g_1(x) \leq 0, \end{aligned} \tag{2.1}$$

which can be formulated as the following SDP with dual variables μ_1, μ_2:

$$\begin{aligned} &\underset{\mu_1,\mu_2}{\text{maximize}} && \mu_2 \\ &\text{subject to} && \mu_1 \begin{bmatrix} \gamma_1 & q_1^\top \\ q_1 & Q_1 \end{bmatrix} - \mu_2 \begin{bmatrix} 1 & \mathbf{0}_n^\top \\ \mathbf{0}_n & O \end{bmatrix} \succeq - \begin{bmatrix} \gamma_2 & q_2^\top \\ q_2 & Q_2 \end{bmatrix} \\ &&& \mu_1 \geq 0. \end{aligned} \tag{2.2}$$

Under the condition that $g_1(x) \leq 0$ is strictly feasible, there is no duality gap (see [22]) and the optimal value of (2.1) can be obtained by solving its dual problem, SDP (2.2). If the optimal value is positive, then the original problem 2QCQP is infeasible. In this way, we could check whether 2QCQP is feasible or not if $Q_1 \succ O$. Furthermore, as we see above, if $g_1(x) \leq 0$ is not strictly feasible, 2QCQP is trivial, i.e., the only possible solution is $x = -Q_1^{-1}q_1$. Therefore, throughout the following discussion, we consider 2QCQP such that the constraint $g_1(x) \leq 0$ is strictly feasible. Assumption 2.1 also ensures that the optimal value of a feasible 2QCQP is bounded by the Weierstrass extreme value theorem. In Appendix A we derive an explicit lower bound of the optimal value. Furthermore, we use the nonsingularity of Q_1 in theoretical analysis of our proposed algorithm.

ASSUMPTION 2.2. *The linear independence constraint qualification (LICQ) holds on the intersection of the boundaries of two constraints, i.e.,* $\nabla g_1(\bar{x})$ *and* $\nabla g_2(\bar{x})$ *are linearly independent for every* $\bar{x}$ *satisfying* $g_1(\bar{x}) = g_2(\bar{x}) = 0$.

Assumption 2.2 requires that if both constraints are active at $\bar{x}$, the gradients $\nabla g_1(\bar{x}) = 2(Q_1\bar{x} + q_1)$ and $\nabla g_2(\bar{x}) = 2(Q_2\bar{x} + q_2)$ are linearly independent. Some constraint qualification (CQ) such as LICQ or Mangasarian-Fromovitz CQ is necessary to ensure that the Karush-Kuhn-Tucher (KKT) conditions are necessary optimality conditions for 2QCQP. The LICQ is known to be the weakest CQ that guarantees the existence and uniqueness of Lagrange multipliers (see [23]). Nonetheless, this assumption can be relaxed if necessary: we discuss later in Appendix B how to find local solutions that violate Assumption 2.2, i.e., the LICQ, by using the Karush-John (sometimes called the Fritz John) optimality conditions.

2.2. The KKT conditions for the problem. As discussed above, we focus on the case where 2QCQP has an optimal solution. Under Assumption 2.2, if $x \in \mathbb{R}^n$ is a local solution of 2QCQP, then there exists a pair of Lagrange multipliers $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ satisfying the KKT conditions:

$$H(\lambda_1, \lambda_2)x = y, \tag{2.3}$$

$$x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2), \tag{2.4}$$

$$\lambda_i \left(x^\top Q_i x + 2q_i^\top x + \gamma_i\right) = 0 \quad (i = 1, 2), \tag{2.5}$$

$$\lambda_i \ge 0 \quad (i = 1, 2), \tag{2.6}$$

where

$$H(\lambda_1, \lambda_2) := Q_0 + \lambda_1 Q_1 + \lambda_2 Q_2 \tag{2.7}$$

and

$$y := -(q_0 + \lambda_1 q_1 + \lambda_2 q_2). \tag{2.8}$$

We note that just like $H(\lambda_1, \lambda_2)$, the vector y depends on λ_1 and λ_2, but for notational simplicity we just write y in what follows. We also remark that the matrix $H(\lambda_1, \lambda_2)$ is the Hessian of the Lagrangian.

2.3. Formulation as a pair of bivariate matrix equations. The variable x satisfying (2.3) can be expressed in terms of λ_1, λ_2. By substituting such x into (2.5), we obtain two bivariate rational equations with respect to λ_1, λ_2 of the form

$$\lambda_i y^\top H(\lambda_1, \lambda_2)^{-1} Q_i H(\lambda_1, \lambda_2)^{-1} y + 2q_i^\top H(\lambda_1, \lambda_2)^{-1} y + \gamma_i = 0, \quad i = 1, 2. \tag{2.9}$$

These can be reduced to a pair of bivariate polynomial equations if the numerator and denominator polynomials in (2.9) are known explicitly, but since this is not the case, solving (2.9) for λ_1, λ_2 is challenging. Instead, we formulate a pair of matrix equations that provide appropriate multipliers: we introduce a pair of matrices $M_1(\lambda_1, \lambda_2)$ and $M_2(\lambda_1, \lambda_2)$ defined by

$$M_i(\lambda_1, \lambda_2) := \begin{bmatrix} Q_i & -H(\lambda_1, \lambda_2) & q_i \\ -H(\lambda_1, \lambda_2) & O & y \\ q_i^\top & y^\top & \gamma_i \end{bmatrix} \quad (i = 1, 2). \tag{2.10}$$

LEMMA 2.1. *For every x that satisfies the KKT conditions* (2.3)–(2.6) *with Lagrange multipliers λ_1 and λ_2, we have $\lambda_i \det M_i(\lambda_1, \lambda_2) = 0$ $(i = 1, 2)$.*

Proof. By equation (2.3), y must belong to $\operatorname{Im} H(\lambda_1, \lambda_2)$. Therefore, if $H(\lambda_1, \lambda_2)$ is singular, we have rank $\begin{bmatrix} -H(\lambda_1, \lambda_2) & y \end{bmatrix} < n$, which implies that $M_i(\lambda_1, \lambda_2)$ is singular. Therefore, we obtain $\lambda_i \det M_i(\lambda_1, \lambda_2) = 0$.

Now suppose that $H(\lambda_1, \lambda_2)$ is nonsingular. For the computation of $\det M_i(\lambda_1, \lambda_2)$, we use the Schur complement of $M_i(\lambda_1, \lambda_2)$ with respect to

$$A_i := \begin{bmatrix} Q_i & -H(\lambda_1, \lambda_2) \\ -H(\lambda_1, \lambda_2) & O \end{bmatrix}.$$

Since

$$A_i^{-1} = \begin{bmatrix} O & -H(\lambda_1, \lambda_2)^{-1} \\ -H(\lambda_1, \lambda_2)^{-1} & -H(\lambda_1, \lambda_2)^{-1} Q_i H(\lambda_1, \lambda_2)^{-1} \end{bmatrix},$$

we have

$$\begin{aligned} \det M_i(\lambda_1, \lambda_2) = (-1)^n \det H(\lambda_1, \lambda_2)^2 \times \\ \left(\gamma_i + q_i^\top H(\lambda_1, \lambda_2)^{-1} y + y^\top H(\lambda_1, \lambda_2)^{-1} Q_i H(\lambda_1, \lambda_2)^{-1} y\right). \end{aligned}$$

Thus, using (2.3) for the above equation, we obtain

$$\det M_i(\lambda_1, \lambda_2) = (-1)^n \det H(\lambda_1, \lambda_2)^2 \left(x^\top Q_i x + 2 q_i^\top x + \gamma_i\right). \tag{2.11}$$

It then follows from (2.5) that $\lambda_i \det M_i(\lambda_1, \lambda_2) = 0$. ∎

Lemma 2.1 suggests computing all possible pairs of Lagrange multipliers λ_1 and λ_2 for the KKT points by solving the bivariate determinantal equations

$$\lambda_1 \det M_1(\lambda_1, \lambda_2) = \lambda_2 \det M_2(\lambda_1, \lambda_2) = 0. \tag{2.12}$$

We will discuss how to solve (2.12) in Sections 2.4 and 3. Note that not all solutions (λ_1, λ_2) for (2.12) are Lagrange multipliers for our 2QCQP, but as long as the number of solutions is finite, we can find the 2QCQP solution via checking the feasibility and comparing the objective values.

For each pair of nonnegative multipliers λ_1 and λ_2 thus obtained, one can compute x by solving the linear equation (2.3). If $H(\lambda_1, \lambda_2)$ is nonsingular, then the vectors x are uniquely determined, and they naturally satisfy the feasibility conditions (1.2). If $H(\lambda_1, \lambda_2)$ is singular, we select a solution that satisfies (1.2) among all solutions of (2.3) and verify that (2.5) holds. Specifically, let $H_0 \in \mathbb{R}^{n \times r}$ be a basis for the null space of $H(\lambda_1, \lambda_2)$ with rank $(H_0) = r$ and v be an arbitrary vector in $\mathbb{R}^r$. Then, solutions of (2.3) can be written as $x_* + H_0 v$ where x_* is any vector satisfying (2.3); for example the least-squares solution. If $x_* + H_0 v$ satisfies the KKT conditions, the objective function can be written without v as follows:

$$\begin{aligned} f(x_* + H_0 v) &= f(x_* + H_0 v) + \lambda_1 g_1(x_* + H_0 v) + \lambda_2 g_2(x_* + H_0 v) \\ &= (x_* + H_0 v)^\top H(\lambda_1, \lambda_2)(x_* + H_0 v) - 2y^\top (x_* + H_0 v) + \gamma_0 + \gamma_1 + \gamma_2. \\ &= x_*^\top H(\lambda_1, \lambda_2) x_* + x_*^\top H(\lambda_1, \lambda_2) H_0 v - 2y^\top (x_* + H_0 v) + \gamma_0 + \gamma_1 + \gamma_2. \\ &= x_*^\top H(\lambda_1, \lambda_2) x_* - 2y^\top x_* + 2(x_*^\top H(\lambda_1, \lambda_2) - y^\top) H_0 v + \gamma_0 + \gamma_1 + \gamma_2 \\ &= x_*^\top H(\lambda_1, \lambda_2) x_* - 2y^\top x_* + \gamma_0 + \gamma_1 + \gamma_2. \end{aligned}$$

This means that the objective function values are the same for all v such that $x_* + H_0 v$ satisfies the KKT conditions. Therefore, by selecting such v, we obtain one of the global solutions. We discuss how to obtain such v in Appendix C.

