SIAM J. OPTIM.
Vol. 26, No. 3, pp. 1669–1694

SOLVING GENERALIZED CDT PROBLEMS VIA TWO-PARAMETER EIGENVALUES*

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

Abstract. We consider solving a nonconvex quadratic minimization problem with two quadratic constraints, one of which being convex. This problem is a generalization of the Celis–Denis–Tapia (CDT) problem and thus we refer to it as GCDT (Generalized CDT). 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 another algorithm for GCDT, which is guaranteed to find a global solution for almost all GCDT instances (and slightly perturbed ones in some exceptionally rare cases), in exact arithmetic (including eigenvalue computation). Our algorithm is based on the approach proposed by Iwata, Nakatsukasa, and Takeda (2015) for computing the signed distance between overlapping ellipsoids. Our algorithm computes all the Lagrange multipliers of GCDT 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. In practice, in finite precision arithmetic, our algorithm requires $O(n^6 \log\log u^{-1})$ computational time, where n is the number of variables and u is the unit roundoff. Although we derive our algorithm under the unrealistic assumption that exact eigenvalues can be computed, numerical experiments illustrate that our algorithm performs well in finite precision arithmetic.

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

AMS subject classifications. 49M37, 65K05, 90C26, 90C30

DOI. 10.1137/15100624X

1. Introduction.

In this paper, we consider solving the quadratic minimization problem with two quadratic constraints

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

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

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. This problem includes the Celis–Dennis–Tapia (CDT) problem as a special case, which minimizes a nonconvex quadratic function over the intersection of two convex quadratic constraints such that Q_1 is positive definite and Q_2 is positive semidefinite. We call this problem GCDT (generalized CDT).

The CDT problem was proposed by Celis, Dennis, and Tapia [9] 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 [26]; an optimal solution for TRS can be obtained from an optimal solution of the

*Received by the editors January 30, 2015; accepted for publication (in revised form) May 2, 2016; published electronically August 24, 2016.
http://www.siam.org/journals/siopt/26-3/100624.html
Funding: This research was supported by Grants-in-Aid for Scientific Research (Challenging Exploratory Research, 26540007) from Japan Society for Promotion of Science.
†Department of Mathematical Informatics, The University of Tokyo, Tokyo 113-8656, Japan (shinsaku_sakaue@mist.i.u-tokyo.ac.jp, nakatsukasa@mist.i.u-tokyo.ac.jp, takeda@mist.i.u-tokyo.ac.jp, iwata@mist.i.u-tokyo.ac.jp).

polynomial-time solvable SDP problem. Moreover, the polynomial solvability property is extended by Sturm and Zhang [27] 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 the TRS. The CDT problem can have a duality gap, in general (see, e.g., [24]). Ai and Zhang [2] derived easily verifiable conditions to characterize when the CDT 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 [22]). 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 7 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, 24] and the location of Lagrangian multipliers corresponding to a local minimizer [10].

Li and Yuan [18] 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 [30] 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 [8] 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 [29] 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 with ϵ-accuracy, that is, a solution guaranteed to satisfy the relaxed constraints: $x^\top Q_i x + 2q_i^\top x + \gamma_i \leq \epsilon$, and in exact arithmetic the solution has objective value within ϵ from the optimal. Unfortunately, however, Bienstock's polynomial-time algorithm does not appear to be very practical, because the polynomial-time feasibility algorithm looks difficult to implement. To the best of our knowledge, no polynomial-time algorithm for the CDT problem has been implemented and used to solve large-scale instances.

In this paper we derive a practical algorithm that is guaranteed to find a global solution for almost all GCDT instances (and slightly perturbed ones in some exceptionally rare cases), in exact arithmetic including exact eigenvalue computation; we make this precise shortly. 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 [25] for equality constrained optimization.

Our algorithm is based on the one developed in [14] for computing the signed distance between overlapping ellipsoids via solving a special case of GCDT. We generalize the algorithm to solve GCDT. The approach is to find the Lagrange multipliers of GCDT from the Karush–Kuhn–Tucker (KKT) conditions. The KKT conditions of GCDT 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.

Our algorithm thus computes a global solution for GCDT (aside from rare cases where perturbation is necessary), if the eigenvalue problems are solved exactly and so are the other computations (such as linear systems). Of course, in practice in finite precision arithmetic, solving linear systems exactly is unrealistic, let alone computing exact eigenvalues. In practice, generalized eigenvalue problems are usually solved approximately by the QZ algorithm [20], which is known to be backward stable. If one were to attempt to compute exact eigenvalues by the QZ algorithm, infinitely many iterations employing exact arithmetic would be necessary, and thus the computational complexity of our algorithm is not bounded in this setting. In practice in finite precision arithmetic with unit roundoff u, the computational complexity of the QZ algorithm for the generalized eigenvalue problem $\det(A - \lambda B) = 0$ $(A, B \in \mathbb{R}^{N\times N})$ is bounded by $O(N^3 \log\log u^{-1})$, and the overall computational complexity of our algorithm for GCDT is $O(n^6 \log\log u^{-1})$. Even with the approximate eigenvalues computed by QZ in standard double precision arithmetic $u \approx 1.1\times 10^{-16}$, we illustrate through experiments that our algorithm works well in practice.

As we shall see, in nongeneric cases our algorithm encounters singular matrix pencils, and we also discuss how to handle such issues via perturbation. In finite precision arithmetic, the solution has accuracy $O(\kappa u)$, where κ is the so-called condition number; see section 6.3.

This paper is organized as follows. In section 2, we derive the KKT conditions of GCDT 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 shows 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 discuss how to handle a troublesome case where the KKT conditions are not necessary for local optimality. In section 6, we summarize our algorithm and show that the running time of our algorithm is $O(n^6 \log\log u^{-1})$. Finally, in section 7, we present numerical experiments to demonstrate the practical performance of our algorithm.

Until section 6 we assume exact arithmetic (including exact generalized eigenvalues), and section 6 discusses the behavior of our algorithm in finite precision arithmetic, taking into account the unit roundoff.

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.

Until section 5, we consider finding all local solutions $\bar{x}$ that satisfy the linear independence constraint qualification (LICQ). Namely, 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. For such local solutions, the KKT conditions are necessary for local optimality, which means all local solutions satisfying the LICQ are obtained by computing all the KKT points. In section 5, we discuss how to find local solutions that violate the LICQ.