2.4. Three cases of (λ_1, λ_2) for the determinantal equations. We rewrite $M_i(\lambda_1, \lambda_2)$ defined by (2.10) in the following matrix polynomial form:

$$M_i(\lambda_1, \lambda_2) = C_i + \lambda_1 D_1 + \lambda_2 D_2 \tag{2.13}$$

where

$$C_i := \begin{bmatrix} Q_i & -Q_0 & q_i \\ -Q_0 & O & -q_0 \\ q_i^\top & -q_0^\top & \gamma_i \end{bmatrix}, \quad D_i := \begin{bmatrix} O & -Q_i & \mathbf{0}_n \\ -Q_i & O & -q_i \\ \mathbf{0}_n^\top & -q_i^\top & 0 \end{bmatrix} \quad (i = 1, 2). \tag{2.14}$$

To obtain all the solutions of (2.12), we now separately consider three cases of (λ_1, λ_2) depending on whether λ_1, λ_2 are zero or not.

1. The pair of zero multipliers $(\lambda_1, \lambda_2) = (0, 0)$ satisfies (2.12). Therefore, we have $(\lambda_1, \lambda_2) = (0, 0)$ as one of the solutions of (2.12).
2. Exactly one of λ_1 or λ_2 is nonzero. If $\lambda_1 \neq 0$ and $\lambda_2 = 0$, (2.12) can be written as

   $$\det M_1(\lambda_1, 0) = \det(C_1 + \lambda_1 D_1) = 0. \tag{2.15}$$

   This can be solved for λ_1 as a linear generalized eigenvalue problem. Similarly, if $\lambda_1 = 0$ and $\lambda_2 \neq 0$, we obtain the values of λ_2 corresponding to $\lambda_1 = 0$ by solving

   $$\det M_2(0, \lambda_2) = \det(C_2 + \lambda_2 D_2) = 0. \tag{2.16}$$

   In some rare cases, $M_1(\lambda_1, 0)$ or $M_2(0, \lambda_2)$ is a singular matrix pencil (e.g., $\det M_1(\lambda_1, 0) = 0$ for all λ_1) and (2.15) or (2.16) has infinitely many solutions. We deal with this case by slightly perturbing some of the matrices so that the matrix pencils $M_1(\lambda_1, 0)$ and $M_2(0, \lambda_2)$ become regular. Details are described in Section 4.
3. $\lambda_1 \lambda_2 \neq 0$. Then (2.12) is equivalent to the bivariate determinantal equations expressed as

   $$\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0. \tag{2.17}$$

   We will discuss in detail how to solve (2.17) in Section 3.

3. Solving the bivariate determinantal equations. From (2.13), we see that (2.17) is a two-parameter eigenvalue problem expressed as

$$\det(C_1 + \lambda_1 D_1 + \lambda_2 D_2) = 0, \tag{3.1}$$

$$\det(C_2 + \lambda_1 D_1 + \lambda_2 D_2) = 0. \tag{3.2}$$

We now discuss how to solve this system of equations for λ_1 and λ_2.

3.1. Reduction to univariate linear eigenvalue problems. The $(2n+1) \times (2n+1)$ two-parameter eigenvalue problem (3.1), (3.2) can be solved via the following $(2n+1)^2 \times (2n+1)^2$ linear generalized eigenvalue problems:

$$\det B(\lambda_1) = \det\left((D_2 \otimes C_1 - C_2 \otimes D_2) + \lambda_1 (D_2 \otimes D_1 - D_1 \otimes D_2)\right) = 0, \tag{3.3}$$

$$\det B(\lambda_2) = \det\left((C_1 \otimes D_1 - D_1 \otimes C_2) + \lambda_2 (D_2 \otimes D_1 - D_1 \otimes D_2)\right) = 0. \tag{3.4}$$

This is a standard process of solving two-parameter eigenvalue problems [2]. As we discuss later, the matrices $B(\lambda_1), B(\lambda_2)$ are the Bézout matrices [15].

Suppose v_1 and v_2 are nonzero eigenvectors of $M_1(\lambda_1,\lambda_2)$ and $M_2(\lambda_1,\lambda_2)$ at the eigenvalue (λ_1,λ_2) respectively, i.e., $M_1(\lambda_1,\lambda_2)v_1 = M_2(\lambda_1,\lambda_2)v_2 = 0$, then $v_1 \otimes v_2$ is a common eigenvector of $B(\lambda_1)$ and $B(\lambda_2)$ (see, e.g., [2]). Therefore, if the matrix pencil $B(\lambda_2)$ is regular, we obtain the solutions (λ_1,λ_2) of (2.17) as follows: we solve (3.4) to obtain all candidates of λ_2 satisfying (2.17), which has finitely many solutions λ_2 if $B(\lambda_2)$ is regular. Then we solve the pair of ordinary eigenvalue problems $\det M_1(\lambda_1,\hat{\lambda}_2) = \det M_2(\lambda_1,\hat{\lambda}_2) = 0$ among the positive $\hat{\lambda}_2$ thus obtained to get the corresponding λ_1, if any. If these determinantal equations have a common solution λ_1, we have (λ_1,λ_2) as a solution of (2.17) (alternatively, we can start from finding λ_1 by solving (3.3)). There is a minor issue here: it turns out that $B(\lambda_2)$ has null space independent of the value of λ_2 and therefore (3.4) has infinitely many solutions. In Section 3.3 we discuss how to overcome this issue by removing the null space of $B(\lambda_2)$.

3.2. Connections with Bézoutians. We now mention a connection of the above process to Bézoutians, which we also use later. In fact, forming $B(\lambda_1), B(\lambda_2)$ in (3.3), (3.4) from M_1, M_2 is equivalent to taking the Bézoutian for the two matrix polynomials $I_{2n+1}\otimes M_1(\lambda_1,\lambda_2)$ and $M_2(\lambda_1,\lambda_2)\otimes I_{2n+1}$, which are of size $(2n+1)^2 \times (2n+1)^2$. Here the Kronecker products are taken to achieve commutativity, which facilitate the formulation of the Bézoutian for matrix polynomials [15].

Two matrix polynomials P_1 and P_2 are said to commute if $P_1(\xi)P_2(\xi) = P_2(\xi)P_1(\xi)$ holds for every value of ξ. The Bézoutian for commuting regular matrix polynomials P_1, P_2 of size $n\times n$ and degree k is defined by the bivariate matrix polynomial

$$\mathcal{B}(s,t) = \frac{P_1(s)P_2(t) - P_2(s)P_1(t)}{s-t} = \sum_{i,j=0}^{k-1,k-1} B_{ij}s^it^j \tag{3.5}$$

in s and t. Here B_{ij} is the $n\times n$ coefficient matrix corresponding to the term s^it^j in $\mathcal{B}(s,t)$. Then the block symmetric $nk\times nk$ matrix

$$B = \begin{bmatrix} B_{0,0} & \cdots & B_{0,k-1} \\ \vdots & \ddots & \vdots \\ B_{k-1,0} & \cdots & B_{k-1,k-1} \end{bmatrix}$$

is called the *Bézout matrix*.

LEMMA 3.1 ([15, Theorem 1.1]). *Suppose that P_1 and P_2 are regular matrix polynomials, i.e., $\det P_1(\xi_1) \neq 0$ and $\det P_2(\xi_2) \neq 0$ for some ξ_1 and ξ_2. Then, the Bézout matrix B is singular if and only if P_1 and P_2 share an eigenpair (ξ, v), i.e., a scalar ξ and a vector $v \neq 0$ such that $P_1(\xi)v = P_2(\xi)v = 0$.*

More generally, the null space of the Bézoutian is related to the so-called common restriction [10, 15] (this fact is not needed for what follows).

According to the definition, for any fixed λ_2, the Bézout matrix between $P_1(\lambda_1) := M_2(\lambda_1,\lambda_2)\otimes I_{2n+1}$ and $P_2(\lambda_1) := I_{2n+1}\otimes M_1(\lambda_1,\lambda_2)$ can be written as

$$(C_1 + \lambda_2 D_2)\otimes D_1 - D_1\otimes(C_2+\lambda_2 D_2),$$

which is equivalent to $B(\lambda_2)$ in (3.4). Lemma 3.1 suggests that we can find the λ_2-values for the solution of $\det M_1(\lambda_1,\lambda_2) = \det M_2(\lambda_1,\lambda_2) = 0$ by computing the values of λ_2 for which $\det B(\lambda_2) = 0$. The discussion for $B(\lambda_1)$ is completely analogous.

3.3. Removing the null space of the Bézout matrix. As discussed above, the solutions of (2.17) can be obtained via solving $\det B(\lambda_2) = 0$ if $B(\lambda_2)$ is a regular matrix pencil. However, as we show below, $B(\lambda_2)$ has nonempty null space independent of the value of λ_2. We now describe how to remove the null space of $B(\lambda_2)$ to obtain a regular matrix pencil so that the number of solutions computed from $\det B(\lambda_2) = 0$ is finite.