2.1. Checking feasibility. The assumption that Q_1 is positive definite enables us to check the feasibility of GCDT. 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 $Q_1 \succ O$, GCDT 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 GCDT by solving the Lagrangian dual of

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

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

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

Under the condition that $g_1(x) \leq 0$ is strictly feasible, there is no duality gap (see [27]) 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 GCDT is infeasible. In this way, we could check whether the GCDT is feasible or not if $Q_1 \succ O$.

Furthermore, as we see above, if $g_1(x) \leq 0$ is not strictly feasible, GCDT is trivial, i.e., the only possible solution is $x = -Q_1^{-1}q_1$. Therefore, throughout the following discussion, we consider the GCDT such that the constraint $g_1(x) \leq 0$ is strictly feasible.

Since $Q_1 \succ O$, the feasible region is closed and bounded, which implies by the Weierstrass extreme value theorem that a feasible GCDT has an optimal solution. 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.

2.2. The KKT conditions for the problem. As discussed above, we focus on the case where GCDT has an optimal solution. Let $x \in \mathbb{R}^n$ be a local solution of GCDT that satisfies the LICQ. Then there exists a pair of Lagrange multipliers $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ satisfying the KKT conditions:

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

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

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

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

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 (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

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

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

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

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 GCDT, but as long as the number

of solutions is finite, we can find the GCDT 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 $\operatorname{rank}(H_0) = r$ 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:

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

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 a v, we obtain one of the global solutions. We discuss how to obtain such a v in Appendix B.

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. As above, (2.17) may have infinitely many solutions (λ_1, λ_2), and in such cases, we need to perturb some of the matrices. We discuss how to perturb the matrices in detail in section 4.

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:

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

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

As we discuss later, the matrices $B(\lambda_1), B(\lambda_2)$ are the Bézout matrices [17]. The binary operator $\otimes$ is the Kronecker product for matrices $X \in \mathbb{R}^{k\times l}, Y \in \mathbb{R}^{m\times n}$, defined as

$$X \otimes Y = \begin{bmatrix} X_{11}Y & X_{12}Y & \cdots & X_{1l}Y \\ X_{21}Y & X_{22}Y & \cdots & X_{2l}Y \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1}Y & X_{k2}Y & \cdots & X_{kl}Y \end{bmatrix}.$$

Note that the Kronecker product satisfies $(X_1 \otimes Y_1)(X_2 \otimes Y_2) = X_1X_2 \otimes Y_1Y_2$. The following property motivates us to use $B(\lambda_1), B(\lambda_2)$ for reducing the two-parameter eigenvalue problem (3.1), (3.2) to the generalized eigenvalue problems (3.3), (3.4): let v_1, v_2 be eigenvectors for the two-parameter eigenvalue problems (3.1), (3.2) with eigenvalues λ_1, λ_2, i.e.,

$$\begin{aligned} M_1(\lambda_1, \lambda_2)v_1 &= (C_1 + \lambda_1 D_1 + \lambda_2 D_2)v_1 = 0, \\ M_2(\lambda_1, \lambda_2)v_2 &= (C_2 + \lambda_1 D_1 + \lambda_2 D_2)v_2 = 0. \end{aligned}$$

For such λ_1, λ_2, a vector $v_1 \otimes v_2$ is an eigenvector of $B(\lambda_2)$ as follows (see, e.g., [7]):

$$\begin{aligned} B(\lambda_2)(v_1 \otimes v_2) &= ((C_1 \otimes D_1 - D_1 \otimes C_2) + \lambda_2(D_2 \otimes D_1 - D_1 \otimes D_2))(v_1 \otimes v_2) \\ &= (C_1v_1 \otimes D_1v_2 - D_1v_1 \otimes C_2v_2) + \lambda_2(D_2v_1 \otimes D_1v_2 - D_1v_1 \otimes D_2v_2) \\ &= ((C_1 + \lambda_1 D_1 + \lambda_2 D_2)v_1) \otimes D_1v_2 - D_1v_1 \otimes ((C_2 + \lambda_1 D_1 + \lambda_2 D_2)v_2) \\ &= 0. \end{aligned}$$

We can verify $B(\lambda_1)(v_1 \otimes v_2) = 0$ similarly. 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

$$\det M_1(\lambda_1, \hat{\lambda}_2) = \det M_2(\lambda_1, \hat{\lambda}_2) = 0 \tag{3.5}$$

among the $\hat{\lambda}_2$ thus obtained to get the corresponding λ_1, if any. If singular matrix pencils $M_i(\lambda_1, \hat{\lambda}_2)$ appear, we apply the perturbation technique shown in section 4, and we regard the values of λ_1 obtained from $\det M_1(\lambda_1, \hat{\lambda}_2) = 0$ and $\det M_2(\lambda_1, \hat{\lambda}_2) = 0$ as equal if they are within the error estimate, which reflects the magnitude of perturbation as we will discuss in section 6.3.

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) if $Q_2 \succ O$; the discussion in section 3.3 focuses on solving (3.4) using $Q_1 \succ O$). 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 facilitates the formulation of the Bézoutian for matrix polynomials [17].

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

$$\mathcal{B}(s,t) = \frac{P_1(s)P_2(t) - P_2(s)P_1(t)}{s-t} = \sum_{i,j=0}^{k-1,k-1} B_{i,j} s^i t^j \tag{3.6}$$

in s and t. Here $B_{i,j}$ is the $n \times n$ coefficient matrix corresponding to the term $s^i t^j$ 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 (see [17, 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 [11, 17] (this fact is not needed for what follows).

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

$$\begin{aligned}\mathcal{B}(s,t) &= \frac{P_1(s)P_2(t) - P_2(s)P_1(t)}{s-t} \\ &= \frac{M_1(t,\lambda_2) \otimes M_2(s,\lambda_2) - M_1(s,\lambda_2) \otimes M_2(t,\lambda_2)}{s-t} \\ &= (C_1 + \lambda_2 D_2) \otimes D_1 - D_1 \otimes (C_2 + \lambda_2 D_2) \\ &= B_{0,0}.\end{aligned}$$

Thus the Bézout matrix for $P_1(\lambda_1), P_2(\lambda_1)$ is expressed as

$$B = [B_{0,0}] = (C_1 \otimes D_1 - D_1 \otimes C_2) + \lambda_2(D_2 \otimes D_1 - D_1 \otimes D_2),$$

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 a 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, Q_1 is nonsingular. Then we see that D_1 has the null vector

$$v = \begin{bmatrix} Q_1^{-1} q_1 \\ \mathbf{0}_n \\ -1 \end{bmatrix} \tag{3.7}$$

since

$$D_1 v = \begin{bmatrix} O & -Q_1 & \mathbf{0}_n \\ -Q_1 & O & -q_1 \\ \mathbf{0}_n^\top & -q_1^\top & 0 \end{bmatrix} \begin{bmatrix} Q_1^{-1} q_1 \\ \mathbf{0}_n \\ -1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n \\ \mathbf{0}_n \\ 0 \end{bmatrix}. \tag{3.8}$$

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

$$w := \frac{v \otimes v}{\|v\|^2} \tag{3.9}$$

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* (3.7) *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 $g_1(x) \leq 0$ is strictly feasible in addition to $Q_1 \succ O$. 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

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

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

$$\widetilde{B}(\lambda_2) := W^\top B(\lambda_2) W. \tag{3.10}$$

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

$$\det \widetilde{B}(\lambda_2) = 0. \tag{3.11}$$

This means we obtain all solutions of (2.17) via solving (3.11).

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* (3.10) *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 (3.9). Since v as in (3.7) 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 $\|u\|_2 \ne 0$. By multiplying u with $B(\lambda_2)$, 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 $\widetilde{B}(\lambda_2)$:

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

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, independently 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 [12, Chap. 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 (3.10).
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 $\hat{\lambda}_2$ obtained in step 3, solve $\det M_1(\lambda_1, \hat{\lambda}_2) = \det M_2(\lambda_1, \hat{\lambda}_2) = 0$. If $M_i(\lambda_1, \hat{\lambda}_2)$ is singular, we introduce perturbation as in step 1 of Algorithm 4.1. 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, if we have $Q_2 \succ O$, 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 (3.11), (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.,

$$M_1(\lambda_1, \lambda_2)z = (C_1 + \lambda_1 D_1 + \lambda_2 D_2)z = \xi(1 + \lambda_1 + \lambda_2)z,$$
$$M_2(\lambda_1, \lambda_2)z = (C_2 + \lambda_1 D_1 + \lambda_2 D_2)z = \xi(1 + \lambda_1 + \lambda_2)z.$$

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

This issue arises also in the algorithm for the signed distance problem [14], 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 [13], 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 on the order of working precision.

4.1. Perturbation process. We now discuss how to introduce perturbation to obtain a regular projected Bézout matrix pencil $\widetilde{B}(\lambda_2)$. 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, which is used if singular matrix pencils $M_i(\lambda_1, \hat{\lambda}_2)$ appear in (3.5). 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

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

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

To ensure $M_1(\lambda_1, 0)$ and $M_2(\lambda_1, 0)$ are regular, we perturb γ_1 and γ_2. Specifically, note from their structure that $\det M_1(\lambda_1, 0)$ and $\det M_2(\lambda_1, 0)$ are of degree at most $2n$ as polynomials in λ_1, and the degree-$2n$ term is

$$\lambda_1^{2n} \det \begin{bmatrix} Q_i & -Q_1 & q_i \\ -Q_1 & O & -q_1 \\ q_i^\top & -q_1^\top & \gamma_i \end{bmatrix} \tag{4.3}$$

for $i = 1, 2$, which can be forced nonzero by slightly decreasing γ_i (we decrease them to prevent the problem from becoming infeasible). We can thus force $M_1(\lambda_1, 0)$ and $M_2(\lambda_1, 0)$ to be regular matrix pencils. We remark here that the degree-$2n$ term λ_1^{2n} of $\det M_i(\lambda_1, \hat{\lambda}_2)$ with $\hat{\lambda}_2 \neq 0$ fixed is also given by (4.3), and thus the above perturbation of γ_i can be employed also to force $M_i(\lambda_1, \hat{\lambda}_2)$ to be regular when necessary.

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 $|\lambda_1| = \infty$.

If the matrix pencil $Q_0 + \lambda_1 Q_1$ has a multiple eigenvalue λ_{1*}, then both $M_1(\lambda_{1*}, 0)$ and $M_2(\lambda_{1*}, 0)$ are singular. In this case, we perturb Q_0 to eliminate the multiple eigenvalue: let $x_1, \ldots, x_s$ form a basis of the null space of $Q_0 + \lambda_{1*} Q_1$ and update $Q_0 := Q_0 + \sum_{j=1}^{s} \alpha_j x_j x_j^\top Q_1$ with $\alpha_1, \ldots, \alpha_s$ being distinct positive values at most $O(u)$.

We now suppose that the eigenvalues of $Q_0 + \lambda_1 Q_1$ are all simple. For any value of λ_{1*} such that $\det M_1(\lambda_{1*}, 0) = 0$ and $\det(Q_0 + \lambda_{1*} Q_1) \neq 0$, we can easily perturb γ_2 so that $\det M_2(\lambda_{1*}, 0) \neq 0$. If $\det M_1(\lambda_{1*}, 0) = 0$ and $\det(Q_0 + \lambda_{1*} Q_1) = 0$ for some value of λ_{1*}, we must have

$$\operatorname{rank} \begin{bmatrix} q_0^\top + \lambda_{1*} q_1^\top \\ Q_0 + \lambda_{1*} Q_1 \end{bmatrix} = n - 1. \tag{4.4}$$

To verify (4.4), let $M_1(\lambda_{1*}, 0) \left[\begin{smallmatrix} w_1 \\ w_2 \\ w_3 \end{smallmatrix}\right] = 0$ with $\left[\begin{smallmatrix} w_1 \\ w_2 \\ w_3 \end{smallmatrix}\right] \neq 0$ and note from the second block that $(Q_0 + \lambda_{1*} Q_1) w_1 + (q_0 + \lambda_{1*} q_1) w_3 = 0$. If the scalar $w_3 \neq 0$ then this implies

$q_0+\lambda_{1*}q_1 \in \mathrm{Range}(Q_0+\lambda_{1*}Q_1)$, hence (4.4) holds. If $w_3 = 0$ then w_1 is parallel to the unique null vector of $Q_0+\lambda_{1*}Q_1$. By the first block we have $Q_1 w_1 = (Q_0+\lambda_{1*}Q_1)w_2$. Left multiplying $w_1^\top$ we obtain $w_1^\top Q_1 w_1 = 0$, which cannot happen unless $w_1 = \mathbf{0}_n$. Thus we conclude w_1, w_3 are both zero, from which we again obtain (4.4).