Since Q_1 is positive definite by Assumption 2.1, Q_1 is nonsingular. Then we see that D_1 has the null vector

$$v = \begin{bmatrix} Q_1^{-1}q_1 & \mathbf{0}_n^\top & -1 \end{bmatrix}^\top \tag{3.6}$$

since

$$D_1 v = \begin{bmatrix} O & -Q_1 & \mathbf{0}_n \\ -Q_1 & O & -q_1 \\ \mathbf{0}_n^\top & -q_1^\top & 0 \end{bmatrix} \begin{bmatrix} Q_1^{-1}q_1 \\ \mathbf{0}_n \\ -1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n \\ \mathbf{0}_n \\ 0 \end{bmatrix}. \tag{3.7}$$

Therefore, for every fixed λ_2, both $M_1(\infty, \lambda_2)$ and $M_2(\infty, \lambda_2)$ have v as a null vector. This means the Bézout matrix $B(\lambda_2)$ has a null vector

$$w := \frac{v \otimes v}{\|v\|^2} \tag{3.8}$$

for every λ_2. In order to obtain a regular matrix pencil from $B(\lambda_2)$ that retains the relevant information, we "project out" the null vector.

First, we show that the null vector v is not the common eigenvector of M_1 and M_2. This means the projection process described later does not spoil the solvability of (2.17).

LEMMA 3.2. *The vector v in* (3.6) *is not a common eigenvector of M_1 and M_2, i.e., $M_1(\lambda_1, \lambda_2)v = M_2(\lambda_1, \lambda_2)v = 0$ does not hold for any finite pair (λ_1, λ_2).*

Proof. Recall that we only need to consider the case where the constraint $g_1(x) \le 0$ is strictly feasible in addition to $Q_1 \succ O$ by Assumption 2.1. We now suppose to the contrary that $M_1 v = M_2 v = 0$ holds with eigenvalues (λ_1, λ_2). Recalling (2.13) and using $D_1 v = 0$, we express these equations as

$$\begin{aligned} &(C_1 + \lambda_2 D_2)v = (C_2 + \lambda_2 D_2)v = 0 \\ \iff &\begin{cases} Q_2 Q_1^{-1} q_1 - q_2 = 0, \\ Q_0 Q_1^{-1} q_1 - q_0 + \lambda_2 (Q_2 Q_1^{-1} q_1 - q_2) = 0, \\ q_1^\top Q_1^{-1} q_1 - \gamma_1 = q_2^\top Q_1^{-1} q_1 - \gamma_2 = 0 \end{cases} \\ \iff &\begin{cases} Q_0 Q_1^{-1} q_1 - q_0 = Q_2 Q_1^{-1} q_1 - q_2 = 0, \\ q_1^\top Q_1^{-1} q_1 - \gamma_1 = q_2^\top Q_1^{-1} q_1 - \gamma_2 = 0. \end{cases} \end{aligned}$$

Using the third equality we obtain

$$\begin{aligned} g_1(x) \le 0 &\iff (x + Q_1^{-1}q_1)^\top Q_1 (x + Q_1^{-1}q_1) - q_1^\top Q_1^{-1} q_1 + \gamma_1 \le 0 \\ &\iff (x + Q_1^{-1}q_1)^\top Q_1 (x + Q_1^{-1}q_1) \le 0. \end{aligned}$$

Since $Q_1 \succ O$, the feasible region of $g_1(x) \le 0$ is a singleton $x = -Q_1^{-1}q_1$, which contradicts the strict feasibility of $g_1(x) \le 0$. ∎

We now consider how to project out the null vector v. Mathematically, the projection is done as follows: form a square orthogonal matrix $[w, W]$ where $w = (v \otimes v)/\|v\|^2$, and define the projected Bézout matrix

$$\widetilde{B}(\lambda_2) := W^\top B(\lambda_2) W. \tag{3.9}$$

Now we show the solutions of (2.17) with $|\lambda_1| < \infty$ satisfies

$$\det \widetilde{B}(\lambda_2) = 0. \tag{3.10}$$

This means we obtain all solutions of (2.17) via solving (3.10).

LEMMA 3.3. *Suppose that λ_2 is a solution of* (2.17) *with some corresponding finite λ_1 and nonzero eigenvectors v_1, v_2, i.e., $M_1(\lambda_1, \lambda_2)v_1 = M_2(\lambda_1, \lambda_2)v_2 = 0$. Then, the projected Bézout matrix $\widetilde{B}(\lambda_2)$ in* (3.9) *is singular.*

Proof. We first show that $B(\lambda_2)$ has the following null vector $u \in \mathbb{R}^{(2n+1)^2}$:

$$u = v_1 \otimes v_2 - ((v_1 \otimes v_2)^\top w)w,$$

where $w = (v \otimes v)/\|v\|^2$ as in (3.8). Since v as in (3.6) is not a common eigenvector of M_1, M_2 by Lemma 3.2, we see that $v_1 \otimes v_2$ is linearly independent of w, thus $\|u\|_2 \neq 0$. By multiplying u to $B(\lambda_2)$, we have

$$\begin{aligned}
B(\lambda_2)u &= B(\lambda_2)(v_1 \otimes v_2) - ((v_1 \otimes v_2)^\top w)B(\lambda_2)w \\
&= -((v_1 \otimes v_2)^\top w)B(\lambda_2)w \\
&= -((v_1 \otimes v_2)^\top w)\left((C_1 \otimes D_1 - D_1 \otimes C_2) + \lambda_2(D_2 \otimes D_1 - D_1 \otimes D_2)\right)w \\
&= -((v_1 \otimes v_2)^\top w)\left((C_1 + \lambda_2 D_2) \otimes D_1 - D_1 \otimes (C_2 + \lambda_2 D_2)\right)(v \otimes v) \\
&= \mathbf{0}_{(2n+1)^2}.
\end{aligned}$$

Moreover, since $\|w\|_2 = 1$, we have $u^\top w = 0$, which means u is orthogonal to w. Therefore, we can rewrite u with some nonzero coefficient vector $c \in \mathbb{R}^{(2n+1)^2-1}$ as $u = Wc$. We now observe that this c is a null vector of $\widetilde{B}(\lambda_2)$:

$$\widetilde{B}(\lambda_2)c = W^\top B(\lambda_2)Wc = W^\top B(\lambda_2)u = \mathbf{0}_{(2n+1)^2-1}.$$

This completes the proof. ∎

In most cases, w is the only null vector of $B(\lambda_2)$ and the projected Bézout matrix $\widetilde{B}(\lambda_2)$ is regular; then we can solve (2.17) via solving $\det \widetilde{B}(\lambda_2) = 0$. However, in some rare cases, $\widetilde{B}(\lambda_2)$ is still singular independent of the value of λ_2. We deal with such cases in Section 4.

A direct computation of W requires $O(n^6)$ operations since the size of B is $O(n^2)$, which can be a significant computational cost in our algorithm. Fortunately, however, $\widetilde{B}(\lambda_2)$ can be computed in $O(n^4)$ flops using Householder transformations [11, Ch. 5]. Specifically, we first form a Householder reflector $P \in \mathbb{R}^{(2n+1)^2 \times (2n+1)^2}$ of the form $P = I_{(2n+1)^2} - 2pp^\top$ where p is a $(2n+1)^2$-dimensional vector with $\|p\|_2 = 1$. To multiply P by a matrix X of size $(2n+1)^2 \times (2n+1)^2$ efficiently, we use the identities $PX = X - 2p(p^\top X)$ and $XP = X - 2(Xp)p^\top$. We use a reflector P that satisfies $Pe_1 = w$ where $e_1 := (1, 0 \ldots, 0)^\top \in \mathbb{R}^{(2n+1)^2}$, so that the first row (and column) of P is equal to w: taking $p = (e_1 - w)/\|e_1 - w\|_2$ accomplishes this. Then we obtain $\widetilde{B}(\lambda_2)$ simply by forming $P^\top B(\lambda_2)P$ and removing the first row and column, which are all zero.

3.4. Pseudocode for $\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0$**.** Summarizing the section, below is the algorithm for solving the bivariate determinantal equations $\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0$.

Algorithm 3.1 Algorithm for solving $\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0$.

1: Form the Bézout matrix pencil $B(\lambda_2)$ according to the definition (3.4).
2: Form the projection matrix W to get $\widetilde{B}(\lambda_2) = W^\top B(\lambda_2) W$ as in (3.9).
3: Solve $\det \widetilde{B}(\lambda_2) = 0$ to obtain the candidates $\hat{\lambda}_2$. In doing so, introduce perturbation as in Algorithm 4.1 if necessary.
4: For all positive $\hat{\lambda}_2$ obtained in Step 3, solve $\det M_1(\lambda_1, \hat{\lambda}_2) = \det M_2(\lambda_1, \hat{\lambda}_2) = 0$. If these two equations hold for the same value of λ_1, return (λ_1, λ_2) as a solution.

As discussed in Section 3.1, we can alternatively start from finding λ_1 by solving $\det \widetilde{B}(\lambda_1) = 0$.

4. Perturbing the matrices to obtain a regular Bézout matrix pencil. Unfortunately, our algorithm faces difficulty in certain cases that result in the matrix pencils (3.10), (2.15) or (2.16) being singular. In such cases there are infinitely many solutions to the determinantal equations, and hence our algorithm fails to find a finite number of candidates for the Lagrange multipliers. Such cases arise for example when C_i, D_i $(i = 1, 2)$ defined by (2.14) have a common eigenpair (ξ, z), i.e.,

$$
\begin{aligned}
M_1(\lambda_1, \lambda_2) z &= (C_1 + \lambda_1 D_1 + \lambda_2 D_2) z = \xi(1 + \lambda_1 + \lambda_2) z, \\
M_2(\lambda_1, \lambda_2) z &= (C_2 + \lambda_1 D_1 + \lambda_2 D_2) z = \xi(1 + \lambda_1 + \lambda_2) z.
\end{aligned}
$$

In this case, (2.17) holds for all (λ_1, λ_2) satisfying $1 + \lambda_1 + \lambda_2 = 0$. Therefore, infinitely many values λ_2 satisfy $B(\lambda_2)(z \otimes z) = 0$, which means $B(\lambda_2)$ is singular for every value of λ_2. If z is linearly independent of $v = \begin{bmatrix} Q_1^{-1} q_1 & \mathbf{0}_n^\top & -1 \end{bmatrix}^\top$, the null vector of D_1, we cannot remove $z \otimes z$ by the projection discussed in Section 3.3. Thus $\det \widetilde{B}(\lambda_2) = 0$ has infinitely many solutions λ_2.