In view of (4.4), we can easily perturb q_0 so that $q_0^\top + \lambda_{1*}q_1$ is not orthogonal to the corresponding eigenvector x of $Q_0 + \lambda_1 Q_1$.

Thus, one can obtain a regular projected Bézout matrix by small perturbations of γ_1, γ_2, Q_0, or q_0.

4.2. Summary of perturbation. 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 (3.11) has finitely many eigenvalues. We summarize the perturbation process in Algorithm 4.1.

Algorithm 4.1. Perturbation process to enforce regularity when necessary.

1: Perturb (decrease) γ_1, γ_2 if necessary so that $\det \begin{bmatrix} Q_i & -Q_1 & q_i \\ -Q_1 & O & -q_1 \\ q_i^\top & -q_1^\top & \gamma_i \end{bmatrix} \neq 0$.

2: Compute eigenvalues of $Q_0+\lambda_1 Q_1$. For each multiple eigenvalue, compute a basis $x_1, \ldots, x_s$ of the eigenvectors and update $Q_0 := Q_0 + \sum_{j=1}^{s} \alpha_j x_j x_j^\top Q_1$.

3: Perturb q_0 as necessary so that $\mathrm{rank} \begin{bmatrix} q_0^\top + \lambda_{1*} q_1^\top \\ Q_0 + \lambda_{1*} Q_1 \end{bmatrix} = n$.

Note that the argument following (4.1), (4.2) shows how to force $M_1(\lambda_1, 0)$ to be regular matrix pencils by perturbing γ_1. Similarly, we can force $M_2(0, \lambda_2)$ to be regular by perturbing γ_2, along with Q_0 (if it is singular, then we add a small positive definite perturbation) so that C_2 is nonsingular. This is how we deal with the singular case of (2.15) and (2.16). As mentioned above, step 1 can be used also to modify singular matrix pencils $M_i(\lambda_1, \hat{\lambda}_2)$ in (3.5) to be regular. This completes the description of the perturbation process.

Let us emphasize that the above perturbation is required extremely rarely. Thus, in theory, our algorithm is almost always guaranteed to find global solutions without perturbation. This was indeed the case in all numerical experiments in section 7.

4.3. Effect of perturbation. We have argued that introducing perturbation is numerically harmless in the sense that they can be regarded as small backward errors. Nonetheless, the perturbation that we introduce does affect the solution x and the optimal value. For example, by perturbing $Q_0 := Q_0 + \Delta Q_0$ and $q_0 := q_0 + \Delta q_0$, the optimal objective value can change at most by $\|\Delta Q_0\|_2 \|x_*\|^2 + 2\|\Delta q_0\|_2 \|x_*\|$, where $\|x_*\|$ is a bound for the norm of the solution. For example, since $g_1(x) = (x + Q_1^{-1}q_1)^\top Q_1 (x + Q_1^{-1}q_1) - q_1^\top Q_1^{-1} q_1 + \gamma_1$, we have the upper bound $\|x_*\| \leq \|Q_1^{-1}q_1\| + (q_1^\top Q_1^{-1} q_1 - \gamma_1)/\sqrt{\sigma_{\min}(Q_1)}$.

We also allow perturbing γ_1, γ_2 resulting in a slightly larger feasible region; its effect on the objective value is nontrivial. However, after obtaining a global solution by perturbing γ_1, γ_2, we can perform a sensitivity analysis and estimate the optimal value of the original problem. Let $f^*(\epsilon)$ be the optimal value for the problem perturbed by adding ϵ_1, ϵ_2 (< 0) to γ_1, γ_2, and x_* a global solution and $\lambda_{1*}, \lambda_{2*}$ its

Lagrange multipliers. Then we evaluate the optimal value $f^*(\epsilon+\Delta)$ using the solution information of $f^*(\epsilon)$ for $(\Delta_1, \Delta_2) = (-\epsilon_1, -\epsilon_2)$ such as

$$\begin{aligned} f^*(0) = f^*(\epsilon + \Delta) &\approx f^*(\epsilon) + \lambda_{1*}\Delta_1 + \lambda_{2*}\Delta_2 \\ &= f^*(\epsilon) - \lambda_{1*}\epsilon_1 - \lambda_{2*}\epsilon_2, \end{aligned}$$

in which the approximation is accurate when Δ is sufficiently small that the inactive constraints are still inactive and the Lagrange multipliers are not much affected (these conditions hold if x_* satisfies the second-order sufficient condition for an isolated local solution together with strict complementarity; see, e.g., [23, section 12.8]). Therefore, as long as ϵ is sufficiently small, the perturbation does not cause a large change in the optimal value.

5. How to find local solutions that violate the LICQ. Here we discuss how to find local solutions that violate the LICQ via the Karush–John (sometimes called the Fritz John) optimality conditions, which are the necessary conditions for local optimality without any constraint qualification. The Karush–John optimality conditions for GCDT can be written as follows [15, 19]:

$$(\lambda_0 Q_0 + \lambda_1 Q_1 + \lambda_2 Q_2)x = -(\lambda_0 q_0 + \lambda_1 q_1 + \lambda_2 q_2), \tag{5.1}$$

$$x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2), \tag{5.2}$$

$$\lambda_i \left(x^\top Q_i x + 2q_i^\top x + \gamma_i\right) = 0 \quad (i = 1, 2), \tag{5.3}$$

$$\lambda_i \ge 0 \quad (i = 0, 1, 2), \tag{5.4}$$

$$(\lambda_0, \lambda_1, \lambda_2) \ne (0, 0, 0). \tag{5.5}$$

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

$$G(\lambda_1, \lambda_2)x = z(\lambda_1, \lambda_2), \tag{5.6}$$

$$x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2), \tag{5.7}$$

$$\lambda_i \left(x^\top Q_i x + 2q_i^\top x + \gamma_i\right) = 0 \quad (i = 1, 2), \tag{5.8}$$

$$\lambda_i \ge 0 \quad (i = 1, 2), \tag{5.9}$$

$$(\lambda_1, \lambda_2) \ne (0, 0). \tag{5.10}$$

Note that these conditions are similar to the KKT conditions shown in section 2.2: in fact, they become equivalent by removing (5.10) and 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 (5.6)–(5.10). 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 (5.6)–(5.10) 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 (5.6)–(5.10). 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 $\mu := \lambda_1/\lambda_2$, we see the conditions (5.6)–(5.10) are written as follows:

$$G(\mu, 1)x = z(\mu, 1), \tag{5.11}$$

$$x^\top Q_i x + 2q_i^\top x + \gamma_i = 0 \quad (i = 1, 2). \tag{5.12}$$

Then, as in Lemma 2.1, the following determinantal equations hold for every μ satisfying (5.11), (5.12):

$$\det N_1(\mu) = \det N_2(\mu) = 0, \tag{5.13}$$

where

$$N_i(\mu) := \begin{bmatrix} Q_i & -G(\mu,1) & q_i \\ -G(\mu,1) & O & z(\mu,1) \\ q_i^\top & z(\mu,1)^\top & \gamma_i \end{bmatrix} \quad (i = 1, 2). \tag{5.14}$$

The proof is completely analogous to that of Lemma 2.1. These equations can be solved for μ as generalized eigenvalue problems. Here, solving $\det N_1(\mu) = 0$ can be regarded as finding all candidates of Lagrange multiplier μ for the following TRS via the generalized eigenvalue problem

$$\underset{x}{\text{minimize}} \quad x^\top Q_2 x + 2q_2^\top x + \gamma_2 \tag{5.15}$$

$$\text{subject to} \quad x^\top Q_1 x + 2q_1^\top x + \gamma_1 = 0. \tag{5.16}$$

As shown in [1], Lagrange multipliers μ for the TRS can be computed via a $(2n+1) \times (2n+1)$ generalized eigenvalue problem without requiring perturbation. Thus the number of candidates μ obtained as solutions of (5.13) is finite. By finding positive solutions of (5.13), we get $(\lambda_1, \lambda_2) = (\mu, 1)$ as candidates for λ_1, λ_2 satisfying (5.6)–(5.10) (Note that, if $(\lambda_1, \lambda_2) = (\mu, 1)$ is a multiplier pair satisfying (5.11), (5.12), then $(\alpha\mu, \alpha)$ also satisfies them for all $\alpha > 0$. Here we have $(\lambda_1, \lambda_2) = (\mu, 1)$ as a representative pair.)

For each (λ_1, λ_2) thus obtained, we compute x from (5.6) and check whether it satisfies (5.7)–(5.10). If $G(\lambda_1, \lambda_2)$ is singular, we compute x just as when we compute the 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 B replacing H by G. Adding these x to the KKT points computed as shown in the above sections, we have all candidates for x satisfying the Karush–John optimality conditions, which means all candidates for a global solution are obtained regardless of whether the LICQ is satisfied or not. We then compute the objective function values for these x and output the x that gives the smallest one.

Note that the LICQ is violated so rarely that in practice it is almost always satisfied for randomly generated instances. This was indeed the case in all the numerical experiments in section 7.

6. Summary and analysis of the algorithm. In this section, we summarize the entire algorithm for solving GCDT, and analyze its behavior on a more realistic computation model. Throughout this section we assume finite precision arithmetic with unit roundoff u, so the eigenvalues λ_1, λ_2 are obtained approximately (not exactly). Complexity analysis is given to see that the runtime of our algorithm is $O(n^6 \log\log u^{-1})$. We also describe some remarks on the accuracy of our solution.

6.1. Outline of the algorithm. We now show the pseudocode for the whole algorithm for solving GCDT.

Algorithm 6.1. Outline of algorithm for solving GCDT.

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 $\det M_1(\lambda_1, \lambda_2) = \det M_2(\lambda_1, \lambda_2) = 0$ for $\lambda_1, \lambda_2 > 0$ by Algorithm 3.1 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: Solve $\det N_1(\mu) = \det N_2(\mu) = 0$ and compute x for $(\lambda_1, \lambda_2) = (0,1), (1,0), (\mu, 1)$.
7: For every x obtained in steps 5 and 6, check the feasibility and rule out infeasible x.
8: For every x obtained in step 7, compute the objective function values: the vector x corresponding to the smallest is a global solution.

Remark 6.1. 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 $g_1(x) \leq 0$ and $g_2(x) \leq 0$. 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

$$\begin{bmatrix} x^\top Q_1 + q_1^\top \\ x^\top Q_2 + q_2^\top \end{bmatrix} \Delta x = - \begin{bmatrix} \max(0, g_1(x)/2) \\ \max(0, g_2(x)/2) \end{bmatrix}. \tag{6.1}$$

If Δx satisfies (6.1) 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.

6.2. Complexity analysis. The algorithm requires a solution of the linear generalized eigenvalue problem $\det \widetilde{B}(\lambda_2) = 0$, 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 [12, section 7.7.7] (this estimate is based on the experimental rule of thumb that on average two QZ steps is enough to find one eigenvalue), 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 6.1: $H(\lambda_1, \lambda_2)x = y$ 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., [16]), 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.

The above complexity analysis is correct when the machine precision u is fixed; strictly speaking, since computing eigenvalues is an iterative process, the complexity depends on the working precision, and thus also on the desired accuracy. Let us remark on the relation between the complexity and accuracy. Recall from section 6.3 that the accuracy scales roughly with the precision as $O(\kappa u)$, which, asymptotically as $u \to 0$, can be regarded as $O(u)$. The complexity, when the machine precision u is taken as a variable, is $O(n^6 \log\log u^{-1})$. Here the $\log\log u^{-1}$ term accounts for the number of iterations in the eigensolver, which converge quadratically or faster [12, Chap. 7]. In virtually any practical application, $\log\log u^{-1}$ can be regarded as a constant.