This issue arises also in the algorithm for the signed distance problem [13], for which a slight perturbation is used as a remedy. We will similarly introduce a perturbation strategy that overcomes this issue. Although the perturbation does alter the problem, it can be regarded as a small backward error in the solution [12], and backward stability is generally the best a numerical algorithm can hope to achieve. Hence numerically its use is acceptable as long as the perturbation size is in the order of working precision.

We now discuss how to perturb the matrices to obtain a regular Bézout matrix pencil. Recall that $M_i(\lambda_1, \lambda_2) = C_i + \lambda_1 D_1 + \lambda_2 D_2$ where C_i, D_i are as defined in (2.14).

To show that the projected Bézout matrix $\widetilde{B}(\lambda_2)$ is a regular matrix pencil, it suffices to ensure that for one fixed value of λ_2, $\widetilde{B}(\lambda_2)$ is a nonsingular matrix. For simplicity, let us take $\lambda_2 = 0$; we can take λ_2 to be any fixed value by replacing C_i with $C_i + \lambda_2 D_2$ in what follows. Our goal is to derive a sufficient condition so that (or rather a perturbation strategy to guarantee) $\widetilde{B}(0)$ is nonsingular, which implies that $\widetilde{B}(\lambda_2)$ is a regular matrix pencil. Note that

$$M_1(\lambda_1, 0) = C_1 + \lambda_1 D_1, \tag{4.1}$$

$$M_2(\lambda_1, 0) = C_2 + \lambda_1 D_1. \tag{4.2}$$

We shall first perturb C_1 and C_2 so that they are nonsingular; this can be done by perturbing Q_0 (and γ_i if they are zero), as the determinant expansion of C_i contains a term of the form $\pm\det(Q_0)\gamma_i$. This forces $M_1(\lambda_1,0)$ and $M_2(\lambda_1,0)$ to be regular matrix pencils. Therefore, Lemma 3.1 holds for $P_1(\lambda_1) = I_{2n+1}\otimes M_1(\lambda_1,0)$ and $P_2(\lambda_1) = M_2(\lambda_1,0)\otimes I_{2n+1}$. Moreover, in the formulation of $\widetilde{B}(\lambda_2)$, we projected out v, which is a common eigenvector of $M_1(\infty,0)$ and $M_2(\infty,0)$. Hence, by Lemma 3.1 again, $\widetilde{B}(0)$ is nonsingular if and only if $M_1(\lambda_1,0)$ and $M_2(\lambda_1,0)$ share no eigenvalue λ_1 other than $|\lambda_1| = \infty$.

Now suppose to the contrary that $\det M_1(\lambda_1,0) = \det M_2(\lambda_1,0) = 0$ has a solution with $|\lambda_1| < \infty$, and there exist nonzero vectors x, y such that

$$M_1(\lambda_1,0)x = (C_1+\lambda_1 D_1)x = 0, \tag{4.3}$$

$$M_2(\lambda_1,0)y = (C_2+\lambda_1 D_1)y = 0. \tag{4.4}$$

More generally, we need to allow $M_1(\lambda_1,0), M_2(\lambda_1,0)$ to have nullity possibly larger than one: suppose that $M_1(\lambda_1,0)[x_1,\dots,x_s] = O, M_2(\lambda_1,0)[y_1,\dots,y_t] = O$ for some integers s and t.

We shall introduce a perturbation strategy that preserves the structure in (2.14), in particular such that the zero terms remain zero.

Partition x_i conformally to C_1 as $\begin{bmatrix} x_{i1}\\ x_{i2}\\ x_{i3}\end{bmatrix}$. Then take the SVD of the matrix of top and bottom parts: $\begin{bmatrix} x_{11},\dots,x_{s1}\\ x_{13},\dots,x_{s3}\end{bmatrix} = U\Sigma V^\top$. We then define $\hat{x}_i$ by $[\hat{x}_1,\dots,\hat{x}_s] = [x_1,\dots,x_s]V$, so that the first s_1 vectors of the top and bottom parts have full column rank. Note that $[\hat{x}_1,\dots,\hat{x}_s]$ still forms a complete set of eigenvectors for $M_1(\lambda_1,0)$. Generically we have $s_1 = s = 1$.

Now we describe the perturbation. To force the first s_1 vectors to yield $M_1(\lambda_1,0)\hat{x}_i \neq 0$, we can perturb C_1 as follows: if $x_1 \neq 0$ then add $\epsilon X_1 X_1^\top$ to Q_1, where $X_1 = [\hat{x}_1,\dots,\hat{x}_{s_1}]$. Note that $(M_1(\lambda_1,0)+\epsilon X_1X_1^\top)[\hat{x}_1,\dots,\hat{x}_{s_1}]$ is full rank, so no linear combination of $\hat{x}_1,\dots,\hat{x}_{s_1}$ is a null vector of $(M_1(\lambda_1,0)+\epsilon X_1X_1^\top)$, and by taking ϵ small enough we can ensure the perturbation in the other eigenvalues of $M_1(\lambda_1,0)$ do not spawn a newly generated solution for $\det M_1(\lambda_1,0) = \det M_2(\lambda_1,0) = 0$.

If $s > s_1$ holds, for the remaining $s - s_1$ vectors, we have $\hat{x}_i = \begin{bmatrix}\mathbf{0}_n\\ \hat{x}_{i2}\\ 0\end{bmatrix}$ where $i = s_1+1,\dots,s$. So it turns out that (4.3) and (4.4) both lead to the same equation

$$\left(\begin{bmatrix}-Q_0\\ O\\ -q_0^\top\end{bmatrix} + \lambda_1\begin{bmatrix}-Q_1\\ O\\ -q_1^\top\end{bmatrix}\right)\hat{x}_{i2} = 0.$$

We rewrite this as

$$-Q_0\hat{x}_{i2} = \lambda_1 Q_1 \hat{x}_{i2}, \tag{4.5}$$

$$(-q_0-\lambda_1 q_1)^\top \hat{x}_{i2} = 0. \tag{4.6}$$

In words, $\hat{x}_{i2}$ must be an eigenvector of the matrix pencil $-Q_0-\lambda Q_1$ for the eigenvalue $\lambda = \lambda_1$, and orthogonal to the vector $-q_0-\lambda_1 q_1$. Since $Q_1 \succ O$ by Assumption 2.1, we see the matrix pencil $-Q_0-\lambda Q_1$ is regular. Therefore, we can perturb q_0 to rule out the existence of such $\hat{x}_{i2}$ and λ_1.

By the perturbation just described, we obtain a Bézout matrix pencil $B(\lambda_2)$ such that $B(0)$ has no null vector except for the one corresponding to $|\lambda_1| = \infty$, i.e., $v \otimes v$ where v is the null vector of D_1. As discussed in Section 3.3, we project out $w = v \otimes v$ from $B(\lambda_2)$ to form $\widetilde{B}(\lambda_2)$. Therefore, by perturbing matrices and projecting out the null vector $w = v \otimes v$, we obtain the projected Bézout matrix $\widetilde{B}(\lambda_2)$ such that $\det \widetilde{B}(0) \neq 0$. This means the projected Bézout matrix pencil $\widetilde{B}(\lambda_2)$ is regular and (3.10) has finitely many eigenvalues.

Note that the above argument shows how to force $M_1(\lambda_1, 0), M_2(0, \lambda_2)$ to be regular matrix pencils. We use the same process to deal with the singular case of (2.15) and (2.16). This completes the description of the perturbation process as we summarize below.

Algorithm 4.1 Perturbation process to enforce regularity when necessary.

1: Solve $\det M_1(\lambda_1, 0) = \det M_2(\lambda_1, 0) = 0$ via the QZ algorithm. If there is no solution, no need to perturb; exit.
2: If C_1 and C_2 are singular, perturb Q_0 (and γ_i if they are zero) to modify its determinant so that $\det C_1 C_2 \neq 0$. Solve $\det M_1(\lambda_1, 0) = \det M_2(\lambda_1, 0) = 0$ again; if no solution, exit. Otherwise proceed.
3: Find null space $[x_1, \ldots, x_s]$ of $M_1(\lambda_1, 0)$.
4: Compute the SVD $\begin{bmatrix} x_{11}, \ldots, x_{s1} \\ x_{13}, \ldots, x_{s3} \end{bmatrix} = U \Sigma V^\top$. Let s_1 be the number of positive singular values.
5: Perturb $C_1 := C_1 + \epsilon X_1 X_1^\top$.
6: If $s > s_1$, compute the eigenvectors of $-Q_0 - \lambda_1 Q_1$ and perturb q_0 so that it is orthogonal to none of them.

5. Summary of the algorithm. In this section, we summarize the whole algorithm for solving 2QCQP. Complexity analysis is given to see that the runtime of our algorithm is $O(n^6)$.

5.1. Outline of the algorithm. We now show the pseudocode for the whole algorithm for solving 2QCQP.

Algorithm 5.1 Outline of algorithm for solving 2QCQP.