6.3. Error analysis. Until section 6 we have developed our algorithm assuming exact arithmetic including exact eigenvalue computation. We now turn to the influence of using finite precision arithmetic. Numerical errors arise during our algorithm, in particular when computing the eigenvalues and solving linear systems. In general, an ε-perturbation in the inputs would change the solution by $O(\kappa\|\varepsilon\|)$, where κ is the so-called condition number [13, section 1.6] of the problem. A complete accuracy analysis would ideally account for all the numerical errors and conditioning that accrue throughout the process. This is far beyond the scope of this work. Instead of attempting a complete conditioning analysis for GCDT here we discuss a practical means to estimate the accuracy of a solution computed by our algorithm. There are several sources of errors: (i) error in computing the eigenvalue λ_2, (ii) error in computing λ_1 after λ_2 is computed, and (iii) forming x.

The error in the computed eigenvalue can be estimated as follows. Let $(\hat{\lambda}, \hat{x}, \hat{y})$ be an eigenpair of a generalized eigenvalue problem $Ax = \lambda Bx$, $y^\top A = \lambda y^\top B$. The absolute condition number of the eigenvalue λ is defined as

$$\kappa_\lambda = \lim_{\varepsilon \to 0} \sup_{\|\Delta A\| \le \varepsilon} \frac{\min_i |\lambda - \lambda_i(A + \Delta A, B)|}{\varepsilon}, \tag{6.2}$$

where $\lambda_i(A + \Delta A, B)$ denotes an eigenvalue of the pencil $(A + \Delta A) - \lambda B$. In (6.2) we perturb only A as this is all we need to bound the error in the computed eigenvalue below. It follows immediately from the proof of [28, Thm. 5] that $\kappa_\lambda = \|y\|\|x\|/|y^\top Bx|$. If the computed eigenpair $\hat{\lambda}, \hat{x}$ has residual $r = A\hat{x} - \hat{\lambda}B\hat{x} \neq 0$, then by taking $E = -r\hat{x}^\top/\|\hat{x}\|^2$ we have $(A + E)\hat{x} = \hat{\lambda}B\hat{x}$, so $\hat{\lambda}$ is an exact eigenvalue of the perturbed pencil $(A + E) - \lambda B$. Combining the fact $\|E\| = \|r\|/\|x\|$ with the condition number $\|y\|\|x\|/|y^\top Bx|$, we see that the error $|\hat{\lambda} - \lambda|$ can be bounded as (to first order in $\|r\|$)

$$|\hat{\lambda} - \lambda| \lesssim \kappa_\lambda \|E\| = \frac{\|y\|\|A\hat{x} - \hat{\lambda}B\hat{x}\|}{|y^\top Bx|}. \tag{6.3}$$

Thus the error in the computed $\hat{\lambda}_2$ and $\hat{\lambda}_1$ can be estimated by (6.3) using the computed eigenvectors $\hat{x}, \hat{y}$ along with the eigenvalues. Such estimates can be used for determining which λ_1 are shared in step 4 of Algorithm 3.1: it is reasonable to regard those that lie within the error estimates as equal. Similarly, the estimates can be used

when we reject the negative values of λ_2, λ_1: if the computed values are negative but with absolute value within the error estimates, they should be set to 0 and kept as candidates.

We next discuss estimating the error in the computed x. Since x is generically obtained from the linear system $H(\lambda_1, \lambda_2)x = y$, we examine the condition number η of the matrix $H(\hat{\lambda}_1, \hat{\lambda}_2)$. The error $\Delta x = \hat{x} - x$ is then estimated as $\|\Delta x\| \lesssim \eta\|\hat{x}\|$ [13, section 7.1]. This would give a rough estimate for the solution accuracy; strictly speaking, we need to account also for the errors in $\hat{\lambda}_1, \hat{\lambda}_2$, but this would be challenging as we would then need to find an upper bound for the condition number of $H(\lambda_1, \lambda_2)$, where λ_1, λ_2 are allowed to move within the error estimates. Once such an error estimate for $\hat{x}$ is obtained, the error in the objective value can be estimated, for example, as $|f(\hat{x}) - f(x)| \lesssim 2\|\Delta x\|\|Q_0\hat{x}\| + \|Q_0\|\|\Delta x\|^2 + 2\|q_0\|\|\Delta x\|$, though such bounds tend to be overestimates.

Using a fixed machine precision u, the outcome of our algorithm is expected to have accuracy $O(\tau\eta u)$, where τ is the conditioning of the two-parameter eigenvalue problem, if the algorithm does not require perturbation. In the very rare case that it requires perturbation, the computed solution may have additional error (which is discussed in section 4.3) of magnitude $O(\kappa u)$, where κ is the overall condition number of the GCDT problem; the product of conditional numbers at each step, including κ_λ of all perturbed generalized eigenvalue problems and η. However, since these condition numbers are independent of the precision u, we can use higher precision arithmetic to improve the accuracy: by using 10 more digits, we expect to improve the accuracy of the outcome roughly by 10 digits.

In practice, however, using higher precision often involves prohibitively increased computational effort, and the discussion here may be only of theoretical interest.

7. 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 16 GB RAM, in which unit roundoff $u \approx 1.1 \times 10^{-16}$. We solved SDP by SeDuMi 1.3.

As mentioned previously, in our algorithm derivation we assumed exact arithmetic together with exact eigenvalue computation, and hence the performance in finite precision arithmetic needs to be examined. As we see later, all instances in the following experiments are solved by our algorithm, indicating it is in practice a reliable global optimization algorithm for GCDT.

Note that all instances in the following experiments did not violate the LICQ. Thus the KKT conditions always sufficed, and we did not need to deal with the Karush–John conditions described in section 5; neither was the perturbation process in section 4 needed.

7.1. Runtime analysis of our algorithm.

We generated random instances of GCDT for $n = 5, 10, \ldots, 40$ and examined the runtime of our algorithm. The random instances are generated as in Burer and Anstreicher [8] and they are also used in section 7.2.2. 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.

The dominant cost of our algorithm is solving a linear generalized eigenvalue problem

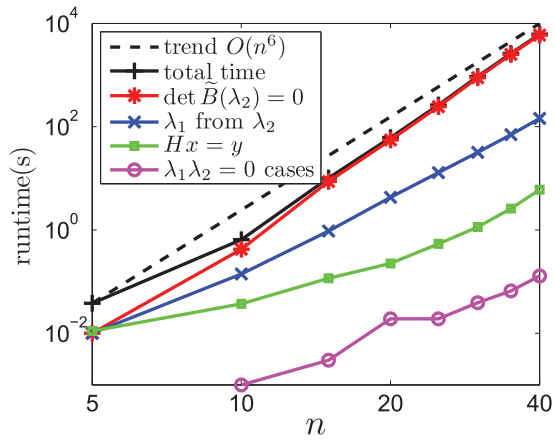

FIG. 1. *Double-logarithmic graph of runtime for varying n. We name the computational time of four parts shown above as "$\det \widetilde{B}(\lambda_2) = 0$," "$\lambda_1$ from λ_2," "$Hx = y$," and "$\lambda_1\lambda_2 = 0$ cases" in the legend in order of appearance.*

$\det \widetilde{B}(\lambda_2) = 0$, 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 1 illustrates that the computational time spent for the whole algorithm and solving $\det \widetilde{B}(\lambda_2) = 0$ scales asymptotically as $O(n^6)$.

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

$$\begin{aligned} &\underset{x}{\text{minimize}} && Q_0 \bullet X + 2q_0^\top x + \gamma_0 \\ &\text{subject to} && Q_i \bullet X + 2q_i^\top x + \gamma_i \leq 0 \quad (i = 1, 2), \\ & && X \succeq xx^\top. \end{aligned} \tag{7.1}$$

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 [2].

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 6.1.

7.2.1. 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 [8]:

$$\begin{aligned} &\underset{x}{\text{minimize}} && x^\top \begin{bmatrix} -4 & 1 \\ 1 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top x \\ &\text{subject to} && \|x\|_2^2 \leq 1, \quad x^\top \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} x \leq 2. \end{aligned} \tag{7.2}$$

We illustrate the objective function value and the feasible region for this problem in

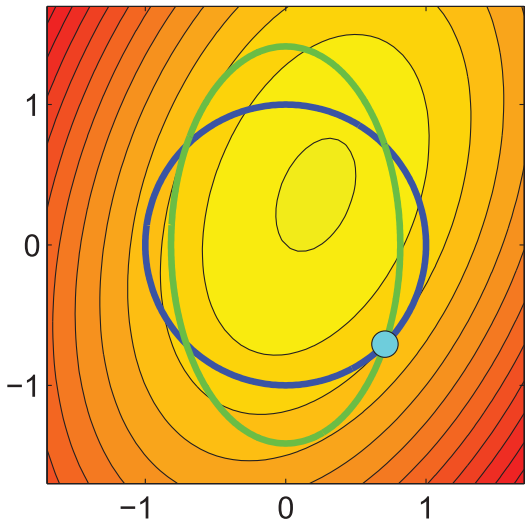

FIG. 2. *Two-dimensional instance from Burer and Anstreicher and its global solution.*

Figure 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 [8] 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.

7.2.2. Random instances from Burer and Anstreicher. In [8], 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 them by our algorithm and the SDP relaxation.

In these instances the eigenvalues are computed with sufficient accuracy, so our algorithm finds the global solutions in all instances and the computed results appear to reflect this. However, to our knowledge, no effective way is available for guaranteeing global optimality given a purported solution. Therefore, we instead check whether solutions obtained by our algorithm satisfy the necessary condition for global optimality as follows. As Yuan [30] proved, the Hessian of Lagrangian $H(\lambda_1, \lambda_2)$ has at most one negative eigenvalue at a global solution of the CDT problem. That is, denoting the number of negative eigenvalues of $H(\lambda_1, \lambda_2)$ at a global solution by $\psi(H(\lambda_1, \lambda_2))$, $\psi(H(\lambda_1, \lambda_2)) \in \{0, 1\}$ is necessary for global optimality for any CDT problem. We confirmed this condition is always satisfied for the solutions computed by our algorithm. Furthermore, as a naive test, we took 10^6 randomly sampled feasible points and compared the objective values with our solution and verified that the computed solution gives the smallest value, indicating the global optimality of our solutions.

Now we consider solving these instances by the SDP relaxation. Indeed, as shown in [2], the SDP relaxation solved the CDT problem if and only if $\psi(H(\lambda_1, \lambda_2)) = 0$.

[1]More precisely, two TRS (TTRS) as Burer and Anstreicher call it in [8]. TTRS is a GCDT with $Q_2 \succ O$, which is a special case of CDT.

TABLE 1
Number of instances solved via the SDP relaxation.

n	CDT	indefinite Q_2
2	99	81
5	90	52
10	90	46
20	92	44

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] The number of CDT instances solved by the SDP relaxation method is shown in Table 1.

7.2.3. Random instances with indefinite Q_2. Based on the idea Burer and Anstreicher [8] present, we generate random instances of GCDT 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 GCDT 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 the 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 GCDT.

For $n = 2, 5, 10, 20$, we generated 100 such instances and solved them by our algorithm and the SDP relaxation. As in section 7.2.2, we confirmed that the objective values at solutions computed by our algorithm are smaller than those at 10^6 randomly sampled feasible points. The result of the SDP relaxation method is shown in Table 1. Compared with the result for CDT instances, 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 GCDT with indefinite Q_2.

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 an 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

[2] In the experiments of Burer and Anstreicher [8], 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 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 7.2.

process in Remark 6.1 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 $= 10^{-8}$).

8. Conclusion and discussion. We have developed an algorithm for finding a global solution of GCDT. Our algorithm solves GCDT 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. We also showed how to find local solutions that violate the LICQ via the Karush–John optimality conditions, thus solving any GCDT problem. In finite precision arithmetic with unit roundoff u, the computational complexity of our algorithm is shown to be $O(n^6 \log\log u^{-1})$ in total. Numerical experiments are conducted to illustrate the runtime of our algorithm and to compare the outcome with the SDP relaxation method. Although we developed the algorithm assuming exact arithmetic including exact eigenvalue computation, all instances in the experiments are solved by our algorithm, indicating it is a reliable global optimization algorithm in practice.

We now remark on possible future work. First, the positive definiteness requirement of Q_1 is not necessary for most of the discussions, and indeed Q_1 being nonsingular would suffice for all the derivations, except for Lemma 3.2, section 4.3, and the discussions following (4.4) and (B.7), all of which treat nongeneric cases. We therefore suspect that with some further analysis and modifications, our approach can be shown to solve a general quadratically constrained quadratic program (QCQP) with two constraints.