1: Test whether the problem is feasible by solving (2.1).
2: Let $(0, 0)$ be one of the candidates of (λ_1, λ_2).
3: Solve $\det M_1(\lambda_1, 0) = 0$ and add its solutions $(\lambda_1, 0)$ to the candidates of Lagrange multipliers. Similarly, solve $\det M_2(0, \lambda_2) = 0$ and get $(0, \lambda_2)$ as candidates.
4: Solve $\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0$ and add its solutions (λ_1, λ_2) to the candidates of Lagrange multipliers.
5: For every (λ_1, λ_2) with $\lambda_1, \lambda_2 \geq 0$ thus obtained, compute the corresponding x.
6: For every x obtained in Step 5, check the feasibility and rule out infeasible x.
7: For every x obtained in Step 6, compute the objective function values: The vector x corresponding to the smallest is a global solution.

REMARK 5.1. *In practice, since our algorithm requires the numerical solutions of eigenvalue problems such as* $\det \widetilde{B}(\lambda_2) = 0$ *and linear systems* $H(\lambda_1, \lambda_2)x = y$, *the computed solution may have relatively large numerical error, depending on the condition numbers. Thus the computed solution may slightly violate the constraints*

$g_1(x) \le 0$ *and* $g_2(x) \le 0$. *To refine the computed solution* x *so that it satisfies the constraints to working precision, we update the computed solution as follows: if our algorithm yields some candidates* x *for the global solution such that* $g_1(x) > \epsilon$ *or* $g_2(x) > \epsilon$, *we increment* x *by a small vector* Δx *so that the refined solution* $x + \Delta x$ *satisfies the constraints to first order. Since* $g_i(x+\Delta x) = g_i(x) + 2(x^\top Q_i + q_i^\top)\Delta x + \Delta x^\top Q_i \Delta x$, *ignoring the quadratic terms in* Δx *we compute* Δx *as the minimum-norm solution of*

$$\begin{bmatrix} x^\top Q_1 + q_1^\top \\ x^\top Q_2 + q_2^\top \end{bmatrix} \Delta x = - \begin{bmatrix} \max(0, g_1(x)/2) \\ \max(0, g_2(x)/2) \end{bmatrix}. \tag{5.1}$$

If Δx *satisfies* (5.1) *exactly and* $\Delta x^\top Q_i \Delta x < \epsilon$ *holds, we have* $g_i(x+\Delta x) < \epsilon$.

We have observed in our experiments that this refinement indeed improves the feasibility and accuracy of the computed solution.

5.2. Complexity analysis. The algorithm requires a solution of the linear generalized eigenvalue problem $\det \widetilde{B}(\lambda_2) = 0$ or $\det \widetilde{B}(\lambda_1) = 0$, whose size is bounded by $(2n+1)^2$. Since the standard QZ algorithm for computing the eigenvalues of an $N \times N$ linear generalized eigenvalue problem requires about $30N^3$ floating point operations [11, §7.7.7], the computational cost is about $30(2n+1)^6 \approx (1.9 \times 10^3)n^6$ flops. This is the dominant cost in our algorithm.

We now examine the computational costs of other steps.

- Step 2 of Algorithm 3.1: As we mentioned in Section 3.3, the projection matrix W can be formed in $O(n^4)$ time by the Householder transformation. Hence Step 2 of Algorithm 3.1 requires $O(n^4)$ time.
- Step 4 of Algorithm 3.1: The number of positive $\hat{\lambda}_2$ obtained by solving $\det \widetilde{B}(\lambda_2) = 0$ is bounded by $(2n+1)^2$. Therefore, in Step 4 of Algorithm 3.1, the two $(2n+1) \times (2n+1)$ linear generalized eigenvalue problems $\det M_1(\lambda_1, \hat{\lambda}_2) = \det M_2(\lambda_1, \hat{\lambda}_2) = 0$ are solved in $O(n^3)$ time at most $(2n+1)^2$ times, which means the computational cost required in this step is at most $O(n^5)$.
- Step 5 of Algorithm 5.1: $H(\lambda_1, \lambda_2)x = y$ is solved for x among all nonnegative pairs of (λ_1, λ_2) satisfying $\lambda_1 \det M_1(\lambda_1, \lambda_2) = \lambda_2 \det M_2(\lambda_1, \lambda_2) = 0$. By Bézout's theorem (e.g., [14]), the number of common solutions satisfying these determinantal equations is bounded by $(2n+2)^2$. Therefore, all KKT points x are computed in $O(n^5)$ time, once the pairs of (λ_1, λ_2) are obtained.

6. Numerical experiments. In this section, we present numerical experiments on runtime of our algorithm and comparison with the SDP relaxation. All experiments were conducted in MATLAB R2010b on a Core i7 machine with 16GB RAM, and we solved SDP by SeDuMi 1.3.

6.1. Runtime analysis of our algorithm. We generated random instances of 2QCQP for $n = 5, 10, \dots, 40$ and examined the runtime of our algorithm. The random instances are generated as in Burer and Anstreicher [7] and they are also used in Section 6.2.2. In addition to the total runtime of our algorithm, we measured the runtime breakdown of the following major parts:

- Solving a linear generalized eigenvalue problem $\det \widetilde{B}(\lambda_2) = 0$.
- Finding λ_1 from the computed $\hat{\lambda}_2$ via $\det M_1(\lambda_1, \hat{\lambda}_2) = \det M_2(\lambda_1, \hat{\lambda}_2) = 0$.
- Computing KKT points x from λ_1, λ_2 by solving $H(\lambda_1, \lambda_2)x = y$.
- Solving $\det M_1(\lambda_1, 0) = 0$ and $\det M_2(0, \lambda_2) = 0$ for the $\lambda_1 \lambda_2 = 0$ cases.

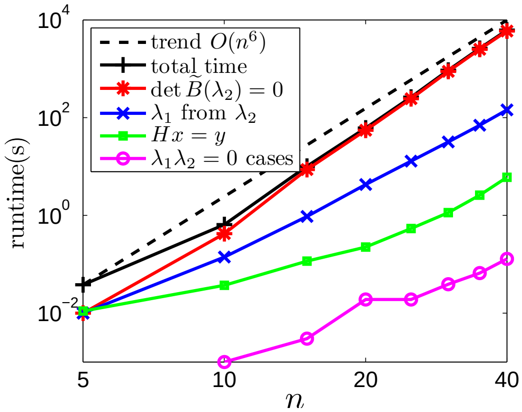

FIG. 6.1. *Double-logarithmic graph of runtime for varying n. We name the computational time of four parts shown above as "$\det \widetilde{B}(\lambda_2) = 0$", "$\lambda_1$ from λ_2", "$Hx = y$", and "$\lambda_1\lambda_2 = 0$ cases" in the legend in order of appearance.*

The dominant cost of our algorithm is solving a linear generalized eigenvalue problem $\det \widetilde{B}(\lambda_2) = 0$, which requires $O(n^6)$ time. In all cases with $n \geq 25$, at least 90% of the runtime was spent on solving $\det \widetilde{B}(\lambda_2) = 0$. Figure 6.1 illustrates that the computational times spent for the whole algorithm and solving $\det \widetilde{B}(\lambda_2) = 0$ scales asymptotically as $O(n^6)$.

6.2. Comparing our algorithm with the SDP relaxation. Here we apply our algorithm and the SDP relaxation to some 2QCQP instances and compare their outcomes. The basic SDP relaxation of 2QCQP is formulated as

$$\begin{aligned} \text{(6.1)} \qquad & \underset{x}{\text{minimize}} && Q_0 \bullet X + 2q_0^\top x + \gamma_0 \\ & \text{subject to} && Q_i \bullet X + 2q_i^\top x + \gamma_i \leq 0 \quad (i = 1, 2) \\ & && X \succeq xx^\top. \end{aligned}$$

It is known that the SDP relaxation is tight if $\operatorname{rank}(X) = 1$. Numerically, we define $\operatorname{rank}(X)$ by the number of eigenvalues of X whose absolute value is larger than 10^{-4}, following Ai and Zhang [1].

We denote the objective function values obtained by our algorithm and the SDP relaxation by v_{prop} and v_{SDP}, respectively. We let $\epsilon = 10^{-8}$ and regard a solution x feasible if $g_1(x) \leq \epsilon$ and $g_2(x) \leq \epsilon$ hold. If our algorithm yields some candidates x for a global solution such that $g_1(x) > \epsilon$ or $g_2(x) > \epsilon$, we apply the refinement method shown in Remark 5.1.

6.2.1. Two-dimensional instance from Burer and Anstreicher. For ease of visualization we first consider the following two-dimensional instance as described in Burer and Anstreicher [7]:

$$\begin{aligned} \text{(6.2)} \qquad & \underset{x}{\text{minimize}} && x^\top \begin{bmatrix} -4 & 1 \\ 1 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top x \\ & \text{subject to} && \|x\|_2^2 \leq 1, \quad x^\top \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} x \leq 2. \end{aligned}$$

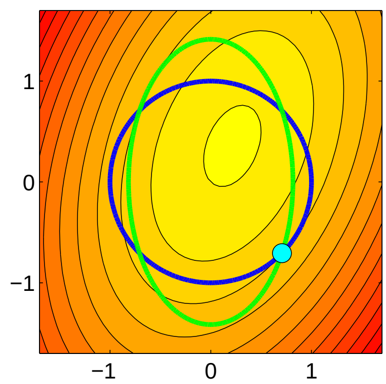

FIG. 6.2. *Two-dimensional instance from Burer and Anstreicher and its global solution.*

We illustrate the objective function value and the feasible region for this problem in Figure 6.2. The global solutions of this problem are $x^* = (\pm 1, \mp 1)/\sqrt{2}$ with objective value -4. Our algorithm computed the solution $(0.70711, -0.70711)$ with $v_{\text{prop}} = -4.0000$. As Burer and Anstreicher [7] show, the SDP relaxation of this problem is not tight and one obtains $v_{\text{SDP}} = -4.25$ by solving this problem via the SDP relaxation. By applying their strengthened approach, one can obtain objective value -4.0360, which still leaves a 0.9% gap from the exact value.