Another issue is the $O(n^6 \log\log u^{-1})$ complexity, which 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 the original GCDT 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):

$$\text{minimize}_{x} \quad f(x) = x^\top Q_0 x + 2q_0^\top x + \gamma_0 \tag{8.1}$$

$$\text{subject to} \quad g_i(x) = x^\top Q_i x + 2q_i^\top x + \gamma_i \le 0 \quad (i = 1, 2, \ldots, m). \tag{8.2}$$

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

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 the objective function is bounded from below as follows:

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

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. 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 B.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 B.2.

B.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 $g_1(x_* + H_0 v) \leq 0$ and $g_2(x_* + H_0 v) \leq 0$, which can be obtained by solving the following TRS for v:

$$
\text{(B.1)} \qquad \begin{array}{ll} \underset{v}{\text{minimize}} & g_2(x_* + H_0 v) \\ \text{subject to} & g_1(x_* + H_0 v) \leq 0. \end{array}
$$

   This problem can be solved via an SDP reformulation (see [27]). If the optimal value of (B.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 $g_2(x_* + H_0 v) \leq 0$, which we do by solving the following minimization problem for v:

$$
\text{(B.2)} \qquad \begin{array}{ll} \underset{v}{\text{minimize}} & g_2(x_* + H_0 v) \\ \text{subject to} & g_1(x_* + H_0 v) = 0. \end{array}
$$

   We show how to solve this quadratic minimization problem with one quadratic equality constraint in Appendix B.2. If the optimal value of (B.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

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

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:

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

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

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

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

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

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 B.2:

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

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

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 (B.6) and (B.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:

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

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

$\mathcal{E}$ 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 $\alpha_1, \alpha_2 \in \mathbb{R}$ satisfying

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

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

So we obtain a vector $v = \alpha_1 v_1 + \alpha_2 v_2$ satisfying (B.6) and (B.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, 21]).

B.2. The quadratic minimization problem with one quadratic equality constraint. In Appendix B.1, we need to solve the following minimization problems (B.2) and (B.3), where $g_1(x_* + H_0 v)$ and $g_2(x_* + H_0 v)$ are defined by the left side of (B.4) and (B.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 (B.2) and (B.3) as follows:

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

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

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

$$\text{(B.16)} \qquad (A_1 + \lambda A_2)v = -(b_1 + \lambda b_2),$$

$$\text{(B.17)} \qquad v^\top A_2 v + 2b_2^\top v + c_2 = 0.$$

Note that, in one constraint minimization problem (B.14) and (B.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 (B.16) and (B.17), where

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

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 γ_i in section 4.1, we perturb 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 (B.17) naturally. If $A(\lambda)$ is singular, we find one of the solutions v satisfying (B.16), (B.17) as follows: Let v_* be the minimum-norm solution of (B.16) and v_0 be an arbitrary null vector of $A(\lambda)$ such that $\|v_0\|_2 = 1$. We see that the solutions of (B.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 (B.17) and solve for t to obtain a solution satisfying (B.16) and (B.17). If no real solution t is obtained, the corresponding λ gives no feasible solution of (B.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.

Acknowledgment. We thank the referees for their helpful suggestions to improve the scope and clarity of the paper.

REFERENCES

[1] S. Adachi, S. Iwata, Y. Nakatsukasa, and A. Takeda, *Solving the Trust Region Subproblem by a Generalized Eigenvalue Problem*, Technical report, METR 2015-14, University of Tokyo, Tokyo, 2015.

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

[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*, Ph.D. thesis, Pierres, Paris, 1779.
[5] D. Bienstock, *A note on polynomial solvability of the CDT problem*, SIAM J. Optim., 26 (2016), pp. 488–498.
[6] I. M. Bomze and M. L. Overton, *Narrowing the difficulty gap for the Celis-Dennis-Tapia problem*, Math. Program. Ser. B, 151 (2015), pp. 459–476.
[7] P. A. Browne, *Numerical Methods for Two Parameter Eigenvalue Problems*, Ph.D. thesis, Department of Mathematical Sciences, University of Bath, Bath, England, 2008.
[8] S. Burer and K. M. Anstreicher, *Second-order-cone constraints for extended trust-region subproblems*, SIAM J. Optim., 23 (2013), pp. 432–451.
[9] 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.
[10] X. Chen and Y.-X. Yuan, *On local solutions of the Celis–Dennis–Tapia subproblem*, SIAM J. Optim., 10 (2000), pp. 359–383.
[11] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Classics App. Math. 58, SIAM, Philadelphia, 2009.
[12] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., The Johns Hopkins University Press, Baltimore, MD, 2012.
[13] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
[14] S. Iwata, Y. Nakatsukasa, and A. Takeda, *Computing the signed distance between overlapping ellipsoids*, SIAM J. Optim., 25 (2015), pp. 2359–2384.
[15] F. John, *Extremum problems with inequalities as subsidiary conditions*, in Traces and Emergence of Nonlinear Programming, G. Giorgi and T. Hoff Kjeldsen, eds., Springer, Basel, 2014, pp. 197–215.
[16] F. C. Kirwan, *Complex Algebraic Curves*, Cambridge University Press, Cambridge, 1992.
[17] L. Lerer and M. Tismenetsky, *The Bézoutian and the eigenvalue-separation problem for matrix polynomials*, Integral Equations Operator Theory, 5 (1982), pp. 386–445.
[18] G. Li and Y. Yuan, *Compute a Celis-Dennis-Tapia step*, J. Comput. Math., 23 (2005), pp. 463–478.
[19] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 17 (1967), pp. 37–47.
[20] C. B. Moler and G. W. Stewart, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256.
[21] 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.
[22] Y. Nesterov, H. Wolkowicz, and Y. Ye, *Noconvex quadratic optimization*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer, New York, 2000, pp. 361–420.
[23] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer, New York, 1999.
[24] 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.
[25] M. J. D. Powell and Y. Yuan, *A trust-region algorithm for equality constrained optimization*, Math. Program., 49 (1991), pp. 189–211.
[26] R. J. Stern and H. Wolkowicz, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.
[27] J. F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, Math. Oper. Res., 28 (2003), pp. 246–267.
[28] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.
[29] B. Yang and S. Burer, *A two-variable approach to the two-trust-region subproblem*, SIAM J. Optim., 26 (2016), pp. 661–680.
[30] Y. Yuan, *On a subproblem of trust region algorithms for constrained optimization*, Math. Program., 47 (1990), pp. 53–63.