6.2.2. Random instances from Burer and Anstreicher. In [7], Burer and Anstreicher also present a practical method to generate CDT[1] random instances. They consider generating CDT instances with several candidates for a global solution, which makes the instances challenging. For $n = 2, 5, 10, 20$, we generated 100 such instances in the same way and solved by our algorithm and the SDP relaxation.

First we check whether a global solution obtained by our algorithm satisfies the necessary condition for global optimality (unfortunately, to our knowledge, no effective way of guaranteeing global optimality is available). As Yuan [25] proved, the Hessian of Lagrangian $H(\lambda_1, \lambda_2)$ has at most one negative eigenvalue at a global solution of the CDT problem. Denoting the number of negative eigenvalues of $H(\lambda_1, \lambda_2)$ at a global solution by $\psi(H(\lambda_1, \lambda_2))$, we show in Table 6.1 the number of instances with $\psi(H(\lambda_1, \lambda_2)) = 0$ and $\psi(H(\lambda_1, \lambda_2)) = 1$ for $H(\lambda_1, \lambda_2)$ at the global solution computed by our algorithm. Table 6.1 shows that $\psi(H(\lambda_1, \lambda_2)) \in \{0, 1\}$ in all cases, indicating that no solution obtained by our algorithm violates the necessary condition for the global optimality.

[1] More precisely, TTRS (two trust-region subproblem) as Burer and Anstreicher call in [7]. TTRS is a 2QCQP with $Q_2 \succ O$, which is a special case of CDT.

TABLE 6.1
The number of instances (among 100) whose $H(\lambda_1, \lambda_2)$ has zero or one negative eigenvalue at the global solutions.

n	$\psi(H(\lambda_1, \lambda_2)) = 0$	$\psi(H(\lambda_1, \lambda_2)) = 1$
2	99	1
5	90	10
10	90	10
20	92	8

Now we consider solving these instances by the SDP relaxation. Indeed, as shown in [1], the SDP relaxation solved CDT if and only if $\psi(H(\lambda_1, \lambda_2)) = 0$. For all instances, we applied the SDP relaxation and checked this fact numerically by confirming that $\operatorname{rank}(X) = 1$ is satisfied if and only if $\psi(H(\lambda_1, \lambda_2)) = 0$ holds[2].

6.2.3. Random instances with indefinite Q_2. Based on the idea Burer and Anstreicher [7] present, we generate random instances of 2QCQP with indefinite Q_2 as follows:

1. Fix the dimension n and set $Q_1 = I, q_1 = \mathbf{0}_n, \gamma_1 = -n^2$ so that $g_1(x) = \|x\|_2^2 - n^2$.
2. Let Q_0 be a diagonal matrix with diagonal entries uniformly distributed in $[-1, 1]$ and generate q_0 with uniform entries in $[-1/2, 1/2]$. Set $\gamma_0 = 0$. Then the objective function is written as $f(x) = x^\top Q_0 x + 2q_0^\top x$.
3. Solve TRS: minimize $f(x)$ subject to $g_1(x) \leq 0$, and save its global solution x^*. Construct an orthogonal matrix V such that $x^*/\|x^*\|_2 = V^\top e_1$. Since $g_1(x^*) = \|x^*\|_2^2 - n^2 = 0$ with high probability, the TRS instances with $f(x) = x^\top V Q_0 V^\top x + 2q_0^\top V^\top x,\ g_1(x) = \|x\|_2^2 - n^2$ has an optimal solution ne_1. Update $Q_0 \leftarrow V Q_0 V^\top, q_0 \leftarrow V q_0$ to facilitate the construction of $g_2(x)$ in the next step.
4. Form an instance of 2QCQP by enforcing the additional quadratic constraint $x^\top Q_2 x + 2q_2^\top x + \gamma_2 \leq 0$, where $q_2 = \mathbf{0}_n, \gamma_2 = -n^2$. Q_2 is a diagonal matrix where first diagonal entry is fixed to be one and the other diagonal entries are generated uniformly in $[-1, 1]$. This construction of Q_2 makes the optimal solution of TRS in step 3, ne_1, infeasible for 2QCQP.

For $n = 2, 5, 10, 20$, we generated 100 such instances and solved them by our algorithm and the SDP relaxation. We summarize the result in Table 6.2. Compared with the number of instances with $\psi(H(\lambda_1, \lambda_2)) = 0$ in Table 6.1, the number of instances solved via the SDP relaxation decreases. On the other hand, our algorithm computes a global solution regardless of whether or not Q_2 is positive definite. These results indicate the effectiveness of our proposed algorithm for computing a global solution of 2QCQP with indefinite Q_2.

[2]In the experiments of Burer and Anstreicher [7], the SDP relaxation method solved 24.6% of 1000 instances for $n = 10$ and 4.1% for $n = 20$. However, for $n = 10, 20$, Table 6.1 shows $\psi(H(\lambda_1, \lambda_2)) = 0$ holds for almost 90% of 100 instances, which means 90% of all instances can be solved by the SDP relaxation. This gap is probably due to the evaluation criteria of $\operatorname{rank}(X)$ in Section 6.2.

TABLE 6.2
Outcome of solving random instances with indefinite Q_2 via the SDP relaxation.

n	% solved by SDP
2	81
5	52
10	46
20	44

Accuracy. Let us remark on the accuracy of the computed solution. Typically our algorithm gives solutions that are more accurate than the SDP-based ones by about 10^{-7}; for example with 100 random instances with $n = 10$, our solution always had objective value smaller than the SDP solution (when it had no relaxation gap) by between $[10^{-9}, 2 \times 10^{-5}]$. To make the comparison fair we used the refinement process in Remark 5.1 also for the SDP solution to ensure that all the computed solutions are feasible to working accuracy. Note that the comparison here depends on SeDuMi precision setting; in our experiments we have used the default setting (pars.eps= 10^{-8}).

7. Conclusion and discussion. We have developed a polynomial-time algorithm for finding a global solution of 2QCQP. Our algorithm solves 2QCQP as follows: find all Lagrange multipliers by solving a system of bivariate determinantal equations, compute the KKT points corresponding to the multipliers and then obtain a global solution with the smallest objective value among the KKT points. The key step of our algorithm is to convert the KKT conditions into a pair of bivariate determinantal equations, which is reduced to a two-parameter eigenvalue problem of size $O(n)$, which in turn is reduced to two linear generalized eigenvalue problems of size $O(n^2)$. For the case where some of these eigenvalue problems are singular, we propose a perturbation process as a remedy. The computational complexity of our algorithm is shown to be $O(n^6)$ in total. Numerical experiments are conducted to illustrate the runtime of our algorithm and to compare the outcome with the SDP relaxation method.

We now remark on possible future work. First, since the $O(n^6)$ complexity is a bottleneck when n is large, the design of a more efficient algorithm is a problem awaiting solution. In addition, our algorithm perturbs the input data of original 2QCQP if some eigenvalue problems are singular and it is desirable to have an approach that does not need such treatment.

Finally, consider applying our algorithm to the nonconvex quadratic minimization problem with m quadratic constraints (mQCQP):

$$\underset{x}{\text{minimize}} \quad f(x) = x^\top Q_0 x + 2q_0^\top x + \gamma_0 \tag{7.1}$$

$$\text{subject to} \quad g_i(x) = x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2, \ldots, m). \tag{7.2}$$

Just as we computed the Lagrange multipliers of 2QCQP via a two-parameter eigenvalue problem, the Lagrange multipliers of mQCQP can be obtained via a m-parameter eigenvalue problem. Therefore, an algorithm for m-parameter eigenvalue problem would enable us to compute a global solution of mQCQP.

REFERENCES

[1] W. Ai and S. Zhang, *Strong duality for the CDT subproblem: a necessary and sufficient condition*, SIAM J. Optim., 19 (2009), pp. 1735–1756.
[2] F. V. Atkinson, *Multiparameter Eigenvalue Problems*, Academic Press, NY, USA, 1972.
[3] A.I. Barvinok, *Feasibility testing for systems of real quadratic equations*, Discrete Comput. Geom., 10 (1993), pp. 1–13.
[4] É. Bézout, *Théorie Générale des Équations Algébriques*, PhD thesis, Pierres, Paris, 1779.
[5] D. Bienstock, *A note on polynomial solvability of the CDT problem*, ArXiv 1406.6429, (2014).
[6] I. M. Bomze and M. L. Overton, *Narrowing the difficulty gap for the Celis-Dennis-Tapia problem*, Preprint, Department of Statistics and Operations Research, Universitaet Wien, Vienna, Austia, (2013).
[7] S. Burer and K. Anstreicher, *Second-order-cone constraints for extended trust-region subproblems*, SIAM J. Optim., 23 (2013), pp. 432–451.
[8] M. R. Celis, J. E. Dennis, and R. A. Tapia, *A trust region strategy for nonlinear equality constrained optimization*, in Numerical Optimization, P. T. Boggs, R. H. Byrd, and R. B. Schnabel, eds., SIAM Philadelphia, 1985, pp. 71–82.
[9] X. Chen and Y.-X. Yuan, *On local solutions of the Celis–Dennis–Tapia subproblem*, SIAM J. Optim., 10 (1999), pp. 359–383.
[10] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, USA, 2009. Unabridged republication of book first published by Academic Press in 1982.
[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 4th ed., 2012.
[12] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, USA, 2nd ed., 2002.
[13] S. Iwata, Y. Nakatsukasa, and A. Takeda, *Computing the signed distance between overlapping ellipsoids*, METR 2014-20, University of Tokyo, July 2014.
[14] F. C. Kirwan, *Complex Algebraic Curves*, Cambridge University Press, Cambridge, 1992.
[15] L. Lerer and M. Tismenetsky, *The Bézoutian and the eigenvalue-separation problem for matrix polynomials*, Integral Equations and Operator Theory, 5 (1982), pp. 386–445.
[16] G. Li and Y. Yuan, *Compute a Celis-Dennis-Tapia step*, J. Comput. Math., 23 (2005), pp. 463–478.
[17] Y. Nakatsukasa, V. Noferini, and A. Townsend, *Computing the common zeros of two bivariate functions via Bézout resultants*, Numer. Math., 129 (2015), pp. 181–209.
[18] Y. Nesterov, H. Wolkowicz, and Y. Ye, *Noconvex quadratic optimization*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer, NY, USA, 2000, pp. 361–420.
[19] J.M. Peng and Y.X. Yuan, *Optimality conditions for the minimization of a quadratic with two quadratic constraints*, SIAM J. Optim., 7 (1997), pp. 579–594.
[20] M.J.D. Powell and Y. Yuan, *A trust-region algorithm for equality constrained optimization*, Math. Programming, 49 (1991), pp. 189–211.
[21] R. Stern and H. Wolkowicz, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.
[22] J.F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, Math. Oper. Res., 28 (2003), pp. 246–267.
[23] G. Wachsmuth, *On LICQ and the uniqueness of Lagrange multipliers*, Oper. Res. Lett., 41 (2013), pp. 78–80.
[24] B. Yang and S. Burer, *A two-variable approach to the two-trust-region subproblem*, Tech. Report, Department of Mathematics, University of Iowa, 2013.
[25] Y. Yuan, *On a subproblem of trust region algorithms for constrained optimization*, Math. Programming, 47 (1990), pp. 53–63.

Appendix A. A lower bound of the optimal value. From (1.2), we have

$$x^\top Q_1 x + 2q_1^\top x + \gamma_1 \le 0 \iff \|Q_1^{1/2}(x + Q_1^{-1}q_1)\|_2^2 \le \|Q_1^{-1/2}q_1\|_2^2 - \gamma_1.$$

Therefore, the value of objective function is bounded from below as follows:

$$\begin{aligned}
&x^\top Q_0 x + 2q_0^\top x + \gamma_0 \\
&= \left(Q_1^{1/2}(x + Q_1^{-1}q_1)\right)^\top Q_1^{-1/2} Q_0 Q_1^{-1/2} \left(Q_1^{1/2}(x + Q_1^{-1}q_1)\right) \\
&\quad + 2\left(Q_1^{-1/2}(q_0 - Q_0 Q_1^{-1} q_1)\right)^\top \left(Q_1^{1/2}(x + Q_1^{-1}q_1)\right) + q_1^\top Q_1^{-1} Q_0 Q_1^{-1} q_1 - 2q_0^\top Q_1^{-1} q_1 + \gamma_0 \\
&\ge \sigma_{\min}(Q_1^{-1/2} Q_0 Q_1^{-1/2}) \left\|Q_1^{1/2}(x + Q_1^{-1}q_1)\right\|_2^2 \\
&\quad - 2\left\|Q_1^{-1/2}(q_0 - Q_0 Q_1^{-1} q_1)\right\|_2 \left\|Q_1^{1/2}(x + Q_1^{-1}q_1)\right\|_2 + q_1^\top Q_1^{-1} Q_0 Q_1^{-1} q_1 - 2q_0^\top Q_1^{-1} q_1 + \gamma_0 \\
&\ge \min\{\sigma_{\min}(Q_1^{-1/2} Q_0 Q_1^{-1/2}), 0\} \left(\|Q_1^{-1/2}q_1\|_2^2 - \gamma_1\right) \\
&\quad - 2\left\|Q_1^{-1/2}(q_0 - Q_0 Q_1^{-1} q_1)\right\|_2 \sqrt{\left\|Q_1^{-1/2}q_1\right\|_2^2 - \gamma_1} + q_1^\top Q_1^{-1} Q_0 Q_1^{-1} q_1 - 2q_0^\top Q_1^{-1} q_1 + \gamma_0,
\end{aligned}$$

where $\sigma_{\min}(Q_1^{-1/2} Q_0 Q_1^{-1/2})$ is the minimum eigenvalue of $Q_1^{-1/2} Q_0 Q_1^{-1/2}$.

Appendix B. Discussion on the Karush-John optimality conditions. Here we discuss how to find local solutions that violate the LICQ via the Karush-John optimality conditions, which is the necessary conditions for local optimality without any constraint qualification. The Karush-John optimality conditions for 2QCQP can be written as follows:

$$(\lambda_0 Q_0 + \lambda_1 Q_1 + \lambda_2 Q_2)x = -(\lambda_0 q_0 + \lambda_1 q_1 + \lambda_2 q_2), \tag{B.1}$$
$$x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2), \tag{B.2}$$
$$\lambda_i \left(x^\top Q_i x + 2q_i^\top x + \gamma_i\right) = 0 \quad (i = 1, 2), \tag{B.3}$$
$$\lambda_i \ge 0 \quad (i = 1, 2). \tag{B.4}$$

We show that a global solution of 2QCQP can be obtained if at least one multiplier of $\lambda_0, \lambda_1, \lambda_2$ in the Karush-John optimality conditions is nonzero.

Note that, if $\lambda_0 \ne 0$ holds, the Karush-John optimality conditions are equivalent to the KKT conditions. Hence, in what follows we consider finding x that satisfies the Karush-John optimality conditions with $\lambda_0 = 0$.

Here we define a matrix $G(\lambda_1, \lambda_2) := \lambda_1 Q_1 + \lambda_2 Q_2$ and a vector $z(\lambda_1, \lambda_2) := -(\lambda_1 q_1 + \lambda_2 q_2)$. Now our goal is to find x such that

$$G(\lambda_1, \lambda_2)x = z(\lambda_1, \lambda_2), \tag{B.5}$$
$$x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2), \tag{B.6}$$
$$\lambda_i \left(x^\top Q_i x + 2q_i^\top x + \gamma_i\right) = 0 \quad (i = 1, 2), \tag{B.7}$$
$$\lambda_i \ge 0 \quad (i = 0, 1, 2). \tag{B.8}$$

Note that these conditions are similar to the KKT conditions shown in Section 2.2: In fact, they become equivalent by replacing $G(\lambda_1, \lambda_2)$ with $H(\lambda_1, \lambda_2)$ and z with y. As we did for finding KKT points, we first consider computing λ_1, λ_2 that satisfy (B.5)–(B.8). We now introduce two cases of λ_1, λ_2.

1. Exactly one of λ_1, λ_2 is zero. For definiteness, suppose $\lambda_1 > 0$ and $\lambda_2 = 0$. In this case, whether the conditions (B.5)–(B.8) hold or not does not depend on the value of λ_1 as long as it is positive. Therefore, we have $(\lambda_1, \lambda_2) = (1, 0)$ as a candidate for the multipliers satisfying (B.5)–(B.8). Similarly, we have $(\lambda_1, \lambda_2) = (0, 1)$ when $\lambda_1 = 0$ and $\lambda_2 > 0$.
2. $\lambda_1 > 0$ and $\lambda_2 > 0$. In this case, defining $\mu := \lambda_2/\lambda_1$, we see the conditions (B.5)–(B.8) are written as follows:

$$G(1, \mu)x = z(1, \mu), \tag{B.9}$$

$$x^\top Q_i x + 2q_i^\top x + \gamma_i = 0 \quad (i = 1, 2). \tag{B.10}$$

Then, as in Lemma 2.1, the following determinantal equations hold for every μ satisfying (B.9),(B.10):

$$\det N_1(\mu) = \det N_2(\mu) = 0 \tag{B.11}$$

where

$$N_i(\mu) := \begin{bmatrix} Q_i & -G(1, \mu) & q_i \\ -G(1, \mu) & O & z(1, \mu) \\ q_i^\top & z(1, \mu)^\top & \gamma_i \end{bmatrix} \quad (i = 1, 2). \tag{B.12}$$

The proof is completely analogous to that of Lemma 2.1. These equations can be solved for μ as generalized eigenvalue problems. If both $N_1(\mu)$ and $N_2(\mu)$ are singular matrix pencils, (B.11) has infinitely many solutions. In this case, we apply the perturbation method described in Section 4 to make $N_1(0)$ or $N_2(0)$ regular. By finding positive solutions of (B.11), we get $(\lambda_1, \lambda_2) = (1, \mu)$ as candidates for λ_1, λ_2 satisfying (B.5)–(B.8).

Note that the multipliers $(\lambda_0, \lambda_1, \lambda_2) = (0, 0, 0)$ satisfies the Karush-John conditions and it may correspond to the global solution. Unfortunately, however, we cannot compute x from $\lambda_0, \lambda_1, \lambda_2$ in this case.

For each (λ_1, λ_2) thus obtained, we compute x from (B.5) and check whether it satisfies (B.6)–(B.8). If $G(\lambda_1, \lambda_2)$ is singular, we compute x as we do for computing KKT points x from $H(\lambda_1, \lambda_2)x = y$ with singular $H(\lambda_1, \lambda_2)$: Specifically, apply what we show in Appendix C replacing H by G. Adding these x to the KKT points computed by the proposed algorithm, we have all candidates for x satisfying the Karush-John optimality conditions except for the one that corresponds to $\lambda_0 = \lambda_1 = \lambda_2 = 0$. We compute objective function values for these x and output x which gives the smallest one.

Appendix C. How to obtain a KKT point from singular $H(\lambda_1, \lambda_2)$. We now discuss how to obtain a KKT point x from $H(\lambda_1, \lambda_2)x = y$ when λ_1, λ_2 are computed but $H(\lambda_1, \lambda_2)$ is singular, which we mentioned in Section 2.2. Specifically, we show how to compute $v \in \mathbb{R}^r$ such that $x = x_* + H_0 v$ satisfies the KKT conditions (2.3)–(2.6). In Section C.1, we first introduce three cases of λ_1, λ_2 depending on whether they are zero or positive. In two of these three cases, we need to solve quadratic optimization problems with one quadratic equality constraint. We discuss how to solve them in Section C.2.

C.1. Three cases of λ_1, λ_2 for finding an appropriate vector v. We now consider the following three cases with respect to λ_1, λ_2 to compute v.

1. $\lambda_1 = \lambda_2 = 0$. In this case, we need to find v satisfying $g_1(x_* + H_0 v) \le 0$ and $g_2(x_* + H_0 v) \le 0$, which can be obtained by solving the following TRS for v:

$$\begin{aligned} &\underset{v}{\text{minimize}} && g_2(x_* + H_0 v) \\ &\text{subject to} && g_1(x_* + H_0 v) \le 0. \end{aligned} \tag{C.1}$$

   This problem can be solved via a SDP reformulation (see [22]). If the optimal value of (C.1) is positive, there is no feasible x for this case and we remove $(\lambda_1, \lambda_2) = (0, 0)$ from the candidates for the Lagrange multipliers.
2. Exactly one of λ_1, λ_2 is zero and the other is positive. For definiteness, suppose $\lambda_1 > 0$ and $\lambda_2 = 0$. In this case, we need to find v satisfying $g_1(x_* + H_0 v) = 0$ and $g_2(x_* + H_0 v) \le 0$, which we do by solving the following minimization problem for v:

$$\begin{aligned} &\underset{v}{\text{minimize}} && g_2(x_* + H_0 v) \\ &\text{subject to} && g_1(x_* + H_0 v) = 0. \end{aligned} \tag{C.2}$$

   We show how to solve this quadratic minimization problem with one quadratic equality constraint in Appendix C.2. If the optimal value of (C.2) is positive, there is no feasible x and we remove the pair (λ_1, λ_2) from the candidates for the Lagrange multipliers. Similarly, if $\lambda_1 = 0$ and $\lambda_2 > 0$ hold, we compute v by solving

$$\begin{aligned} &\underset{v}{\text{minimize}} && g_1(x_* + H_0 v) \\ &\text{subject to} && g_2(x_* + H_0 v) = 0. \end{aligned} \tag{C.3}$$

3. $\lambda_1 > 0$ and $\lambda_2 > 0$. In this case, we need to find v satisfying $g_1(x_* + H_0 v) = 0$ and $g_2(x_* + H_0 v) = 0$, i.e., we solve the following quadratic equations for v:

$$v^\top H_0^\top Q_1 H_0 v + 2(Q_1 x_* + q_1)^\top H_0 v + x_*^\top Q_1 x_* + 2q_1^\top x_* + \gamma_1 = 0, \tag{C.4}$$

$$v^\top H_0^\top Q_2 H_0 v + 2(Q_2 x_* + q_2)^\top H_0 v + x_*^\top Q_2 x_* + 2q_2^\top x_* + \gamma_2 = 0. \tag{C.5}$$

   We denote $A_i = H_0^\top Q_i H_0$, $b_i = Q_i x_* + q_i$, $c_i = x_*^\top Q_i x_* + 2q_i^\top x_* + \gamma_i$ for $i = 1, 2$ and consider solving

$$h_1(v) = v^\top A_1 v + 2b_1^\top v + c_1 = 0, \tag{C.6}$$

$$h_2(v) = v^\top A_2 v + 2b_2^\top v + c_2 = 0 \tag{C.7}$$

   where $A_1 \succ O$. First, we solve the following two quadratic optimization problems with one quadratic equality constraint by applying the technique shown in Appendix C.2:

$$\underset{v}{\text{minimize}} \quad h_2(v) \quad \text{subject to} \quad h_1(v) = 0, \tag{C.8}$$

$$\underset{v}{\text{maximize}} \quad h_2(v) \quad \text{subject to} \quad h_1(v) = 0. \tag{C.9}$$

   Let v_1, v_2 be the optimal solutions of these problems respectively. Since $h_2(v_1) \le 0$ and $h_2(v_2) \ge 0$ must hold in order for (C.6) and (C.7) to have a common solution, we remove (λ_1, λ_2) from the candidates for the Lagrange multipliers if $h_2(v_1) h_2(v_2) > 0$.

Now we define the following two sets:

$$\mathcal{E} = \{v \in \mathbb{R}^r \mid h_1(v) = 0\}, \tag{C.10}$$

$$\mathcal{H} = \{v \in \mathbb{R}^r \mid v = c_1 v_1 + c_2 v_2,\ c_1, c_2 \in \mathbb{R}\}. \tag{C.11}$$

$\mathcal{E}$ represents the boundary of the ellipsoid $h_1(v) = 0$ and $\mathcal{H}$ is a two-dimensional subspace containing the origin and v_1, v_2. In the definition of $\mathcal{H}$, if v_1, v_2 are linearly dependent, we replace v_1 by an arbitrary $v_1' \in \mathcal{E}$ such that v_1, v_2 are linearly independent. Note that the intersection $\mathcal{E} \cap \mathcal{H}$ is connected and $v_1, v_2 \in \mathcal{E} \cap \mathcal{H}$. Since $h_2(v_1) \le 0,\ h_2(v_2) \ge 0$ and the value of $h_2(v)$ changes continuously in $\mathcal{E} \cap \mathcal{H}$, there are some α_1, α_2 satisfying

$$h_1(\alpha_1 v_1 + \alpha_2 v_2) = 0, \tag{C.12}$$

$$h_2(\alpha_1 v_1 + \alpha_2 v_2) = 0. \tag{C.13}$$

So we obtain a vector $v = \alpha_1 v_1 + \alpha_2 v_2$ satisfying (C.6) and (C.7) by solving the following system for α_1, α_2:

$$h_1(\alpha_1 v_1 + \alpha_2 v_2) = (\alpha_1 v_1 + \alpha_2 v_2)^\top A_1 (\alpha_1 v_1 + \alpha_2 v_2) + 2b_1^\top (\alpha_1 v_1 + \alpha_2 v_2) + c_1 = 0,$$
$$h_2(\alpha_1 v_1 + \alpha_2 v_2) = (\alpha_1 v_1 + \alpha_2 v_2)^\top A_2 (\alpha_1 v_1 + \alpha_2 v_2) + 2b_2^\top (\alpha_1 v_1 + \alpha_2 v_2) + c_2 = 0.$$

These bivariate quadratic scalar equations can be solved by taking the Bézoutian of the polynomials $p_1(\alpha_1, \alpha_2) = h_1(\alpha_1 v_1 + \alpha_2 v_2)$ and $p_2(\alpha_1, \alpha_2) = h_2(\alpha_1 v_1 + \alpha_2 v_2)$ (see [4, 17]).

C.2. The quadratic minimization problem with one quadratic equality constraint. In Appendix C.1, we need to solve the following minimization problems (C.2) and (C.3) where $g_1(x_* + H_0 v)$ and $g_2(x_* + H_0 v)$ are defined by the left side of (C.4) and (C.5) respectively. We denote $A_i = H_0^\top Q_i H_0,\ b_i = Q_i x_* + q_i,\ c_i = x_*^\top Q_i x_* + 2q_i^\top x_* + \gamma_i$ and rewrite (C.2) and (C.3) as follows:

$$\underset{v}{\text{minimize}} \quad v^\top A_2 v + 2b_2^\top v + c_2 \quad \text{subject to} \quad v^\top A_1 v + 2b_1^\top v + c_1 = 0, \tag{C.14}$$

$$\underset{v}{\text{minimize}} \quad v^\top A_1 v + 2b_1^\top v + c_1 \quad \text{subject to} \quad v^\top A_2 v + 2b_2^\top v + c_2 = 0. \tag{C.15}$$

Now we focus on solving (C.15); (C.14) can be solved analogously. The technique is similar to the one for computing the Lagrange multipliers in Section 2. The KKT conditions for (C.15) can be written as follows:

$$(A_1 + \lambda A_2)v = -(b_1 + \lambda b_2), \tag{C.16}$$

$$v^\top A_2 v + 2b_2^\top v + c_2 = 0. \tag{C.17}$$

Note that, in one constraint minimization problems (C.14) and (C.15), LICQ is naturally satisfied and the KKT conditions are necessary conditions for global optimality. Let $A(\lambda) = A_1 + \lambda A_2$ and $b(\lambda) = -b_1 - \lambda b_2$. Then, similarly to Lemma 2.1, we see that $\det L(\lambda) = 0$ holds for every λ satisfying (C.16) and (C.17) where

$$L(\lambda) = \begin{bmatrix} A_2 & -A(\lambda) & b_2 \\ -A(\lambda) & O & b(\lambda) \\ b_2^\top & b(\lambda)^\top & c_2 \end{bmatrix}. \tag{C.18}$$

This determinantal equation can be solved for λ as a generalized eigenvalue problem. If $L(\lambda)$ is singular for every λ, we perturb matrices to force $L(\lambda)$ to be a regular matrix pencil; as we perturbed Q_0 (and γ_i if they are zero) in Section 4, we perturb A_1 and c_2 to ensure that $L(0)$ is nonsingular.

For each λ thus obtained, one can compute v by solving $A(\lambda)v = b$. If $A(\lambda)$ is nonsingular, v is uniquely determined and it satisfies (C.17) naturally. If $A(\lambda)$ is singular, we find one of the solutions v satisfying (C.16), (C.17) as follows: Let v_* be the minimum-norm solution of (C.16) and v_0 be an arbitrary null vector of $A(\lambda)$ such that $\|v_0\|_2 = 1$. We see that the solutions of (C.16) can be expressed as $v_* + tv_0$ where $t \in \mathbb{R}$ is an arbitrary constant. Then we substitute $v = v_* + tv_0$ into (C.17) and solve for t to obtain a solution satisfying (C.16) and (C.17). If no real solution t is obtained, the corresponding λ gives no feasible solution of (C.15). By an argument analogous to that in Section 2.2, one can verify that the value of the objective function is independent of t, which means the obtained solution $v = v_* + tv_0$ is one of the global solutions.